Towards PNA Directed Chemical Total Protein Synthesis

Barry Cottam-Howarth

Submitted in accordance with the requirements of the degree of Ph.D

The University of Leeds
Department of Chemistry
September 2007

The candidate confirms that the work submitted is his own and that the appropriate credit has been given where reference has been made to the work of others.

Acknowledgements

I would first like to thank my supervisor Dr Stuart Warriner whose boundless enthusiasm and untiring support have bridged the various chasms and pitfalls I have encountered over the course of my PhD. Professor Adam Nelson, my second advisor, has also been a great source of inspiration and motivation during this time. I would also like to thank Professor Peter Johnson for the work, effort and assistance with SPROUT that made this project possible.

I would also like to thank the technical staff for all their invaluable help and efforts during my time at Leeds. In particular I would especially like to thank Jim Titchmarsh (HPLC), Tanya Marinko-Covell (MS) and Simon Barrett (NMR) whom have all been exceptional. I am also very grateful for the funding supplied by the EPSRC.

In compiling this thesis I would like to recognise and thank Dr Daniel Morton and Dr Anil Nagalingam for their exemplary proof reading services.

A PhD is an intense and highly stressful endeavour, and would be unsurmountable if not for the friendship and support of ones colleagues and peers. Special thanks therefore must be given to Anil Nagalingam, Daniel Morton, Steve Bartlett, Chris Cordier and Jim Tunstall. I would also like to thank members of the Warriner, Nelson and Fishwick groups, both past and present.

A special mention must be made to Chris Cordier whom over the course of the last three years has gone out of his way on occasions too numerous to mention to provide assistance and friendship when it most mattered.

Finally, I would like to thank all of my friends and family for their support and love. In particular, I would like to thank Hélène Duval for her endless love, patience and support, and it is to her whom I dedicate this thesis.

Abstract

This thesis is concerned with the development of a novel methodology for the chemical synthesis of proteins. The methodology is expected to overcome limitations of current techniques for chemical synthesis of proteins that restrict the size of the target molecule to less than 150 amino acids. It is intended that the use of a DNA analogue, PNA, as a directing group will allow different short polypeptides to self-assemble into a desired target protein.

The first chapter is an overview of literature concerning templated synthesis and the applications of PNA. Examples of simple metal-templated synthesis are discussed however there is a particular focus on the work of David Liu and his DNA-templated synthesis. The uses of PNA in a number of diverse applications are presented along with brief comparative descriptions of PNA and DNA. A brief discussion of current methodologies for protein synthesis is then used to set the context for the project aims and a description of the proposed methodology.

The second chapter details the effort made in modifying literature procedures to successfully develop the robust and reproducible chemistry required for multi-gram synthesis of PNA monomers.

The third chapter briefly discusses the computer-aided design of a pair of linker molecules that act as a bridge between polypeptides and their PNA directing groups. The challenging synthesis of these linker molecules is detailed and synthetic routes for the multi-gram synthesis of both are given.

The fourth chapter charts the development of the solid-supported chemistry used to construct PNA-peptide chimeras with close attention to the challenges encountered with incorporating the linker molecules. Of the two chimera targets required to examine PNA-directed ligation, a multi-gram preparative synthetic route is presented for one chimera. A route that provides analytical quantities of the complementary chimera is also discussed.

The fifth chapter presents the experimental procedures used for the synthesis of the compounds involved in this project. Analytical data and the characterisation of compounds are detailed for each compound where possible.

Contents

Glossary

A	Adenine
Ac	Acetate
AEG	Amino-Ethyl Glycine
AIDS	Acquired Immuno-Deficiency Syndrome
ALET	Auto-Ligation Energy Transfer
Bhoc	Benzhydryloxycarbonyl
Bn	Benzyl
Boc	*tert*-ButoxyCarbonyl
BR	Borromean Rings
bs	Broad singlet
Bu	Butyl
C	Cytosine
CBz	CarboxyBenzyl
CD	Circular Dichroism
CMV	CytoMegaloVirus
Cyt	Cytochrome
d	Doublet
DABCO	Diazabicyclo[2,2,2]-octane
DBU	Diazabicyclo[4,5,0]-undec-7-ene
DEPT	Distortionless enhancement by polarisation transfer
DIPEA	Diisopropylethylamine
DMA	*N,N*-Dimethylacetamide
DMAP	*N,N*-Dimethylaminopyridine
DMF	*N,N*-Dimethylformamide
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dsDNA	Double-Stranded Deoxyribonucleic acid
DTS	DNA Templated Synthesis
E	End
EDC	1-Ethyl-3-(dimethylaminopropyl)carbodiimide hydrochloride salt
Et	Ethyl
FISH	Fluorescence In-Situ Hybridisation
fM	femto molar

Fmoc	Fluorenylmethoxycarbonyl
FmocOSu	*N*-(9-Fluorenylmethoxycarbonyloxy)succinimide
FRET	Fluorescence Resonance Energy Transfer
G	Guanine
Gln	Glutamine
Gly	Glycine
H	Hairpin
HATU	N-{(Dimethylamino)-1H-1,2,3-triazolo[4,5-b]-pyridino-1-ylmethylene}-N-methylmethan-aminium hexafluorophosphate
HCTU	2-(6-Chloro-1-H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HF	Hydrofluoric acid
His	Histidine
HIV	Human immunodeficiency Virus
HMBC	Heteronuclear multiple bond correlation
HMQC	Heteronuclear multiple quantum coherence
HOBt	1-Hydroxy-7-benzotriazole
HPLC	High Performance Liquid Chromatography
IBPRS	Iterated Branching Reaction Pathway Synthesis
IIDQ	2-Isobutoxy-1-isobutoxycarbonyl-1,2-dihydroquinoline
LC-MS	Liquid chromatography-Mass spectroscopy
Lys	Lysine
m	Multiplet
m.p.	Melting point
MB	Molecular Beacon
Me	Methyl
Met	Methionine
mRNA	Messenger ribonucleic acid
MS	Mass spectrometry
MU	Mutant
NCL	Native Chemical Ligation
NMM	*N*-methylmorpholine
NMO	*N*-Methyl morpholine*N*-oxide
NMR	Nuclear Magnetic Resonance
Pbf	2,2,4,6,7-Pentamethyl-dihyrobenzofurane-5-sulfonyl
PCR	Polymerase Chain Reaction
Ph	Phenyl
PNA	Peptide Nucleic Acid

q	Quartet
RCM	Ring Closing Metathesis
R_F	Retention factor
RNA	Ribonucleic acid
RT	Room Temperature
s	Singlet
SelenoMet	SelenoMethionine
SPPS	Solid Phase Peptide Synthesis
ssDNA	Single-Strand DeoxyriboNucleic Acid
T	Thymine
t	triplet
TEBAB	Triethylbenzylammonium bromide
TEBAC	Triethylbenzylammonium chloride
TES	Triethylsilane
TFA	Trifluoroacetic acid
TFE	2,2,2-Trifluoroethanol
THF	Tetrahydrofuran
Thr	Threonine
TIC	Total Ion Count
TIS	Triisopropylsilane
T_m	Melting Temperature for DNA/DNA, DNA/PNA or PNA/PNA duplex denaturation (sic)
TMS	Trimethylsilyl
TMSCl	Trimethylsilylchloride
tRNA	Transfer RiboNucleic Acid
Trt	Trityl
Tyr	Tyrosine
UV	Ultra-violet
WT	Wild-Type

Introduction

Chapter 1

1 Peptide Nucleic Acid Directed Chemical Total Protein Synthesis

The importance of research into the understanding of protein chemistry and biology cannot be understated. A massive multinational research effort into decoding the human genome is underway to further our understanding of how life works.[1] The benefits that can be reaped from expanding and applying this knowledge to areas such as medicinal care are profound. The modern chemist is central to these efforts - organic synthesis can offer a path toward novel and interesting biomolecules that are inaccessible by other means. Designer peptides and proteins are one set of tools that can unravel the information presented by nature. The ability to make discrete changes to polypeptides, such as introducing a specific-point mutation or a fluorescent reporter group in an enzyme, provides a powerful way to interrogate these fascinating biomolecules and decipher how they work. There is no methodology that can fully address the synthesis of designer proteins, but we have sought to learn from nature's templated approach to peptide synthesis and thus provide facile access to designer peptides. This project has been concerned with developing a novel approach to overcome some of the limitations of chemical total protein synthesis that currently prevent the synthesis of large protein targets. It is the intent of the project to use templated reactions to organise and assemble smaller polypeptides into proteins much larger than those accessible by current technology. To achieve this goal oligomeric strands of nucleic acids will be used as directing groups that will orchestrate templated synthesis. In this first chapter, a review of some relevant literature is presented that will help set the background to the theory of the project. Initially the general concepts of templated synthesis are discussed with some representative examples that illustrate the level of design and control that can be achieved with this approach to synthesis. The review then continues with some elegant work by Liu that is closer in context with the aims of this project, DNA-templated organic synthesis, and highlights the power of nucleic acids as directing groups for chemical reactions. Next the surveyed literature centres upon Peptide Nucleic Acid (PNA), which provides the template element for this project, and current applications of PNA are presented to demonstrate its utility in chemistry. Finally, a contrasting reminder of how nature effortlessly makes proteins is given to complete the review, and this leads into an explanation of the general theory and idea behind this project.

1.1 Templated Synthesis

Exerting control over how molecules interact and react together becomes increasingly difficult as their size and complexity increase. In the mass and size domain of natural product synthesis, this can be a considerable challenge. The outcome of a reaction between large molecules is often dictated by large entropic molecular energies that lead to poor and disabling reaction kinetics. Consequently, there is a need for methodology that can control the regiochemistry and rate of macromolecular reactions.

An approach used in both nature[2, 3] and the lab[4, 5] to overcome such difficulties is to preorganise reactants upon a template (Figure 1.1). Molecules can be anchored *via* intermolecular forces or covalent tether to a scaffold and brought to a sustained close proximity. The template arranges the reactive centres into a desired configuration, thus reducing the activation entropy of the reaction and increasing the rate of the desired reaction. By ordering the components in this way, control can be exerted over the course of the reaction pathway.

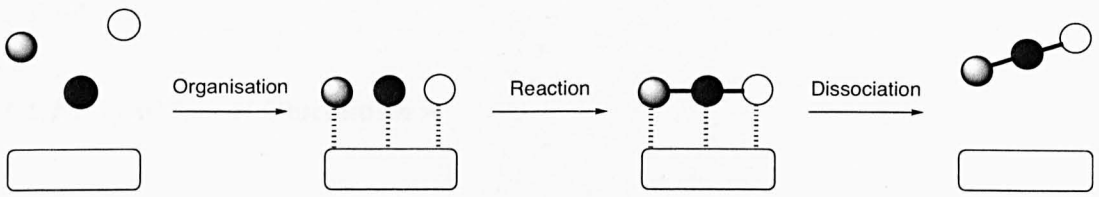


Figure 1.1 Schematic representation of templated synthesis

1.1.1 Metal-mediated templated reactions

1.1.1.1 Crown ether synthesis

A classic example of metal-templated chemistry is the use of potassium salts in the synthesis of crown ethers[4, 6] (Figure 1.2):

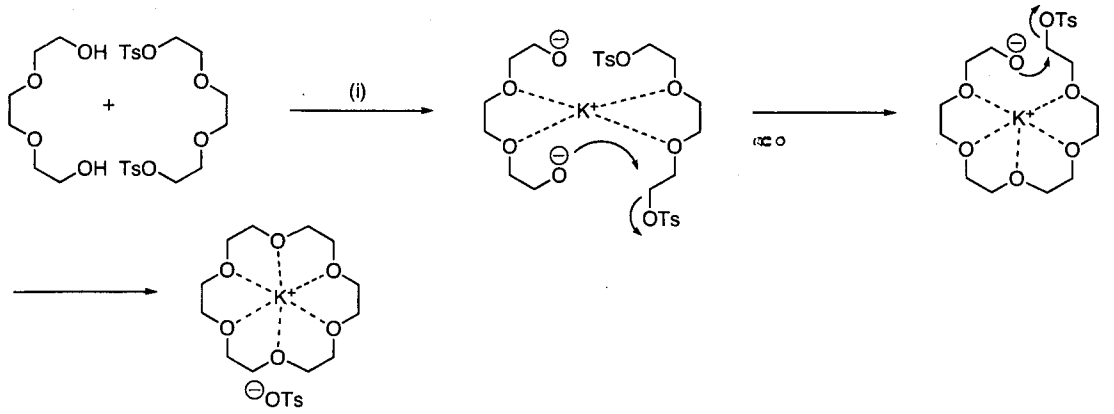
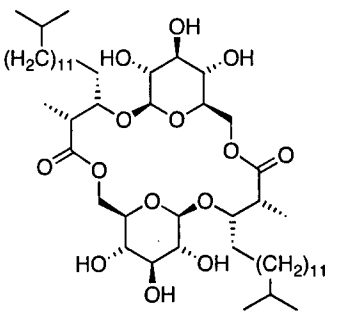


Figure 1.2 Potassium templated synthesis of 18-crown-6

The potassium cation draws the reactants together by coordination to the oxygen atoms, reducing the degrees of freedom the molecules possess. The reacting centres are now held in close proximity, and crucially the *intra*molecular S_N2-like reaction is now favoured ahead of polymerisation *via inter*molecular S_N2. Without the potassium's preorganisation of the reactants, the reaction yields a mixture of linear polyethylene glycols.

1.1.1.2 Synthesis of Glucolipsin A

An impressive extension of this methodology was utilised by Fuerstner *et al.* in the synthesis of the macrodiolide **1,** Glucolipsin A (Figure 1.3).[7, 8]



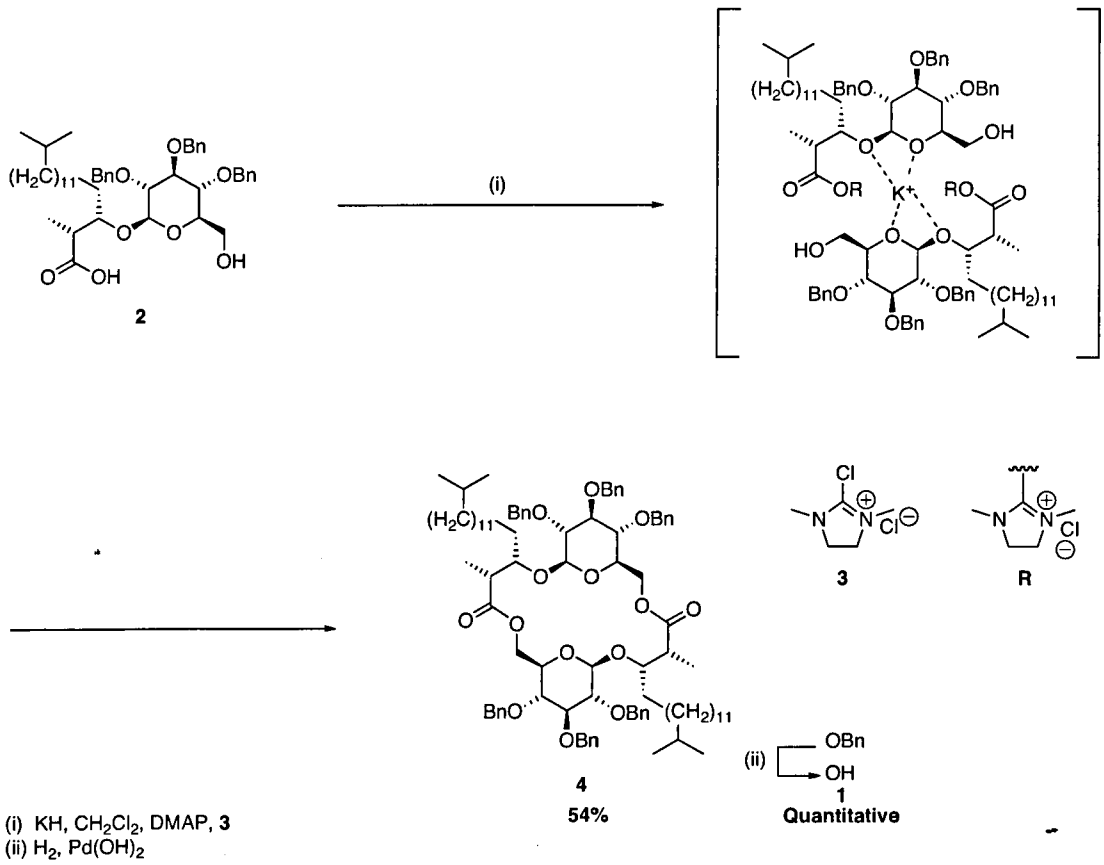
1

Figure 1.3 Glucolipsin A

An obvious disconnection of the core structure reveals it to be a dimer of carbohydrate origin. It was thought that synthesis of the elaborated sugar and subsequent dimerisation

would represent the most expediant route to the target molecule. Prior work suggested that the central dilactone core of the two structures would possess a significant level of ionophoric character; by implication the dimerisation reaction could be templated by a suitable metal cation, akin to the synthesis of crown ethers.

The fragments were synthesised and treated with potassium hydride, which acted with dual capacity to deprotonate the precursor **2** and act as the source of admixed templating cation. The acids were activated with 2-chloro-1,3-dimethylimidazolinium chloride **3**, and the protected form of Glucolipsin A **4** was formed as the major cyclisation product (Scheme 1.1). Quantitative reductive cleavage of the benzyl protecting groups completed the synthesis of Glucolipsin A.



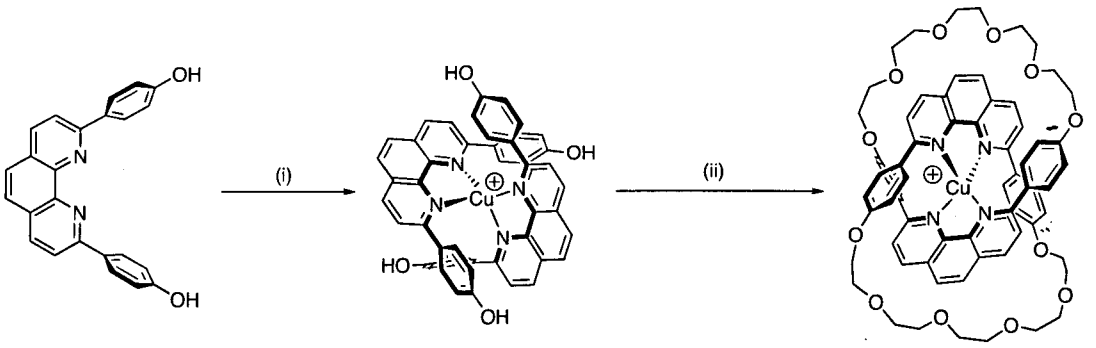
Scheme 1.1 Potassium templated dimerisation of hydroxy glycosyl acid **2**

This approach was also used successfully in the synthesis of some of Glucolipsin's stereoisomers and structurally related compounds such as Cycloviracin B1[9].

1.1.1.3 Catenand synthesis

The use of metals for organising ring forming events is more potent than the feats achieved by Fuerstner *et al.* would suggest. There is great interest in the development of large topologically interesting macromolecules within the field of molecular machinery. The supramolecular assembly of big molecules into structures with useful and controllable properties could one day provide the means to implement the promises of nanotechnology.[10-13] A simple example of a supramolecular structure is the catenand; a chain of two (or more) macrocycles, interlocked *via* the central cavities. Early efforts in this area using non-templated methods were synthetically unproductive, giving minimal amounts of product. The first methodology to give rapid access to useful quantities of an interlocked-ring substance was published by Sauvage *et al.* with the introduction of the Catenands.[14, 15]

It was realized that suitably functionalized organic ligands bound to a metal complex of well defined geometry could be elaborated such that they could encircle each other, forming two linked rings. Copper (I) was demonstrated to bind strongly enough to modified phenanthroline ligand **5** to allow such manipulation. The ligands occupy two coordination sites on the metal, and by virtue of the tetrahedral bonding geometry, are opposing and perpendicular to each other (**6**). The two ligands are converted to macrocycles by joining their common phenolic ends together with a diiodopolyglycol (Scheme 1.2). Cyclisation of one ligand traps its partner in the newly formed macrocyclic central cavity; cyclisation of both ligands therefore gives two linked ring structures (**7**). The copper cation can be extracted quantitatively from the product with tetramethyl ammonium cyanide.



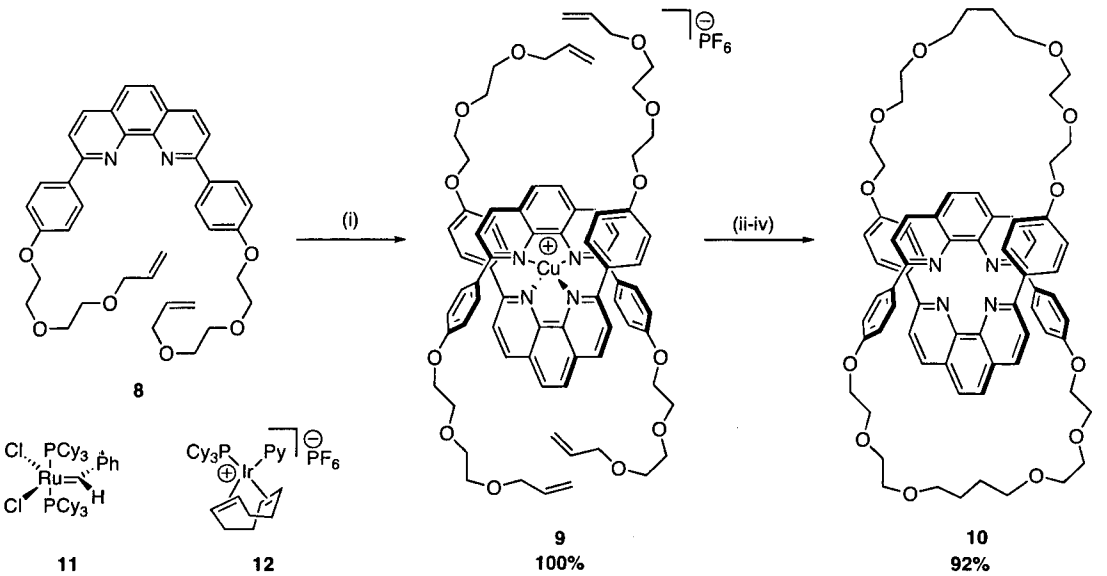
5 **6** **7**
19%

(i) $[Cu(NCCH_3)_4]Cl$
(ii) $ICH_2(CH_2OCH_2)_4CH_2I$

Scheme 1.2 Copper templated synthesis of [2] catenand **7**

This result demonstrated that template control could be used to preorganise substrates to give an otherwise unlikely outcome. The catenand strategy is however still subject to large unfavourable entropic energies during the final cyclisation event; the tethered but long polyglycol iodide can adopt many configurations that do not facilitate closure. This represents the major limitation with this technology.

The advent of Ring Closing Metathesis (RCM) allowed Sauvage *et al.* to revisit the catenand work and develop new methodology to overcome the difficult kinetics of previous attempts.[16] New phenathroline ligands (**8**) were synthesized that possessed phenolic cyclisation motifs and were precomplexed with copper (**9**). RCM of the alkene units and reduction of the cyclised product gave the catenand **10** in much improved yield (Scheme 1.3).

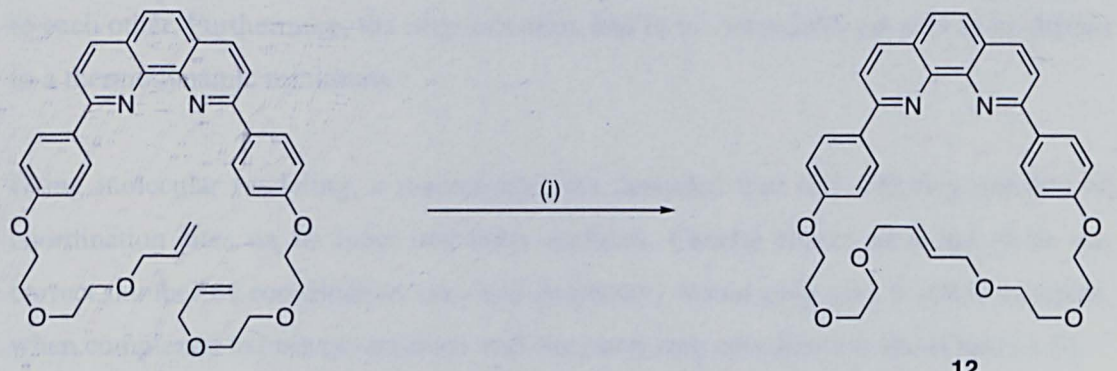


(i) $Cu(MeCN)_4PF_6$
(ii) **11**
(iii) **12**, H_2
(iv) KCN

Scheme 1.3 Improved [2] catenand synthesis

The use of two shorter tethered chains (as opposed to one long chain) reduced the entropy of the system, but the remarkably high efficiency of the cyclisation is attributed to preorganisation of the alkenes.[17] The locked conformation of the ligands is thought to reinforce interactions between the glycol oxygen atoms and the phenanthroline rings that result in reduced mobility of the phenolic arms. Consequentially, it is believed that the alkenes lie in closer proximity than expected, resulting in improved RCM efficiency.

The yields of RCM product **12** without the metal template (non-catenane macrocycle product) are markedly reduced, supporting this theory (Scheme 1.4).



(i) **11** **70%, *E* >> *Z***

Scheme 1.4 Untemplated RCM of modified 1,10 phenanthroline ligands

1.1.1.4 Borromean rings

The most complex structure to date synthesized by use of metal templated synthesis is the Borromean rings (BR) by Stoddart *et al.*[18-20] The BR are three interlocked rings in which scission of any ring separates all three from each other (Figure 1.4).

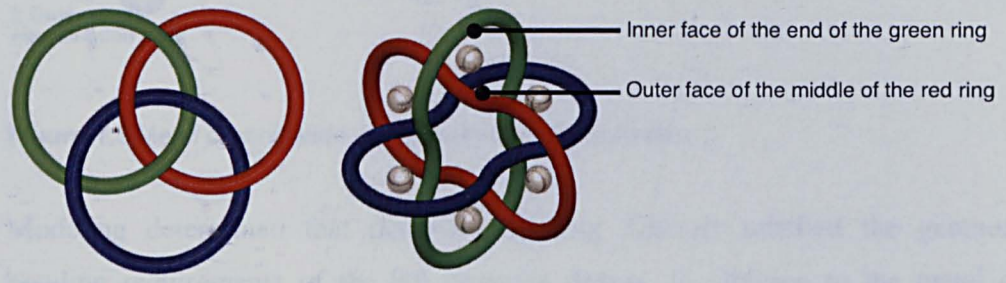


Figure 1.4 Borromean rings[20]

It was envisaged that three perpendicular ellipsoid rings could be constructed whereby the inner ends of one ring were complexed to the outer middle edges of the next ring, in a circular fashion (Figures 1.4, 1.5). To enable a good yield and facilitate BR formation, the complex was designed so that the assembly of the rings is reversible and that the BR are the thermodynamic product.

To achieve this it was necessary that the metal templated the required ligand assembly with extreme prejudice. A metal was needed that could discriminate between the inner middle of one ring and the outer ends of another, and orientate them to be perpendicular to each other. Furthermore, the ring formation had to be reversible, yet able to be driven to a thermodynamic minimum.

Using molecular modeling, a macrocycle was designed that had differing number of coordination sites on its inner and outer surfaces. Careful choice of metal (with the correct number of coordination sites and geometry) would only give a stable complex when complexed to both of one inner and one outer ring coordination site (Figure 1.5).

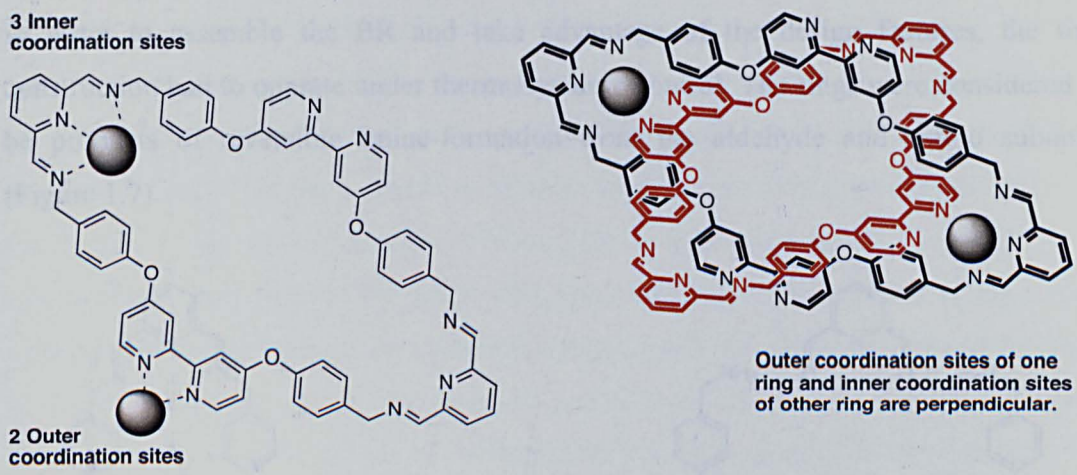


Figure 1.5 Metal organization of imine-derived macrocycles

Modeling determined that penta-coordinating Zinc(II) satisfied the geometric and bonding requirements of the BR complex design. In addition to the metal complex bonding, the BR were designed such that π-π stacking interactions could occur (twelve in total), further stabilizing the design to be the major thermodynamic product (Figure 1.6).

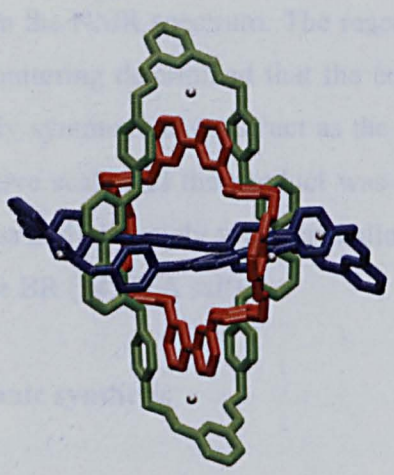


Figure 1.6 Computer modeled synthetic borromean rings

In order to assemble the BR and take advantage of the design features, the ring construction had to operate under thermodynamic control. The rings were considered to be products of reversible imine-formation from the aldehyde and amino subunits (Figure 1.7).

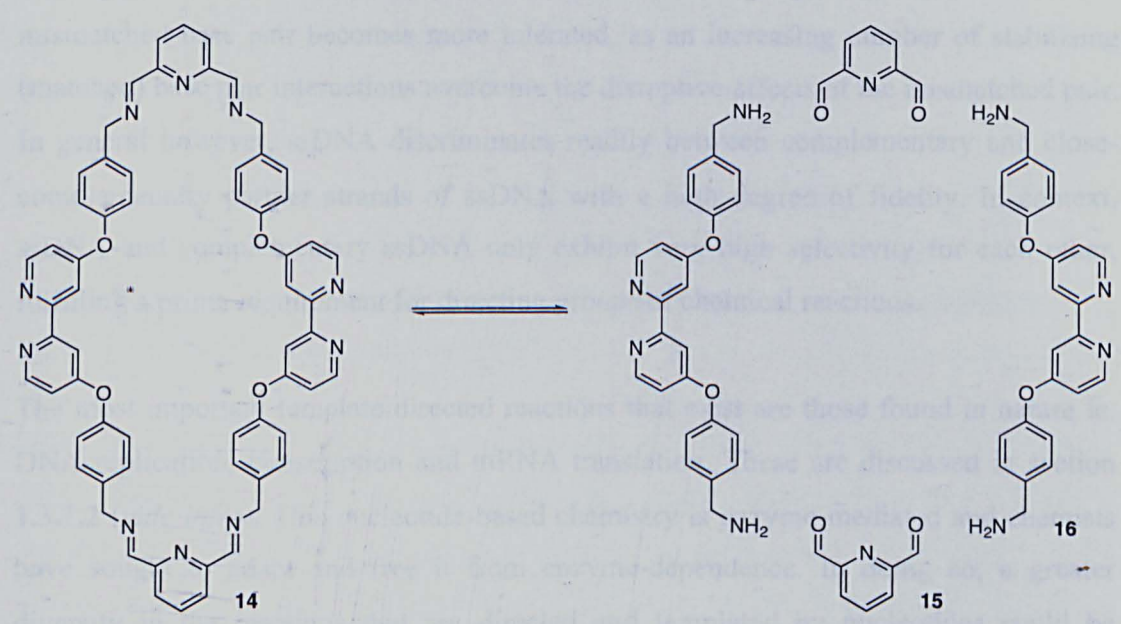


Figure 1.7 Reversible formation of macrocycle

When equimolar amounts of the amine (**16,** as TFA salt) and aldehyde (**15**) components were combined in deuterated methanol, no reaction was detected by NMR studies after several hours of stirring at room temperature. Addition of $Zn(OAc)_2$ immediately

caused a dramatic change in the NMR spectrum. The reaction was heated under reflux for three days; periodic monitoring determined that the equilibrium had been attained after 48 h, giving one highly symmetrical construct as the major product. The reaction was repeated on a preparative scale and the product was obtained in 90% yield after recrystallisation. NMR, mass and ultimately x-ray crystallographic analysis proved that the product obtained was the BR (**14**, TFA salt).

1.1.2. DNA templated organic synthesis

Helical DNA can be considered as an excellent directing group for templated synthesis.[21-29] The binding of two complementary DNA strands yields a predictable and programmable structure. ssDNA will hybridise with its complementary strand to form a stable duplex of dsDNA – however if one of the base pairs in the dsDNA is mismatched, then the resultant duplex is significantly destabilised (relative to a fully matched duplex). For short strands of ssDNA therefore, a stable double helix will only be formed with complementary ssDNA. With increasing length of dsDNA, a mismatched base pair becomes more tolerated, as an increasing number of stabilising (matched) base pair interactions overcome the disruptive affects of the mismatched pair. In general however, ssDNA discriminates readily between complementary and close-complementary partner strands of ssDNA with a high degree of fidelity. In context, ssDNA and complementary ssDNA only exhibit very high selectivity for each other, fulfilling a prime requirement for directing groups of chemical reactions.

The most important template directed reactions that exist are those found in nature ie. DNA replication, transcription and mRNA translation. These are discussed in section **1.3.1.2** (*vide infra*). This nucleotide-based chemistry is enzyme-mediated and chemists have sought to adapt and free it from enzyme-dependence. In doing so, a greater diversity in the reactions that are directed and templated by nucleotides could be achieved, given that enzymes work with a very limited set of substrates and environmental conditions. Seminal work by Orgel examined the parameters for self-replication of simple oligonucleotides under possible prebiotic-like conditions (no enzymes).[30-35] This work followed a natural path toward more sophisticated systems involving unnatural nucleotide polymers and laid the foundation for the use of DNA as directing groups. Von Kiedrowski was also concerned with the prebiotic origins of

oligonucleotide replication and successfully demonstrated nucleic-acid directed replication of DNA analogues without enzymic control.[36-38] Lynn further developed the idea of using nucleotides as directing groups for organic synthesis and produced chemistry for the exponential self-amplification of DNA analogues [39-43] The important groundwork laid out by Orgel, Von Kiedrowski and Lynn paved the way for broader research into nucleotide-directed organic synthesis.

Later attempts to exploit DNA as a template for organic synthesis were initially concerned with the synthesis of unnatural nucleotide polymers *via* modifications of the phosphate backbone. A leading example is the work of Kool *et al.* whom ventured to improve the detection of point mutations (single base mismatch) in DNA.[44]

Established methodology utilised a DNA strand to template the ligation of two shorter oligonucleotides that together comprise the complementary strand. This enzyme (DNA-ligase) mediated reaction can be studied using Fluorescence Resonance Energy Transfer[45-49] (FRET, Figure 1.8) labelling to determine if the strands have been united. FRET is the non-radiative transfer of energy from an excited donor dye to a relaxed acceptor dye.

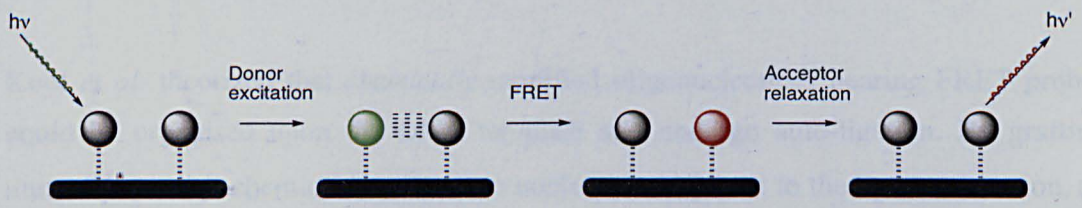


Figure 1.8 Schematic representation of FRET

FRET is initiated by the light-induced excitation of a donor dye molecule. If the FRET donor is isolated, it relaxes by emitting light at a slightly reduced frequency relative to that it absorbed. If an acceptor dye is located within close permanent proximity of the excited donor dye, then FRET can occur. The acceptor dye is excited by the donor molecule and subsequently relaxes with the emission of light. This light is of a much lower frequency (red-shifted) than that emitted by an isolated donor dye. Detection of donor or acceptor emitted light determines if the two probe dyes are held in close proximity or not.

The DNA ligase assay requires each of the templated oligonucleotides to bear a FRET probe near the residues adjacent to the ligation junction. If there is a point mutation present upon the template then the enzyme will not ligate the two DNA strands together, as the requirement for absolute sequence specificity has not been honoured. Without ligation the probes cannot be held in sustained proximity and FRET cannot occur; the donor dye undergoes radiative relaxation and a point mutation is identified in the template strand.

The dependence on enzyme mediation for the ligation process was considered a limitation of this technique for a number of reasons. Shorter oligodeoxynucleotides exhibit greater sequence specificity (point mutation sensitivity), yet ligase enzymes are rendered inefficient when joining oligodeoxynucleotides of less than nine bases in length. The technology cannot be extended to give direct detection of RNA point mutations as DNA ligases show low activities with RNAs. The technique cannot be utilized for *in vitro* studies as the ligase cannot be easily delivered into cellular structures.

1.1.2.1 Auto-ligation of nucleic acid derivatives

Kool *et al.* theorized that *chemically* modified oligonucleotides bearing FRET probes could be organised upon the DNA template and undergo auto-ligation. By grafting mutually reactive chemical handles onto nucleotides adjacent to the ligation junction, an unnatural backbone linkage would form upon templation. This would exclude the need for DNA ligase.

To test this hypothesis auto-ligation probes (a 7-mer and a 13-mer) were designed that would be complementary to a H-*ras* protooncogene template sequence. The sequence contains a characteristic codon-12 G → T mutation that has been well studied. Two 7-mer probes, one wild-type (WT) and one mutant strand (MU, complementary with G → T mutation) were constructed with different acceptor FRET probes(Figure 1.9).[50-52] The WT probe was labeled with yellow-emissive hexachlorofluoroscein (HEX) dye and the MU strand was labeled with red-emissive Rhodamine (ROX) dye. The 7-mer strands were designed so that the point mutation occurred midway along the oligonucleotide, to maximize selectivity of the MU strand. The 13-mer was equipped

with a 'universal' donor probe, the green-emissive 5-carboxyfluorescein (FAM) that can engage in FRET with either 7-mer. The 7-mers were backbone-modified to hold a nucleophilic 3'-Phosphorothioate group; The 13-mer was adapted to contain a 5'-Iodothymidine.

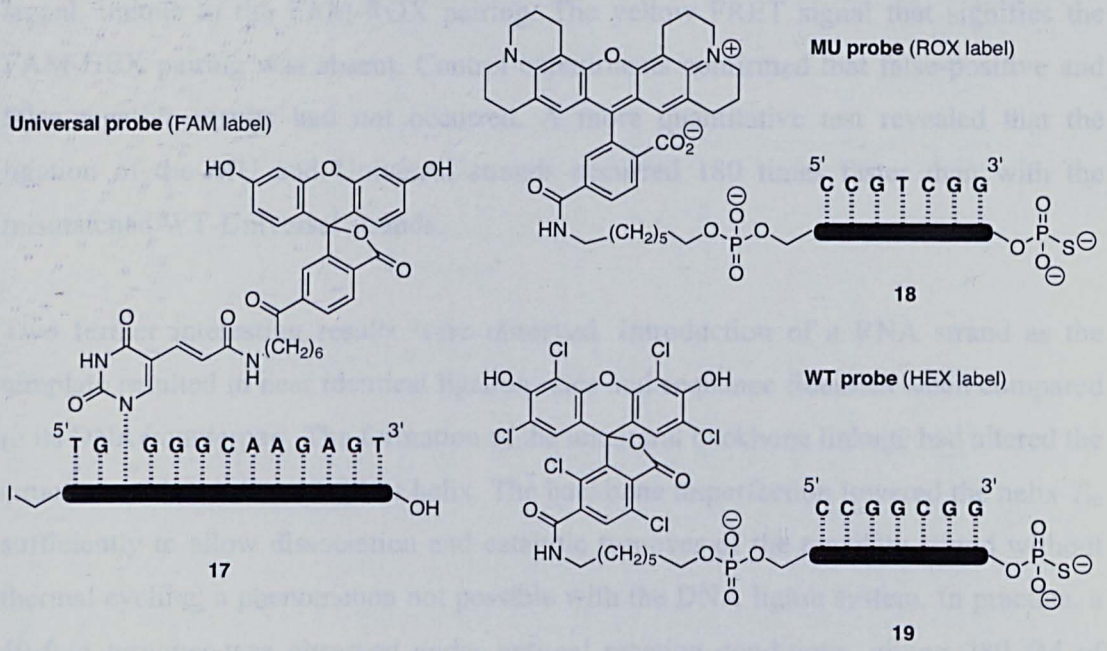


Figure 1.9 DNA tagged FRET probes

Conceptually, the H-*ras* 40-mer would template the ligation of only the MU 7-mer and the universal 13-mer in a mixture containing both 7-mers (Figure 1.10). Displacement of the iodide with the phosphorothioate would form the unnatural backbone linkage, enabling FRET to occur.

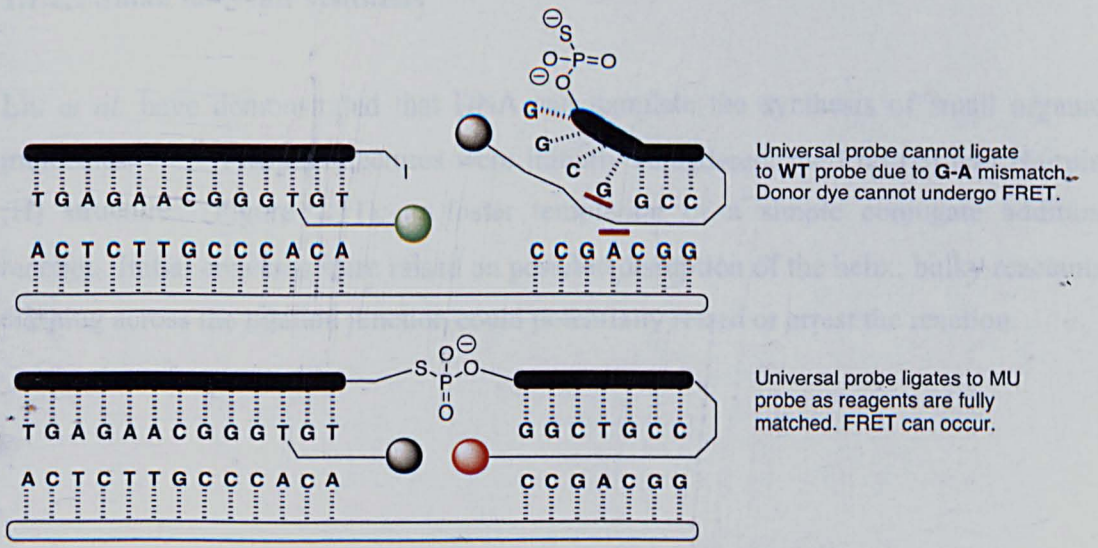


Figure 1.10 Auto-Ligation Energy Transfer experiment

It was found that the Auto-Ligation Energy Transfer (ALET) experiment was successful. A solution containing equimolar amounts of **18**, **19** and the universal DNA strand **17** was agitated and warmed at 32°C over 18h to give an intense red FRET signal, unique to the FAM-ROX pairing. The yellow FRET signal that signifies the FAM-HEX pairing was absent. Control experiments confirmed that false-positive and false-negative results had not occurred. A more quantitative test revealed that the ligation of the MU and Universal strands occurred 180 times faster than with the mismatched WT-Universal strands.

Two further interesting results were observed. Introduction of a RNA strand as the template resulted in near identical ligation rates and sequence fidelities when compared to its DNA counterpart. The formation of the unnatural backbone linkage had altered the structure of the mutant-template helix. The backbone imperfection lowered the helix T_m sufficiently to allow dissociation and catalytic turnover of the template strand without thermal cycling; a phenomenon not possible with the DNA ligase system. In practice, a 40-fold turnover was observed under optimal reaction conditions, giving 280 fM of ligated probes from 7 fM of template (per hour). The formation of unnatural backbone linkages between oligonucleotides had demonstrated the potential for other DNA-templated chemistries across the helix backbone.

1.1.2.2 Small molecule synthesis

Liu *et al.* have demonstrated that DNA can template the synthesis of small organic molecules.[23, 53] Two architectures were initially considered, the End (E) and Hairpin (H) structures (Figure 1.11), to foster templation of a simple conjugate addition reaction. Initial concerns were raised on possible disruption of the helix; bulky reactants clashing across the ligation junction could potentially retard or arrest the reaction.

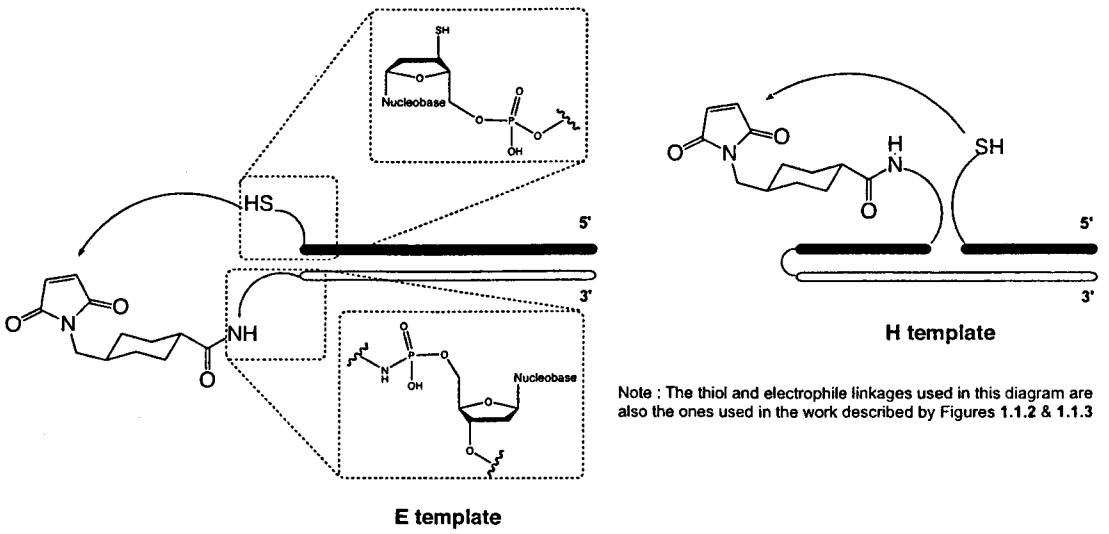


Figure 1.11 DNA templated conjugate addition of a thiol to a maleimide

The reaction was performed in pH 7.5 buffered solution at 25 °C with 60 nM concentration of all reagents. Within minutes, both E and H structures had complete turnover of the starting materials, giving the thiomaleimide as the sole product. It was found that both E and H architectures reacted at identical rates where $k_{app} \sim 10^5 M^{-1}s^{-1}$. The reaction was repeated unsuccessfully with template that had been pre-quenched with β-mercaptoethanol, proving the template thiol was the active nucleophile.

The successful proof of concept study for DNA Templated Synthesis (DTS) was expanded further to probe the scope of electrophile and nucleophile pairings (Figure 1.12).

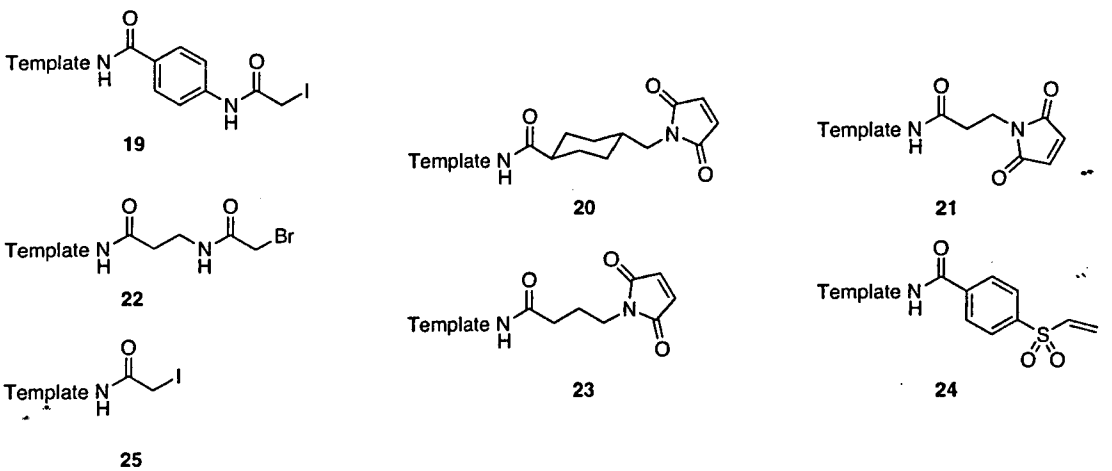


Figure 1.12 Electrophiles reacted *via* DTS with thiol and amine nucleophiles

On both H and E templates, all reagents **19 – 25** reacted in good yield and excellent sequence specificity with sulfur and nitrogen based template nucleophiles. No reaction was observed with mismatched reagents. These results are significant as the wide variance in reaction geometries required for all the possible reactions was tolerated without deleterious effects.

The selectivity of DTS was examined by studying the effects of introducing a single point mutation in the centre of the template for the reaction of **20** with the thiol nucleophile. Mismatched thiol reagents reacted ~ 200 times slower relative to the fully matched case. Furthermore, by elevating the reaction temperature above the T_m of mismatched reagents, the undesired side products were not observed.

Interestingly, a number of different reaction types were found to exhibit 'distance-independence' when templated using DTS. It was found that the rate of these reactions was not affected by the number of intervening bases between the reacting components, for both the E and H architectures (Figure 1.13).

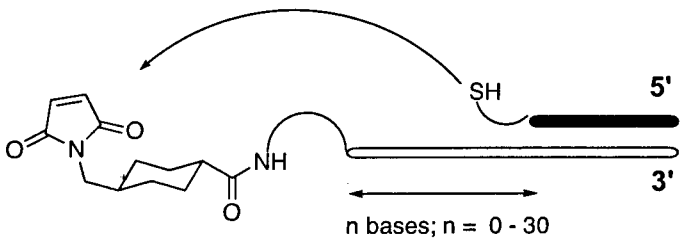


Figure 1.13. Distance-independence of DTS

For the largest gap tested, n = 30, a massively large pseudo-macrocyclic transition state is passed through during the reaction lifetime. The apparent ease with which this occurs is in stark contrast with untemplated macrocyclisations. To discover the origin of distance independence, a number of unnatural strands were synthesised wherein the region of unpaired bases was modified. Different structural motifs (Figure 1.14) replaced the oligodeoxynucleotides to determine any possible contribution of the following parameters toward distance independence: Interbase interactions (**27**); Conformational preferences of the backbone (**28**); Electrostatic interactions of the backbone (**29**) and backbone hydrophilicity (**30**).

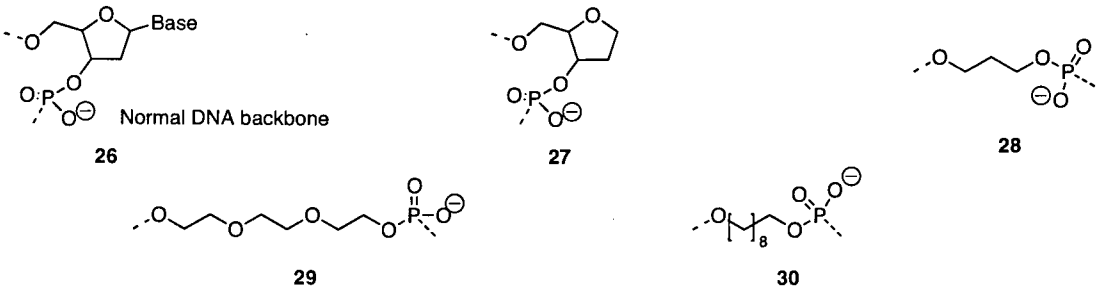


Figure 1.14 DNA backbone surrogates

Substitution of the intervening bases with any of the analogues in Figure 1.14 had no effect on the rate of product formation; the backbone structure of the template has no profound influence on the rate of DTS. However when a 10 base strand of DNA was added (complementary to the central portion of the n = 30 oligomer) the reaction was severely impaired, suggesting that unhindered mobility of this region is critical to distance independence. Based on this evidence, it was proposed that distance independence is a consequence of DNA annealing being rate determining. On the reaction timescale, formation of the helix is much slower than the templated reaction, regardless of the number of bases between the reacting centres. If this hypothesis is correct, the reaction will slow down at lower concentrations of reagent; for n = 1 – 10, there was a marked decrease in rate when reagent concentration was lowered. The observation implies that the rate accelerating effects of DTS can be substantial enough to have DNA annealing become rate determining. Distance independence is an important discovery as potentially one DNA template could code for different reagents at different regions on the strand. This would allow multistep reaction sequences to be programmed for a single substrate on the same template.

1.1.2.3 Amplifiable and evolvable small molecule libraries

Liu *et al.* hold an interest in the creation of amplifiable and evolvable libraries of non-natural small molecules.[22, 26, 54] The programmability of DNA and the potential for multistep reactions presents the opportunity to achieve this using DTS (Figure 1.14).

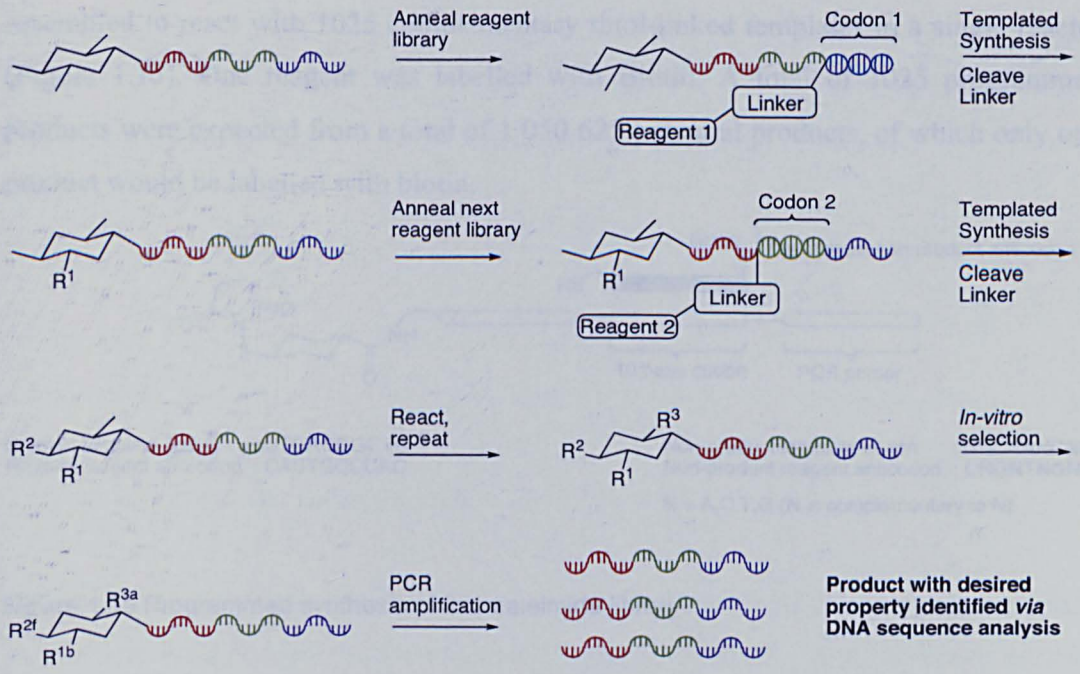


Figure 1.14 Evolvable small molecule libraries using DTS

Conceptually, a pool of unique DNA templates (each consisting of three codons) are tagged onto a basic core structure (one template per structure) and are then treated with a reagent library. Each member of the reagent library is tagged with a single anticodon that is complementary only with a codon located in the codon-1 region of the DNA templates. The tagged reagents locate and anneal with the correct template and react with the core structure that the template is attached to. The vestigial linker and anti-codon are then separated from the template and removed from the pool. Another reagent library is then added to the template pool that is coded to react with the codon-2 region of the templates and the iterative sequence is continued. A third and final reagent library that is encoded for the codon-3 region of the templates is used to complete the chemical elaboration of the original core structure. With a sufficient number of properly encoded templates, the final library of compounds will contain every possible permutation of the reagent-modified core structure. These compounds are screened *in-vitro* for a desired property and isolated. Successful candidates of the screening process are identified by their unique DNA template tag and used as the basis for further chemical evolution.

Initially a proof of concept study was undertaken to validate the potential for multi-step DTS. A library of 1025 E-template (Figure 1.11) linked maleimide reagents were

assembled to react with 1025 complementary thiol-linked templates in a single reactor (Figure 1.15). One reagent was labelled with Biotin. A total of 1025 programmed products were expected from a total of 1 050 625 potential products, of which only one product would be labelled with biotin.

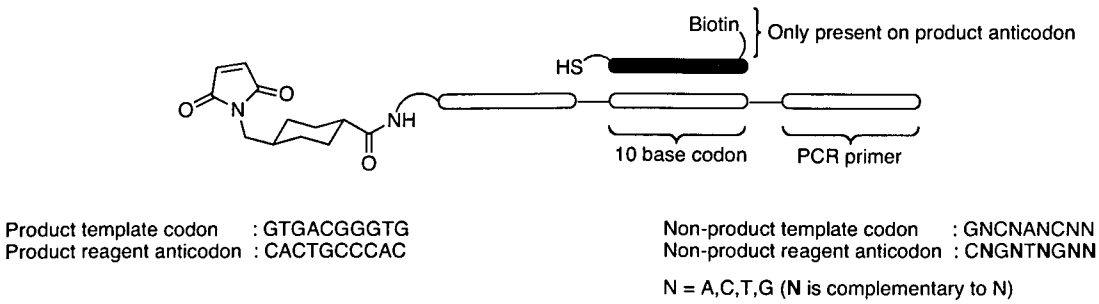


Figure 1.15 Programmed synthesis of thiomaleimide library

Equimolar amounts of the reagents and templates at 60nM *total* concentration were mixed at 25°C for 10 min. The resultant mixed products were selected *in vitro* for binding to immobilised streptavidin; biotin exhibits very strong streptavidin affinity. The surviving molecules were amplified *via* PCR using primers (attached beyond the coding region) and analysed with a restriction endonuclease that digests the biotin anti-codon only. The ratio of biotin-labelled sequences to non-biotin labelled sequences amplified during PCR was found to be 1:1; this is in excess of a 1000-fold enrichment of the biotin containing product. It should be noted that the experiment was designed to elicit success by virtue of the product template never being more than 50% matched to the non-biotin labelled reagents.

Liu *et al.* have expanded the scope of DTS and reported a broader range of chemistries that have been performed upon an E template (Figure 1.16).

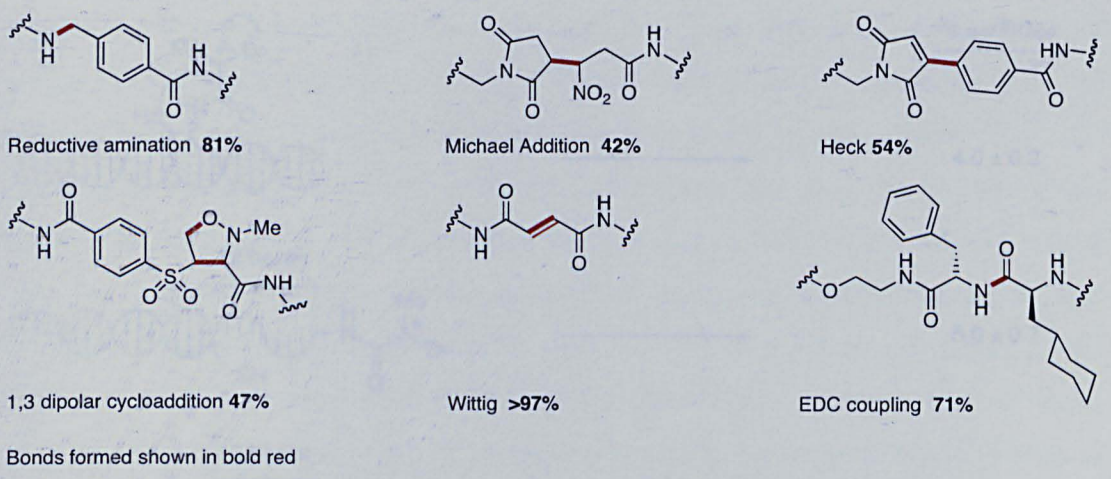


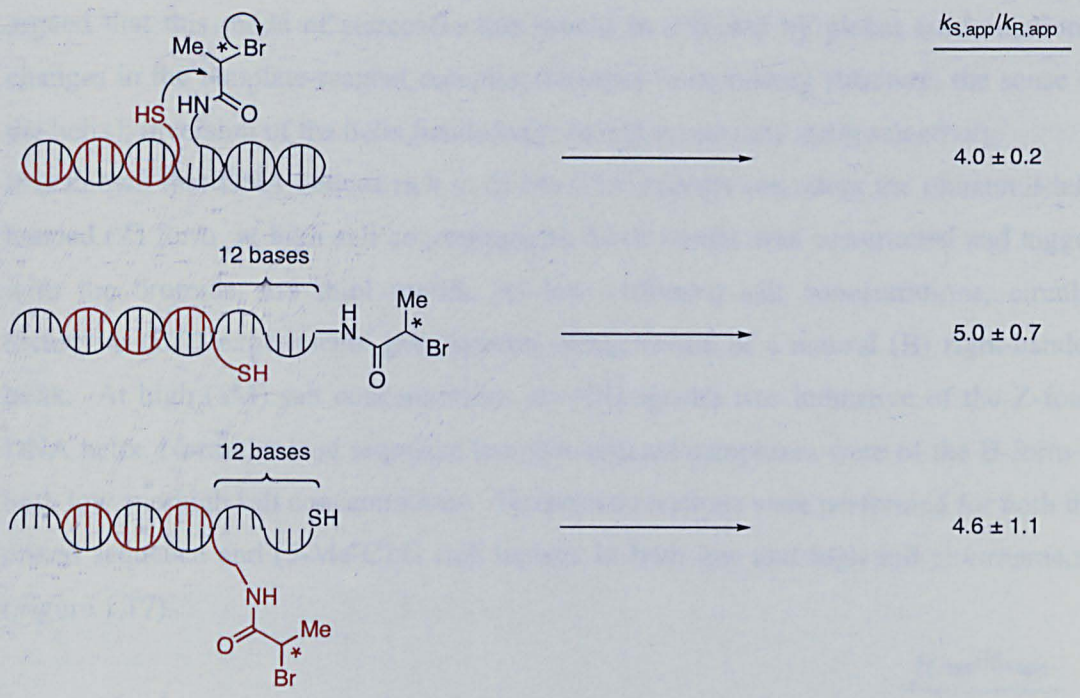
Figure 1.16. Reactions enabled by DTS

Sequence selectivity was absolute in all cases with untemplated or mismatched reaction products not being observed. Yields generally ranged from 40 – 90%, except in the case of cycloadditions of unactivated alkenes (<10%).

It was found that the Wittig, nitroaldol, Heck and EDC coupling reactions operated with distance independence (*vide supra*) for a range of 0-10 intervening bases. There is no conclusive explanation at this time to determine or predict whether a reaction will operate with distance independence or not.

1.1.2.4 Stereoselective DTS

Liu *et al.* reported that DTS exhibited "stereoselectivity" for a simple S_N2 reaction on both the E and H template architectures (Scheme 1.5).[27] In this instance, "stereoselectivity" is used to describe a marked difference in rate of a DTS mediated reaction between a nucleophile and two electrophiles that have their reactive centre set in opposing configurations. By measuring the apparent relative rates of reaction of an alpha-keto bromide with a thiol (for both the *S* and *R* electrophiles), it was possible to determine how structural changes to the DNA helix influenced the stereoselectivity.

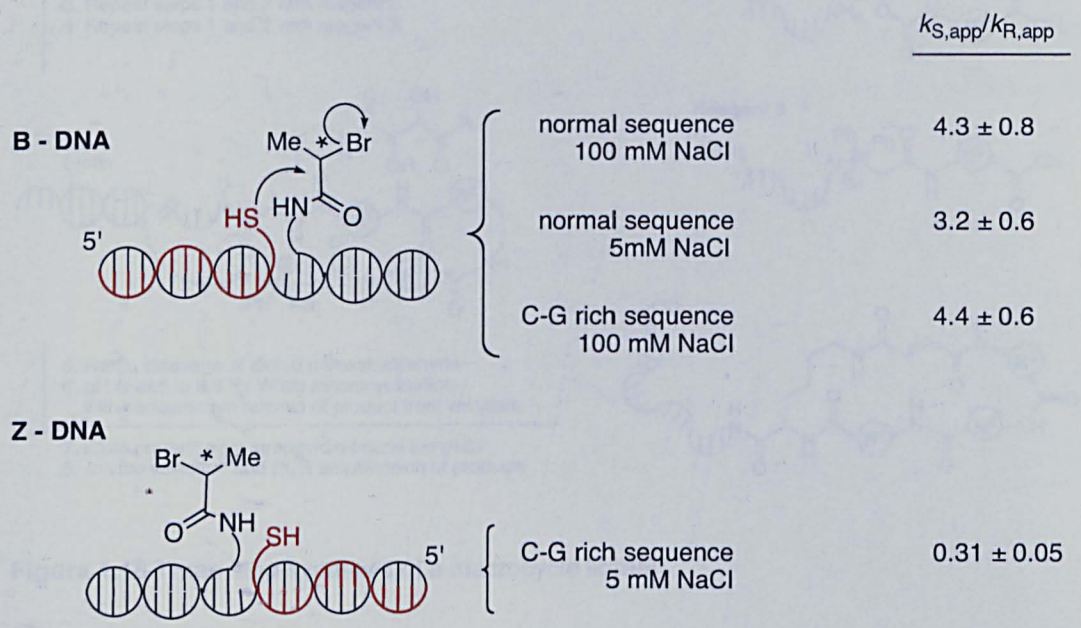


Scheme 1.5 Relative rates of reaction for enantiomers of an α-bromo amide with a thiol

It was found that the selectivity observed was independent of whether the bromide was bound to the template or reagent DNA strand. Separating the reacting components with 12 nucleotides did not appreciably alter the apparent ratio of rate constants. These findings suggest that the conformation of the DNA scaffold imposes itself upon the transition state of the templated reaction. To determine the effects of the backbone conformation on the stereoselectivity, the 12 base unpaired region was replaced with an achiral poly(ethylene glycol) linker of similar length. Stereoselectivity was no longer observed for either the thiol or bromide linked template systems. To further understand this phenomenon, a template was constructed in which 11 of the 12 unpaired oligonucleotides (all except the one immediately adjacent to the reacting centre) had their nucleobases excised and replaced with a proton. Stereoselectivity was lost, demonstrating that the configuration of the terminal template nucleotide is not (solely) directing the stereoselectivity of the reaction. As the nucleotides were sequentially restored to the template (from either 5' to 3' or 3' to 5') stereoselectivity gradually increased, however no effect was observed until at least 5-6 bases were in place. This implies that base-stacking is an important feature in the stabilisation of single-stranded DNA and crucial in the induction of stereoselectivity in DTS. By corollary, it was

argued that this mode of stereoselection would be affected by global conformational changes in the template-reagent complex (changes in secondary structure, the sense of the helix); inversion of the helix handedness should reverse the stereoselectivity.

It is known that DNA helices rich in (5-Me-C):G repeats can adopt the unnatural left-handed (Z) form at high salt concentrations. Such a helix was constructed and tagged with the bromide and thiol motifs. At low (100mM) salt concentrations, circular dichroism (CD) experiments gave spectra characteristic of a natural (B) right-handed helix. At high (5M) salt concentrations, the CD spectra was indicative of the Z-form DNA helix. Normal mixed sequence template/regeant complexes were of the B-form at both low and high salt concentrations. Templated reactions were performed for both the mixed sequence and (5-Me-C):G rich helixes in both low and high salt environments (Figure 1.17).



Scheme 1.17 Stereoselectivity inversion upon reversal of template secondary structure

The Z-DNA form gave a dramatic change in stereoselectivity. The results are consistent with the theory that the conformation of the helix is the stereochemical origin of the stereoselectivity observed with this DTS reaction.

1.1.2.5 Multi-step DTS large molecule libraries

Liu *et al* pursued the potential of DTS for the generation of an evolvable library of molecules.[55, 56] It was considered feasible to construct a diverse library of macrocyclic compounds using a three step templated synthesis (Figure 1.18).

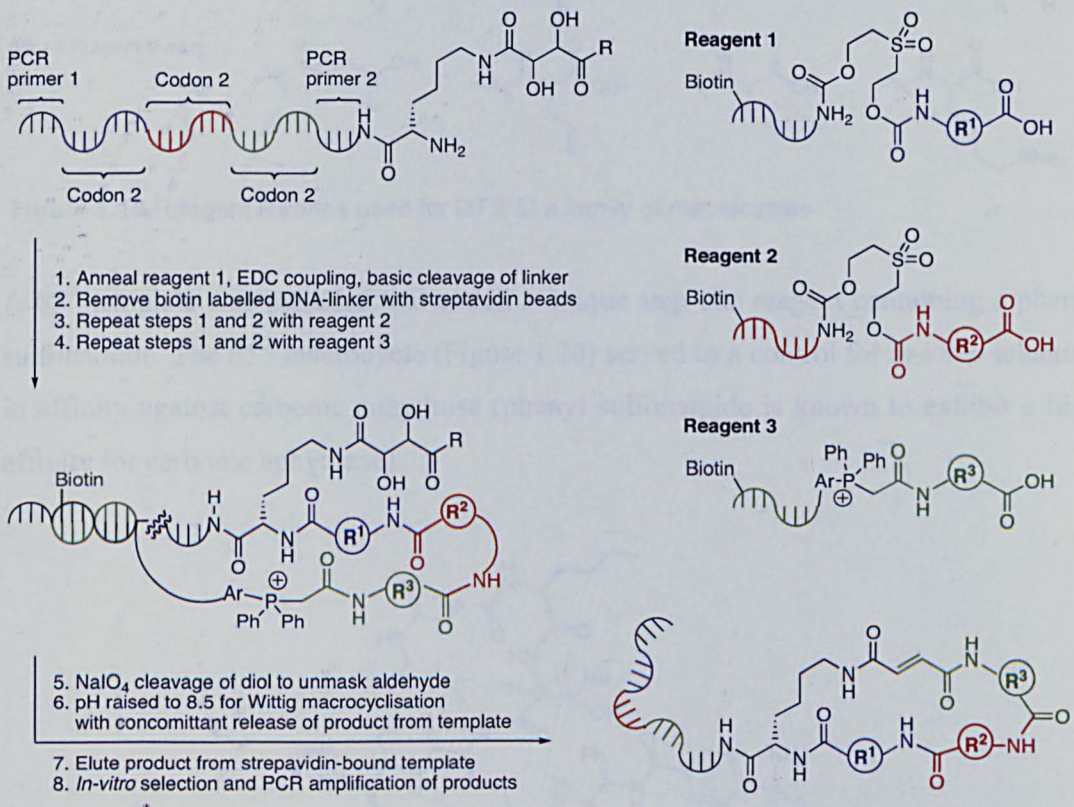


Figure 1.18 Three step synthesis of a macrocycle library

A pool of 64 templates were constructed that coded for each combination of four step 1,2 and 3 building blocks (Figure 1.19).

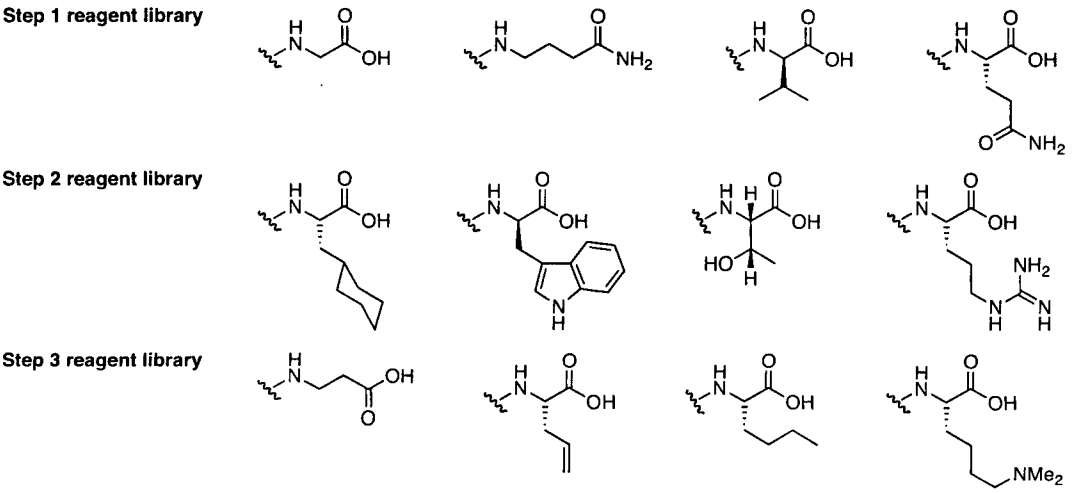


Figure 1.19 Reagent libraries used for DTS of a family of macrocycles

A 65th template was included that coded a unique step one reagent containing a phenyl sulfonamide. The 65th macrocycle (Figure 1.20) served as a control for *in-vitro* selection in affinity against carbonic anhydrase (phenyl sulfonamide is known to exhibit a high affinity for carbonic anhydrase).

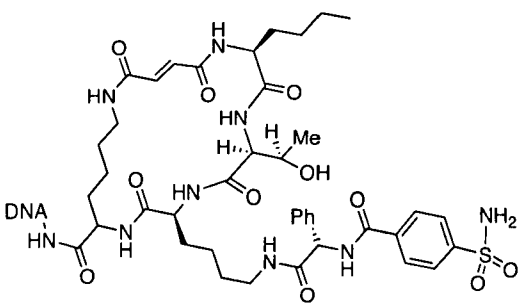


Figure 1.20 Positive control macrocycle

The 64 macrocycles were synthesised and isolated in 1 – 5% overall yield. Comparison of some of the final library compounds to authentic samples (non-DNA linked analogues) proved that the reaction conditions did not induce epimerisation of any stereocentres. Two rounds of *in-vitro* selection (for binding affinity against carbonic anhydrase) followed by PCR amplification and strand analysis gave the positive control macrocycle as the only molecule surviving selection. The work proved that it is possible to rapidly assemble a library of small complex molecules and screen *in-vitro* for protein affinity using DTS.

Using the codon methodology, an array of diverse linear molecules were also prepared to exemplify Iterated Branching Reaction Pathway Synthesis (IBRPS, Figure 1.21) *via* DTS[57]. IBRPS is the generation of a structurally diverse library by means of repeatedly executing multiple reaction pathways from each library member. With each iteration, the number of library members increases.

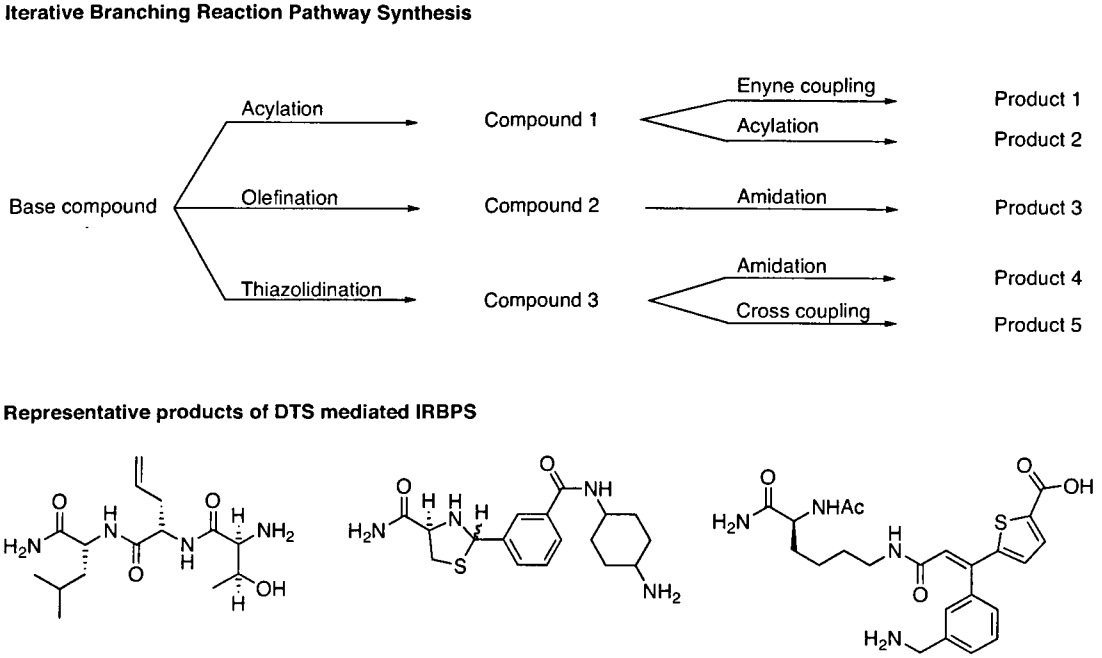


Figure 1.21 IBRPS using three codon DTS technology

The increased chemical complexity and diversity of the final library members illustrates the potential of DTS for small molecule synthesis and (biological) screening.

Recently published work by Liu *et al.* has centred on ordered multistep synthesis in a single solution.[58] By exploiting the temperature and strand length dependencies of DNA hybridisation it was possible to complex three reagents on one template and have them react sequentially in a desired order (Figure 1.22).

At low temperature, the three reagents anneal upon the template and are held so that only two reagents are positioned to react. As the temperature is raised, the desired reaction occurs between reagents 1 and 2, and some chemical structure is passed from reagent 1 to reagent 2. The first reagent DNA strand then dissociates from the template as its T_m is passed; the reagents T_m is controlled by the length of their DNA strand, the

longer it is, the greater it is. At the elevated temperature, the template and remaining reagents rehybridise to bring the second and third reagent strands into reactive proximity. The second reaction takes place and the growing molecule is passed to the third reagent. The temperature is raised again to remove the second reagent DNA strand and allow rehybridisation of the template and third reagent. After the third reaction, the final reagent DNA strand is removed to give the desired linear product. Figure 1.22 shows a representative product of this approach, the result of three sequential Wittig reactions. A trimer peptide has also been prepared using this methodology.

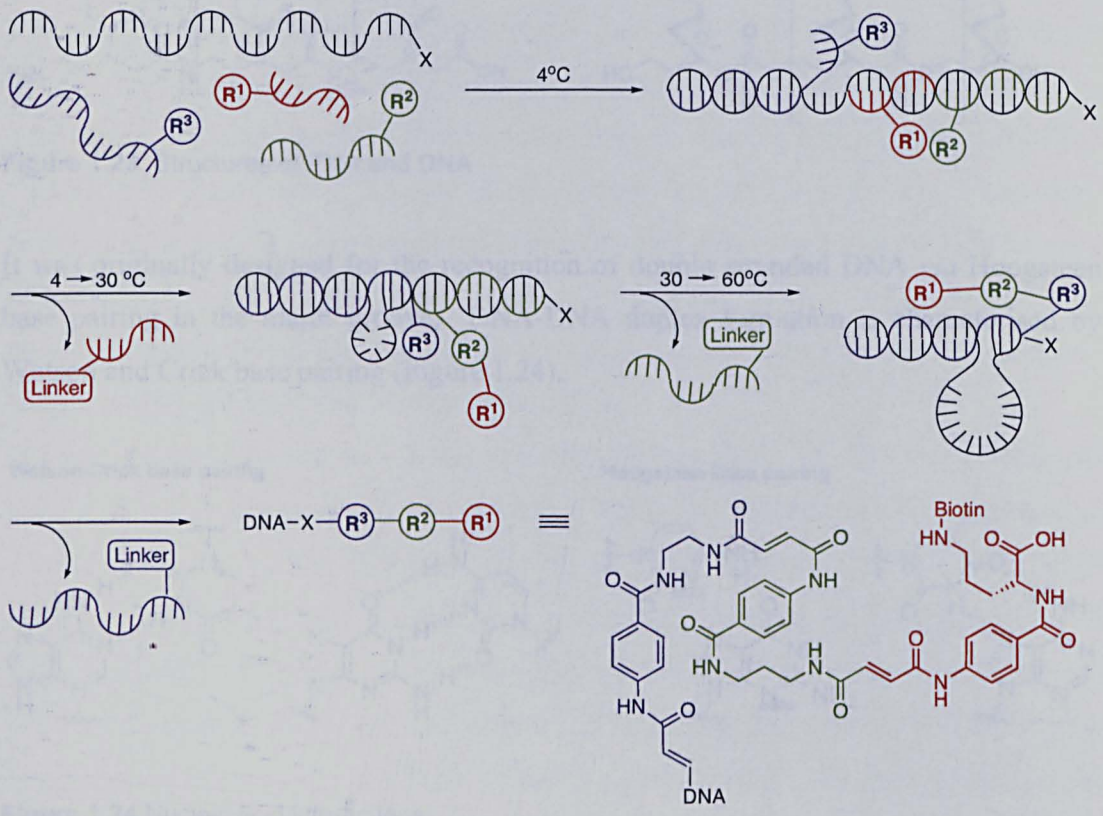


Figure 1.22 Ordered multistep synthesis in a single solution

The power of DTS lies with its ability to program the reactivity of a mixture of reagents to yield exclusively a selection of predicted structures. The DNA template also acts as a readily amplified and unique tag, allowing the conservative use of precious reagents without compromising ease of identification. The templated reactions proceed at a much enhanced rate compared to their untemplated counterparts by virtue of an effective

increase in reagent molarity. DTS has been demonstrated to be a powerful methodology for the rapid assembly of small complex molecules.

1.2 Peptide Nucleic Acid

PNA is a structural mimic of DNA in which the deoxyribose phosphodiester back bone has been replaced with homomorphous aminoethyl glycine (Figure 1.23).[59]

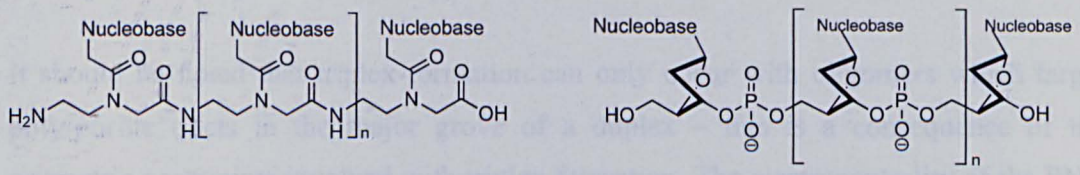


Figure 1.23 Structures of PNA and DNA

It was originally designed for the recognition of double stranded DNA *via* Hoogsteen base pairing in the major groove. DNA-DNA duplex formation is characterised by Watson and Crick base pairing (Figure 1.24).

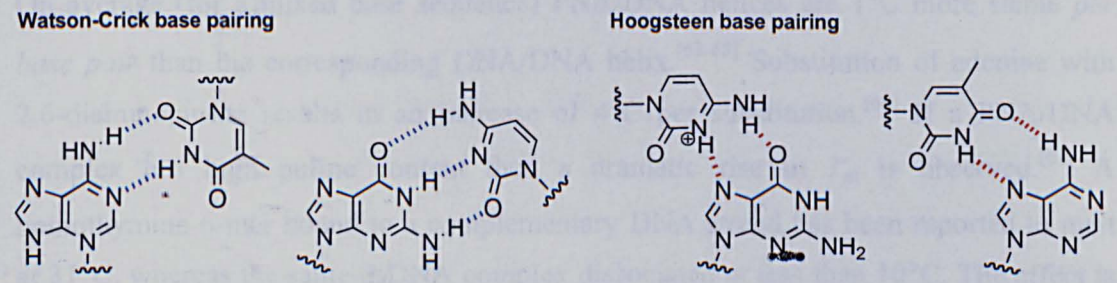
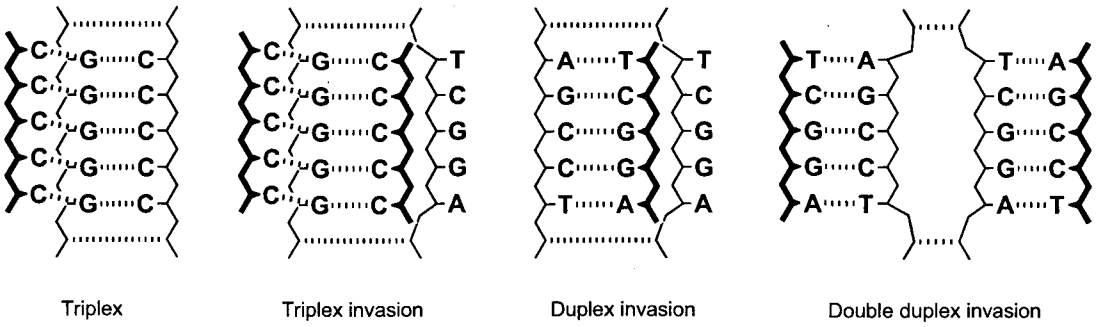


Figure 1.24 Nucleic acid interactions

The PNA backbone is characteristically different to its DNA counterpart being that it is acyclic, achiral and carries no formal charge. PNA oligomers have both parallel and anti-parallel binding modes, although the antiparallel form is strongly preferred.

Single stranded PNA can interact with DNA helices in more than one way (Scheme 1.25).[60-62]



Scheme 1.25 PNA modes of binding with DNA

It should be noted that triplex-formation can only occur with oligomers which target poly-purine tracts in the major grove of a duplex – this is a consequence of the geometric constraints involved with triplex formation. The electroneutrality of the PNA backbone is significant. Annealed DNA strands experience repulsive interactions between the negatively charged phosphate backbones. PNA strands do not suffer these destabilising forces and bind more tightly to DNA and RNA than oligodeoxynucleotides; this is reflected in the higher T_m values for any given strand length of PNA relative to its DNA counterpart.

On average (for a mixed base sequence) PNA/DNA helices are 1°C more stable *per base pair* than the corresponding DNA/DNA helix.[63-65] Substitution of adenine with 2,6-diaminopurine results in an increase of 4°C per substitution.[66] If a PNA/DNA complex has high purine content then a dramatic rise in T_m is observed.[67] A homothymine 6-mer bound to a complementary DNA strand has been reported to melt at 31°C, whereas the same dsDNA complex dissociated at less than 10°C. The effect is more pronounced for a homothymine 10mer – this melts at 73°C as opposed to 23°C for the dsDNA helix.

Mismatched bases destabilise annealed PNA/DNA strands to a far greater extent than their DNA/DNA equivalents; on average the T_m is reduced by 10°C for a DNA/DNA point mutation, and 15°C for the equivalent PNA/DNA helix.[68-70] The corollary of enhanced binding and selectivity is the ability to use shorter PNA strands in lieu of DNA to achieve the same level of affinity with a given DNA sequence.

DNA/RNA hybridisation is dependant on the salt concentration within their medium whereas PNA hybridisation is not.[71] This means at low salt concentrations, DNA and RNA helices denature and unwind. PNA can therefore more easily invade and displace

complementary DNA strands of target sequences at lower salt concentrations as the hybridisation kinetics are more facile.

The amino ethylglycine backbone of PNA is not easily recognised by nucleases or proteases so PNA oligomers exhibit resistance to enzymatic degradation.[72, 73] It is also stable over a wide pH range, unlike DNA. These properties allow an increased lifetime of PNA (relative to DNA) for *in-vitro* and *in-vivo* applications.

Applications of PNA as a DNA based equivalent already exist.[74-78] The enhanced thermal and chemical stabilities of PNA however often translate into tangible advantages in such applications, driving their inclusion. Greater selectivity and binding affinities of PNA strands relative to DNA based probes allow smaller oligomers to be used in most biological assays.

1.2.1 Applications of PNA

As a mimic of DNA, PNA can be used in many DNA-based applications.

1.2.1.1 Molecular Beacons

It is important in biochemical research to have the means for labelling chromosomes in specific places.[79-83] A popular method for the sequence specific labelling of oligonucleotides is the use of a molecular beacon (MB). Molecular beacons are fluorescence based probes that are unable to emit light until they are bound to their targeted sequence (Figure 1.26). The fluorescent reporter group is anchored onto a terminus of a hairpin oligonucleotide. It is held in close proximity to a fluorescence quencher situated on the other strand terminus. While the hairpin structure is maintained, either through nucleobase-nucleobase interactions or hydrophobic forces, the quenching unit stifles any fluorescence from the reporter dye. When the MB successfully anneals and hybridises with the target DNA sequence the dye and quencher are sufficiently separated in space that quenching of the fluorophore is minimal and a signal is observed.

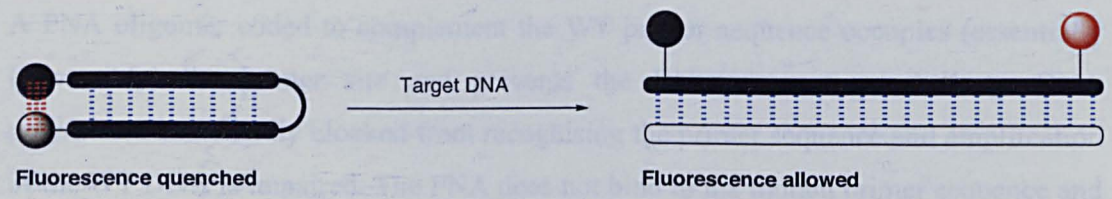


Figure 1.26 PNA beacon fluorescence

There is no obligation to wash away excess unhybridised probes after treatment of the DNA sample because they are unable to affect the signal of the active probes. This translates to a tangible advantage that allows the use of MBs *in-vivo* where washing techniques cannot be applied. Real time monitoring of both location and concentration of target oligonucleotides is possible when administering large quantities of MBs into cellular structures. As a target sequence is replicated it hybridises with the stockpiled MBs; this technique is known as Fluorescence In-Situ Hybridisation (FISH).[84-86] PNA is an adept foundation for MBs and well suited for FISH, displaying on average three times greater signal intensity than its DNA counterpart, as a result of its great binding affinity.[87-92]

1.2.1.2 PCR Clamping

PNA can effect the rapid detection and identification of mutant DNA by selective amplification of the rogue sequences.[93-95] The high selectivity and affinity of PNA allows the discrimination of deviant oligonucleotides that differ by as little as one base pair. Detection of rare mutant targets is achieved by the selective inhibition of wild-type DNA (WT DNA) amplification (Figure 1.27).

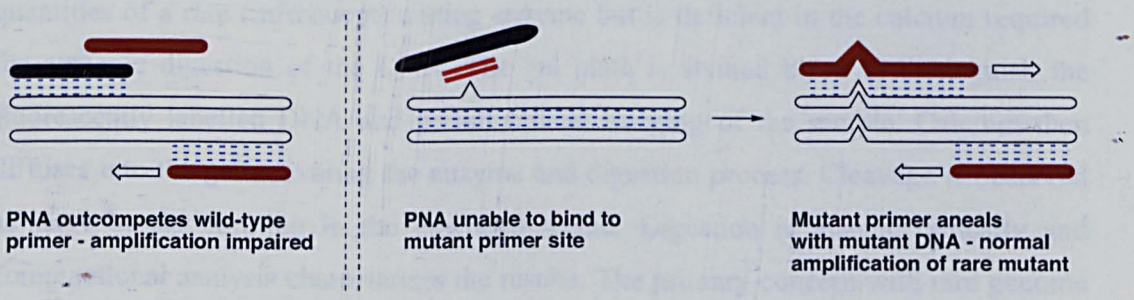


Figure 1.27 PCR clamp assisted enrichment of rare mutant gene

A PNA oligomer coded to complement the WT primer sequence occupies (essentially irreversibly) the primer site and prevents the WT primer from docking. DNA polymerase is sterically blocked from recognising the primer sequence and amplifcation of the WT DNA is impaired. The PNA does not bind to the mutant primer sequence and normal amplification ensues. Under these conditions amplification of the MU DNA dominates over production of the WT amplicons. The final pool of amplicons is massively enriched with the desired mutant sequence. The approach can be applied in other ways; placing the clamp either adjacent to or distal to the primer sequence also acts to arrest primer elongation.

1.2.1.3 Rare genome cutting

Rare genome cutting is the *infrequent* scission of a strand of DNA into a small number of large fragments. An optical imaging technique for the sequencing of large DNA fragments introduced in 1993 by Wang *et al.* hastened genomic mapping and is dependent upon rare-genome cutting (Figure 1.28).[96, 97]

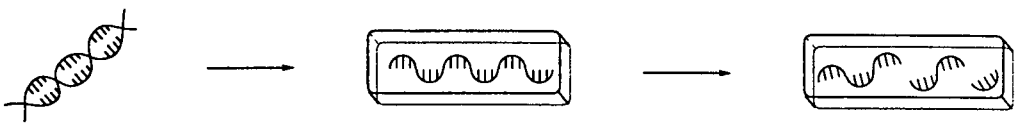


Figure 1.28 Rare genome cutting of immobilised DNA

In the Wang technique, the chromosome of interest is thermally denatured into single strands and trapped in molten agarose gel in an elongated form. The gel contains large quantities of a rare (infrequent) cutting enzyme but is deficient in the calcium required for enzymic digestion of the DNA. The gel plate is stained black to distinguish the fluorescently labelled DNA and permit optical imaging of the sample. Calcium then diffuses into the gel activating the enzyme and digestion process. Cleavage is observed as large breaks forming in the unwound strand. Digestion is imaged optically and computational analysis characterises the results. The primary concern with rare genome cutting is a lack of enzymes that can offer infrequent cutting – it is of benefit to genomic mapping to swell the pool of available rare-cutters.

Nielsen *et al.* devised a strategy to allow the wide range of regular restriction enzymes to act as rare genome cutters.[98, 99] PNA clamps are used to mask arbitrarily a small number of sites that are susceptible to a restriction enzyme. The partially masked DNA is exposed to DNA methylase that alkylates the unmasked nucleotides, thus preventing their recognition by restriction enzymes. Thermal unclamping of the PNA reveals the unmethylated regions of the chromosome that are subsequently the only regions susceptible to recognition by and digestion with the restriction endonuclease. Application of the enzyme results in the scission of the few sites eligible for digestion (Figure 1.29).

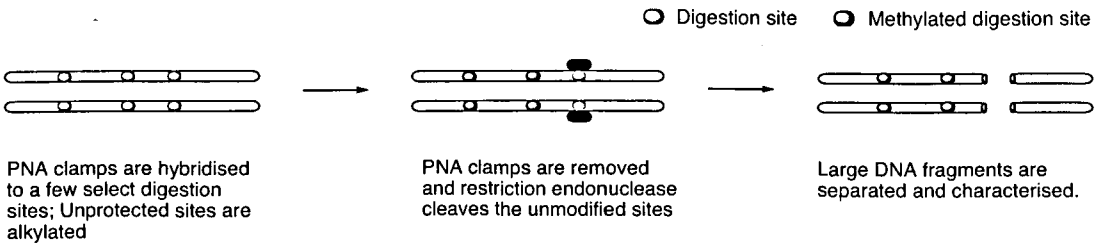


Figure 1.29 Rare cutting of a large gene using a regular restriction endonuclease

Considered choice of PNA clamp and restriction enzyme significantly widens the breadth of rare cutting systems available, expediting the characterisation of large sections of genomic DNA.

1.2.1.4 Antisense Gene therapy

A promising application of PNA lies within antisense gene therapy.[100-102] The transcription of certain genes can have an adverse effects on the human body. Sickle cell anaemia, Hepatitis, cancer and HIV are all potential targets for antisense techniques as a part or whole gene-targeted therapy.[103] Antisense therapy operates by introducing a nucleotide oligomer that either intercepts and incapacitates the mRNA produced by the harmful genes (Figure 1.30)[68] or prevents transcription of mRNA *via* steric blocking of chromosomal DNA binding.[104] Preventing the transcription of mRNA halts the expression of the harmful genes, potentially curing the ailment it induces. Targeting mRNA (instead of genomic DNA) is often the easier and more responsive approach to antisense therapy as delivering PNA inside the cell nucleus is a formidable task.

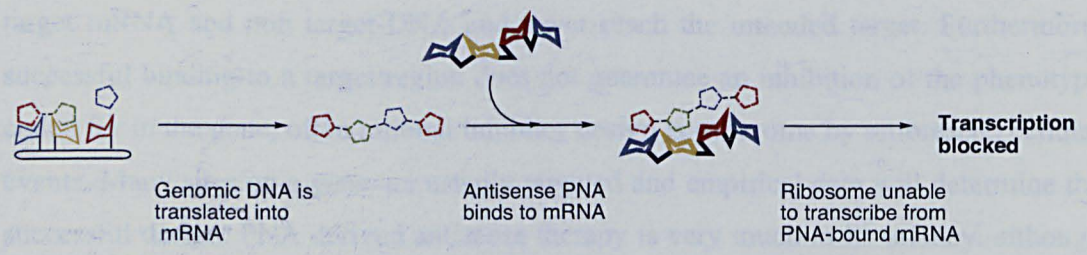


Figure 1.30 Antisense therapy that targets mRNA

Antisense therapy is relatively new and only one drug (based on DNA) has been approved for therapeutic use by the Food and Drug Administration of America – Formivirsen.[105] The drug, marketed as Vitravene®, is an inhibitor of the CytoMegaloVirus (CMV) which can cause retinitis in AIDS sufferers. CMV is a herpes related virus that is found in the vast majority of humans worldwide, but is held in remission by CD_4 immunoglobin cells. AIDS patients whose immune system has been compromised are susceptible to CMV induced retinitis (loss of peripheral and night vision) as they cannot muster a sufficient defence against CMV.[106] Formivirsen intervenes by suppressing the viral replication *via* the mRNA coding for reproduction of the virus. It is postulated that a suitably attenuated PNA derivative would be a more effective drug.

Initial development of PNA-based antisense drugs was stunted by poor cellular uptake of PNA. For simple prokaryote cellular organisms (bacteria) it was found that attachment of a cell-penetrating peptide (KFFKFFKFFK) to the PNA strand raised the efficiency of cellular invasion to a level suitable for therapeutic applications.[107, 108] The mechanism of how the peptide mediates passage through the cell wall is not understood at this time.[109-111] Entry into eukarotic cells requires more advanced chicanery. A PNA-DNA chimera is prepared whereby the conjoined strands are complementary, thus forming a hairpin. The hybrid is treated with a cationic lipid (which is attracted to the anionic DNA backbone) that transports the PNA across the cellular divide.[112-115] Inside the cell, dissociation of the lipid and hairpin nucleotide structure occurs under the physiological conditions to deliver the drug *in-vivo*. It is critical that the stability of the hairpin is less than that of the PNA-target complex but sufficient that it can maintain its integrity to present a hydrophobic package for the lipid to encapsulate. Beyond entry to the cell, other challenges remain to hinder a successful antigene interaction. Nucleotide oligomers often show high affinity for proteins (cell surface and intracellular), non

target mRNA and non target-DNA and never reach the intended target. Furthermore, successful binding to a target region does not guarantee an inhibition of the phenotype coded for in the gene; often rational inhibitor design is overcome by unforeseen cellular events. Many sites on a gene are usually targeted and empirical data will determine the successful design. PNA derived antisense therapy is very much in its infancy, although recent patent literature indicates that the field is moving forward.[116-120] At this time there is very little published work in PNA based antisense therapy, but research is still ongoing.[121]

1.3.0 Peptide synthesis

Peptides can be synthesised by man by either chemical methods or hijacking biological machinery provided by nature.[122-126] Both techniques have relative advantages and disadvantages but ultimately are seen to be complementary and indispensable.

1.3.1 Biological synthesis of peptides

1.3.1.1 Recombinant DNA synthesis

Man is able to take advantage of natures elegant solution to peptide and protein synthesis. Prolonged study of the underlying biology of protein synthesis has allowed the inception of transgenic techniques (Figure 1.31) for the synthesis of unnatural peptide sequence. DNA that codes for a designer peptide is introduced into the genome of a slave organism *via* a vector plasmid.[127, 128]

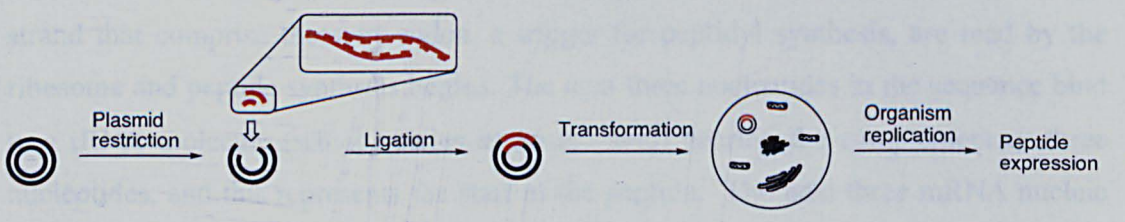


Figure 1.31 Gene cloning

Circular plasmid DNA is cleaved open using restriction endonucleases to give an open-form DNA tapered with "sticky ends. The foreign DNA is complexed with the open plasmid *via* short unpaired nucleotides that comprise the complementary "sticky ends". The plasmid backbone is reformed using a DNA ligase enzyme to give the completed

transgenic vector. The plasmid is introduced to the host organism (eg yeast strain) via endocytosis and assimilated into the cell's anabolic pathways. Culture of the organism multiplies the number of copies of the designer DNA and ultimately magnifies the expression of the desired peptide. Cell lysis of the fermentation broth gives a crude mixture of cellular debris that can be purified by centrifugation and electrophoresis to give the required peptide.

1.3.1.2 Mechanism of peptide synthesis from DNA

The translation of human genetic information from DNA to a final protein is mediated by a school of enzymes and the catalytic ribosomes (Figure 1.32). [126]

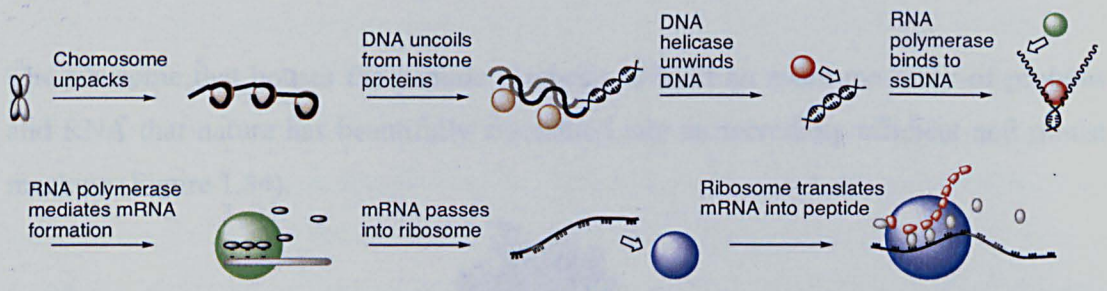


Figure 1.32 Gene translation and transcription

Supercoiled DNA unravels from structural proteins and a series of enzymes unwind the helical dsDNA and transcribe a gene into a mRNA oligomer. The mRNA passes from the nucleus into the cytosol and docks with a ribosome. Within the ribosome, the mRNA acts as a template for peptide construction. Three nucleotides of the mRNA strand that comprise the start codon, a trigger for peptidyl synthesis, are read by the ribosome and peptide synthesis begins. The next three nucleotides in the sequence bind to a tRNA molecule (which carries an amino acid) bearing the complementary three nucleotides, and this represents the start of the peptide. The next three mRNA nucleic acids similarly code for another tRNA molecule which docks along side the first. The two amino acids carried by the tRNA molecules are now in sustained close proximity and react together. The second amino acid initiates an addition-elimination reaction with the first amino acid; amino acid one is liberated from its tRNA carrier molecule and forms an amide bond to amino acid two (Figure 1.33). The process is repeated until the

stop codon is encountered, at which point the final peptide is passed out from the ribosome.

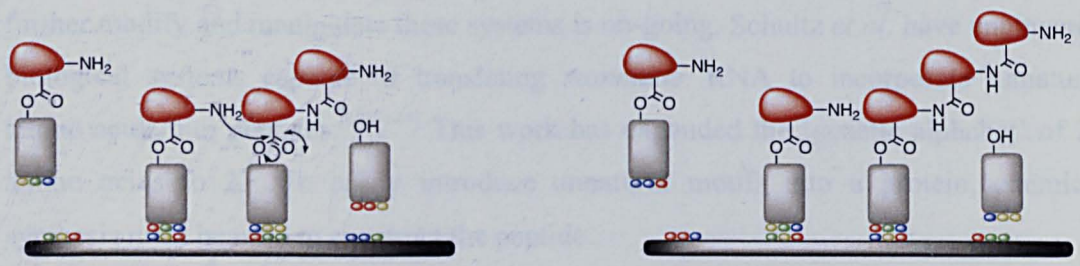


Figure 1.33 mRNA templated synthesis of peptides within the ribosome

The peptide is then transported to other cellular organelles for post-translational modifications to give a fully-fledged protein.

The ribosome that houses the peptide synthesis is itself an awesome array of proteins and RNA that nature has beautifully assembled into an incredibly efficient and robust machine (Figure 1.34).

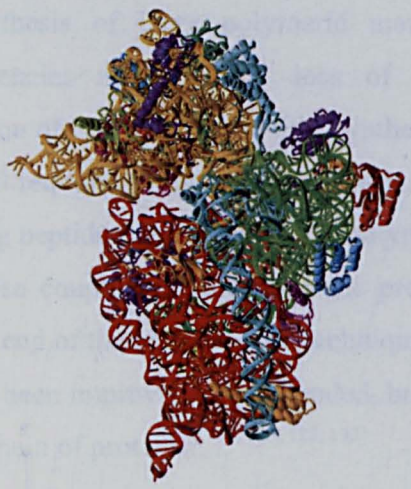


Figure 1.34 Representation of the 30S subunit of a bacterial ribosome

Ramakrishnan *et al.* have resolved ribosomal crystal structures to less than 3Å resolution allowing detailed insight on the mRNA decoding process (i.e. how mRNA and tRNA interact) of ribosomal peptide synthesis.[129, 130] Additionally, the mode of action of some antibiotics has also been clarified.[131] This important work has served to highlight how the biological synthesis of peptides and proteins is immensely more efficient and reliable than any man-made technology.

However, when using the approach of nature, the limitations of this methodology apply. Biological systems are generally constrained to the use of 20 natural amino acids and the selection of post-translational modifications that are provided by nature. Research to further modify and manipulate these systems is on-going. Schultz *et al.* have engineered biological systems capable of translating *manmade* ^{t}RNA to incorporate unnatural amino acids into peptides.[132-134] This work has expanded the 'genetic alphabet' of 20 amino acids to 21. To easily introduce unnatural motifs into a protein, chemical synthesis must be used to construct the peptide.

1.3.2 Chemical synthesis of peptides

The first planned chemical synthesis of a peptide was performed by Fischer in the late 19th century.[135] His goal of chemical total protein synthesis could not be achieved using the techniques available at that time; an 18mer was the greatest polypeptide that was constructed. Fischer was unable to create a protein because traditional liquid phase chemistry cannot provide the level of reaction efficiency required to synthesise these macromolecules. The synthesis of large polymeric materials requires very high monomer coupling efficiencies and minimal loss of product material through purification. The introduction of Solid Phase Peptide Synthesis (SPPS) by Merrifield in 1963 was the paradigm shift required to address these issues.[136] Merrifield realised that immobilisation of a growing peptide upon an insoluble polymer matrix allows the use of excess reagents to maximise coupling efficiencies; the product can be filtered away from excess reagents at the end of the reaction. The techniques and chemistries involved with SPPS have constantly been improved and expanded, but it remains unable to yield peptides within the size domain of proteins.[122, 123, 125, 137]

The linear stepwise assembly of amino acids becomess problematic as the growing peptide chain becomes more protein-like in size and structure. When the polypeptide attains a certain size (usually 5-15 residues in length) it begins to adopt a secondary structure *via* non-covalent forces (eg hydrogen-bonding). The *N*-terminus (upon which the chain is grown) becomes obstructed and coupling efficiencies are dramatically lowered, resulting in failure sequences. This is known as a difficult coupling.[138-141] The accruement of failure sequences is fatal to the synthesis of polypeptides because they are not easily separated from the product peptide. This is generally overcome by

'capping' unreacted amino termini with acetic anhydride/diisopropylamine to give the unreactive *N*-acyl carbamate. Capped sequences do not take part in any further peptide chemistry and provided they do occur near the end of a sunthetic sequence, are separable from the product sequence by preparative HPLC.

Techniques such as backbone-modification or addition of chaotropic salts counteract the aggregation of peptides by disrupting the internal hydrogen bonding network.[142-144] However, the occurrence of difficult couplings becomes increasingly regular with increasing chain length, and linear SPPS is limited in practice to the synthesis of 40-50mers.

To overcome the size limitations of linear SPPS, the convergent synthesis of shorter polypeptide chains is required. Large peptides with *protected* side-chains are not suitable for convergent peptide synthesis because they are not readily soluble. Removing the hydrophobic protecting groups will restore the aqueous solubility of a peptide but also creates the problem of controlling the regiochemistry of the ligating (unprotected) peptides. The solution requires that the ligating termini of the peptides are mutually and exclusively reactive; that is to say unreactive with the unprotected sidechains but not with each other.

A number of strategies have been published that utilise chemoselective reactions to overcome the regiochemical issues when joining peptides together.[145, 146] However these generate unnatural backbone linkages that are not always tolerated for the purposes of the peptide being synthesised. Ultimately, technology is required that provides a natural (native) peptide linkage post-ligation for maximal utility in modern peptide and protein chemistry.

Research by Bauer *et al.* in 1953 set forth the beginnings of Native Chemical Ligation, a chemo and regioselective technology that allows for the ligation of unprotected peptides to give a target polypeptide with no unnatural backbone linkages. [147] Bauer found that (in aqueous solution) valine, activated as the thiophenol ester, would undergo addition-elimination with the sidechain thiol of cysteine to give a *S*-linked adduct of the two amino acids. This adduct would then undergo *S* → *N*-acyl-transfer to give a native Val-Cys dipeptide. (Figure 1.35).

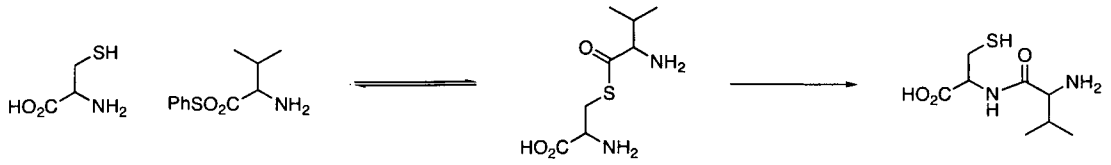


Figure 1.35 Bauer native peptide bond formation

The work of Bauer was first extended to polypeptide synthesis by Kemp *et al.* [148-151] Modifying the thiol exchange/rearrangement idea, a peptide containing a terminal cysteine was captured by a thiol-containing auxiliary that was *O*-linked to another peptide. The auxiliary served to hold the peptide fragments in close proximity to facilitate the acyl transfer (Figure 1.36)

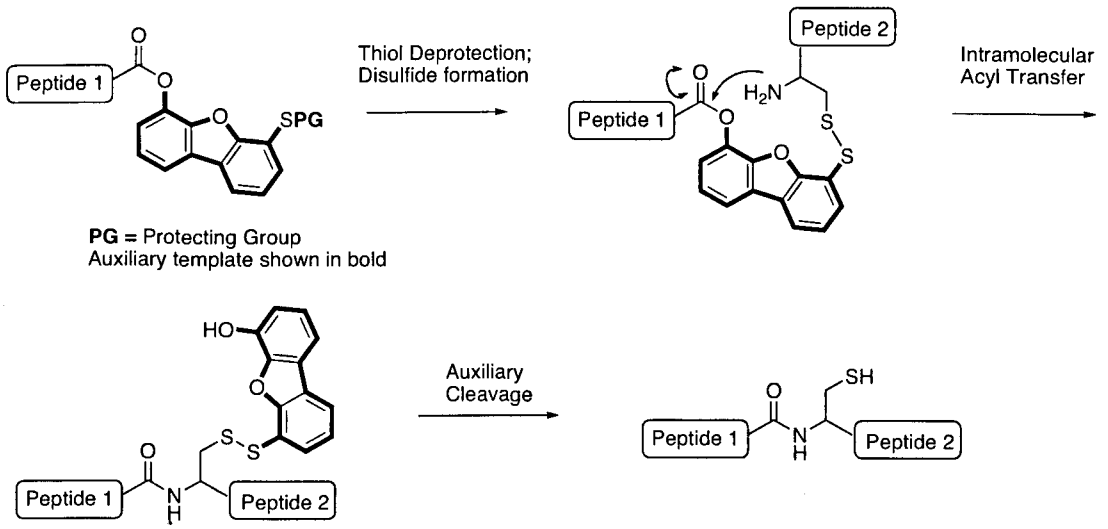


Figure 1.36 Kemps prior thiol capture technology

Although Kemp used this technology in the synthesis of a 29 residue fragment of the bovine pancreatic trypsin inhibitor, ultimately the methodology was too elaborate to be used for general purpose peptide synthesis.[152] The lengthy and non-trivial use of sulfur-deprotection and disulfide chemistry ultimately limited its applicability and potential.

Kent expanded upon the work of Bauer and published an elegant solution with the introduction of cysteine-mediated Native Chemical Ligation (NCL)[153] (Figure 1.37).

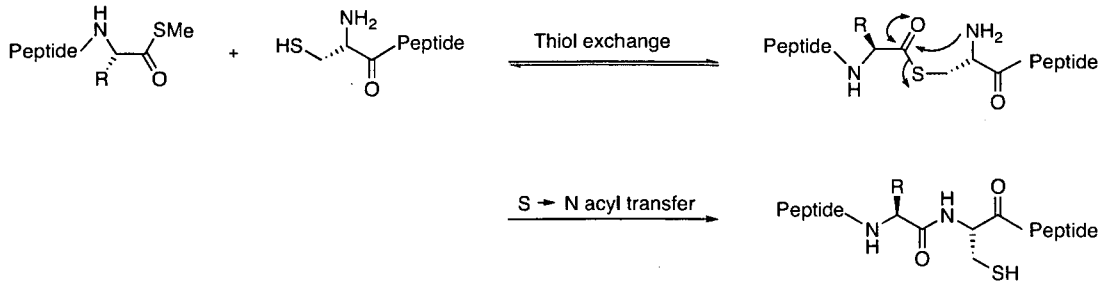


Figure 1.37 Native chemical ligation

Taking the original chemistry reported by Bauer, Kent developed it further into a working technology that is able to ligate (relatively) large peptide fragments.

A peptide bearing a terminal cysteine reacts reversibly with the thioester (under reducing conditions, to prevent disulphide formation) to join the two peptides together *via* thiol exchange. Spontaneous sulfur to nitrogen acyl shift forms a native peptide linkage between the two peptides and regenerates the cysteine side-chain. A number of small proteins have been synthesised using this approach.[154, 155]

There are two potential drawbacks to this technology. The first is the difficulty in preparing large peptides as their thioester analogues.

The second is that a cysteine residue is required to be present at the ligation junction, limiting the choice of target proteins to those that possess a cysteine residue suitably disposed along the peptide chain.

The challenge of preparing peptides as their thioester analogues has been answered with more than one useful solution and both traditional SPPS techniques and biological techniques have been employed for this cause.

Traditional SPPS of peptidyl thioesters requires that a peptide is grown upon a solid support and subsequently cleaved from it under nucleophilic conditions. Using a thiol as a cleavage reagent, the peptide is liberated as a thioester (Figure 1.38).

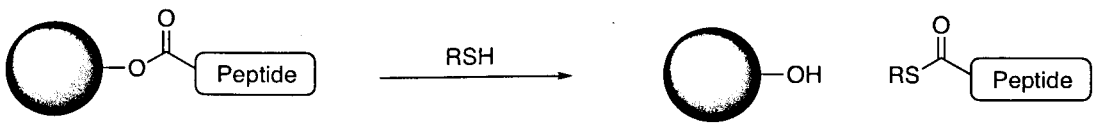


Figure 1.38 Standard SPPS of thioester peptides

The problem with this approach is that it is not compatible with basic fmoc-based SPPS conditions used for the majority of manual peptide synthesis. Ellman *et al* solved this issue by developing the 4-sulfamylbutyl "safety-catch" linker for use in SPPS of

thioesters.[156, 157] The 4-sulfamylbutyl aminomethyl resin used in this method can be loaded with an fmoc-protected amino acid using a standard DIC coupling and peptide elongation can be carried out using standard fmoc SPPS reagents and conditions. The "safety-catch" linkage is invulnerable to nucleophilic attack until it has been chemically activated (*via* mild alkylation conditions), after which the protected peptide can be cleaved by the sulfur nucleophile of choice (Figure 1.39).

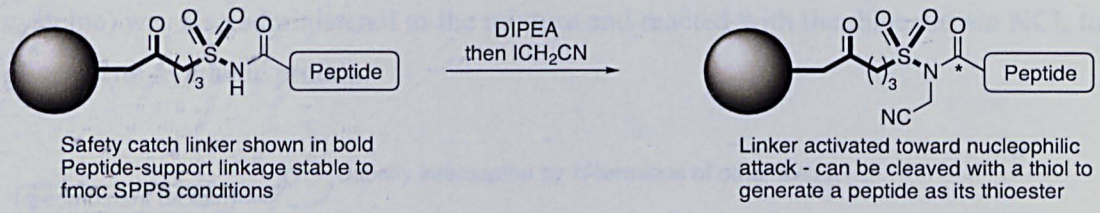


Figure 1.39 Ellman "safety-catch" linker activation

This method is used routinely to prepare sidechain-protected thioester peptide fragments of up to 50 residues (general SPPS size limit for peptides) in length.

While NCL excels at joining short peptide fragments together, it is less able to effect the solution-phase ligation of larger peptide fragments together, with hydrolysis of the activated thioester rapidly becoming dominant with increasing peptide fragment size.[158] Muir developed an elegant semi-synthetic solution, expressed protein ligation, to overcome the stability and reactivity problems of larger peptide thioesters that restrict the size of peptides accessible by NCL.[159, 160] Expressed protein ligation makes use of protein splicing, the process in which a protein undergoes an intramolecular rearrangment to excise an internal sequence (intein) and ligate the sequences that had previously flanked it (exteins) (Figure 1.40).

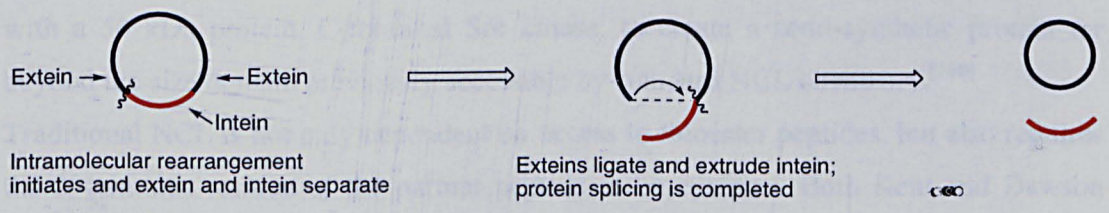


Figure 1.40 Protein splicing – extrusion of intein protein sequence

Protein splicing is mediated *in-vitro* by a splicing protein. The pathway of intein extrusion is directly analogous to NCL, involving an activated thioester intermediate made from the *C*-terminus of one extein and the thiol sidechain of an *N*-terminus

cysteine of the intein (Figure 1.41). Muir evolved a strain of *E. coli* with a defective splicing protein that would express a desired recombinant protein but could only partially undergo intein extrusion. The mutant splicing protein was unable to effect the final acyl transfer step of the NCL-like extein ligation process – thus rendering the large recombinant protein as an activated thioester. Muir intercepted this reactive intermediate with thiophenol to (transiently) produce a very large recombinant peptide as its thioester. In addition to the thiophenol a synthetic peptide (bearing a *N*-terminus cysteine) was also administered to the mixture and reacted with the thioester *via* NCL to give a semi-synthetic protein.

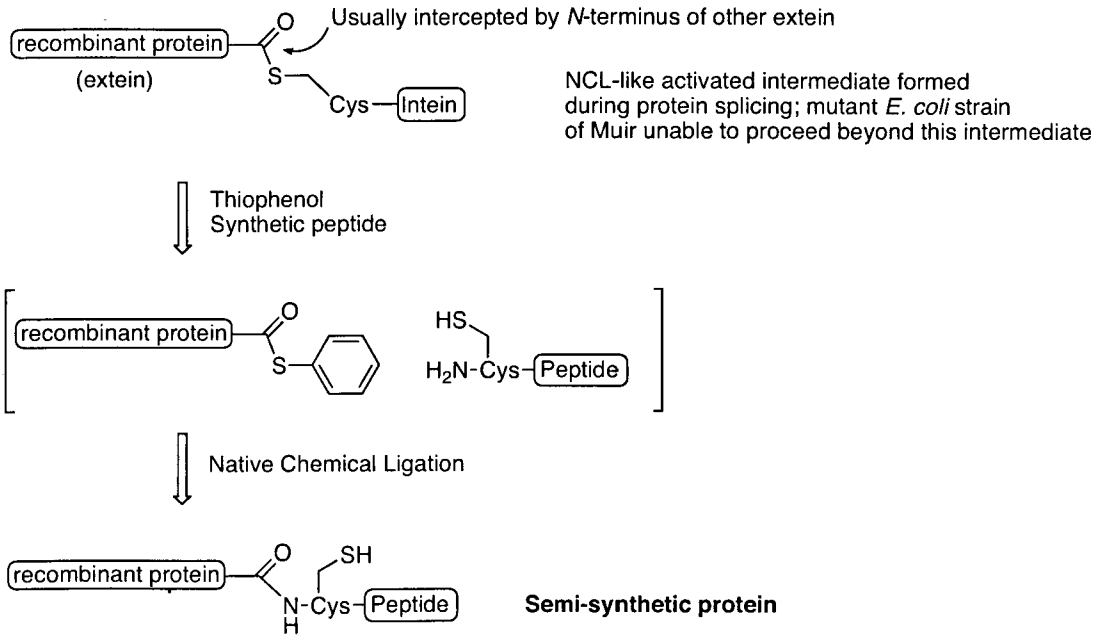
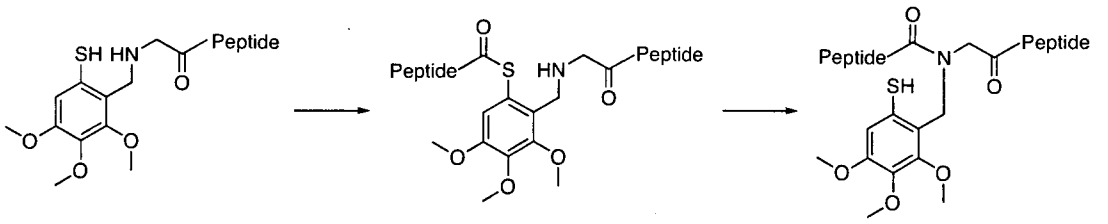


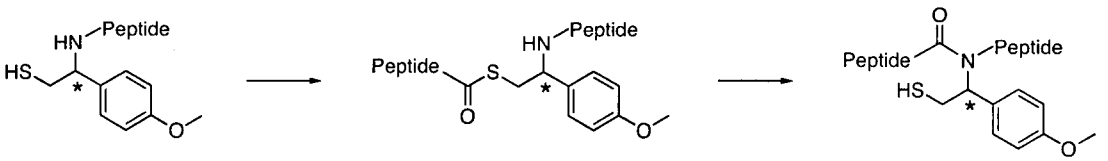
Figure 1.41 Expressed protein ligation

Using expressed protein ligation, Muir successfully ligated a phosphotyrosine peptide with a 50 kDa protein, C-terminal Src kinase, to create a semi-synthetic protein far beyond the size domain previously accessible by standard NCL chemistry.[160]

Traditional NCL is not only dependent on access to thioester peptides, but also requires the *N*-terminus residue of the partner peptide to be cysteine. Both Kent and Dawson moved to overcome the limitation of cysteine-dependence by attaching cleavable thiol auxiliary units to the *N*-terminus of a peptide to imitate the function of a cysteine side-chain. The chemical premise of thiol exchange and acyl shift are preserved, but the amino acid side-chain is no longer involved in the reaction (Figure 1.42).



30 (Dawson auxiliary)



31 (Kent auxiliary)

Figure 1.42 Auxiliaries used to mediate NCL

30 can be cleaved using TFA after acyl transfer has occurred; the benzylic amine bond is rendered sufficiently labile for protolysis upon formation of the amide. **31** requires treatment with HF for cleavage, although if an additional methoxy substituent is placed at the 2 position on the ring, TFA can be used instead. The Kent auxiliary was used successfully in a series of test-ligations and later in the synthesis of a wild-type cytochrome b562 protein[161] and an unnatural variant (Met → SelenoMet substitution at position 7). The 106 residue protein was synthesised in two halves, cyt b562[1-63] and cyt b562[64-106], and converted to the benzyl thioester and auxiliary-linked units respectively. Ligation under aqueous conditions (pH 7.5 buffered) proceeded to give the target protein *via* formation of a His-Gly junction. Dawson applied his auxiliary during the synthesis of the SH3 domain of α-spectrin.[161] The 62 amino acid unit was divided into the fragments α-spectrin[1-27] and α-spectrin[28-62] and converted to the thiophenol thioester and auxiliary linked derivatives, respectively. Ligation across a Lys-Gly junction and subsequent TFA cleavage of the auxiliary proceeded smoothly to give the SH3 domain of α-spectrin.

Auxiliary-mediated NCL has addressed some of the problems associated with convergent peptide synthesis, but there remains scope for improvement. The Kent auxiliary is prepared as a racemate to give diastereomeric peptides when attached; this complicates HPLC and NMR analysis. Both auxiliaries are sensitive to steric bulk about the ligation junction and require one of the ligating residues to be a Gly to effect an

efficient ligation. The synthesis of thioesters is difficult, and as of yet there is no good method to synthesise a protected thioester that would allow multicomponent synthesis.[162]

1.4 Project Aims

The chemical synthesis of proteins holds the potential to access structures not accessible using biological methods. However current chemical methodologies are limited to the synthesis of small proteins of low molecular weight that are not representative of the majority of protein targets. It is clear that a new methodology is required that can enable the *multi-component* convergent synthesis of polypeptides.

This project aims to overcome the size restrictions of current techniques by developing methodology that will enable the programmed ordering (to arrange in a desired sequence) and simultaneous ligation of multiple unprotected polypeptides (Figure 1.43).

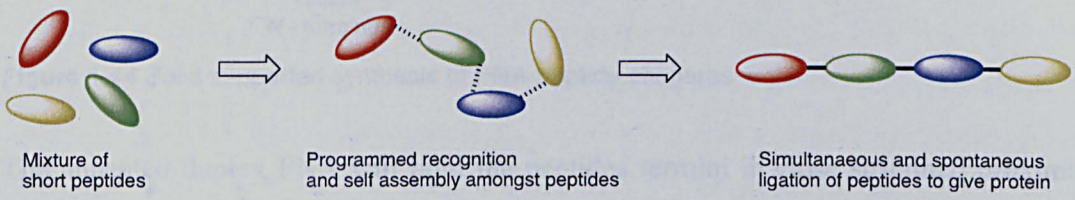


Figure 1.43 Idealised schematic of peptide self-assembly and ligation

It is thought that the simultaneous ligation of several smaller polypeptides will avoid solubility and ligation difficulties encountered with the stepwise assembly of a fewer larger peptides.

1.4.1 Concept and theory

In order to effect the self-assembly of peptides into a specified sequence, the peptides will have to be tagged with programmable auxiliaries. These units will template the directed sequencing and auto-ligation of the peptides. The auxiliaries will ideally be removable under mild chemical conditions.

Liu *et al* have demonstrated that DNA can orchestrate a programmed reaction exclusively between two specific reactants within a diverse mixture of mutually reactive

partners (*vide supra*). However, DNA is not suited for use in this project as the conditions for SPPS and DNA synthesis are very different; creating a peptide-DNA chimera would be non-trivial.[72, 163] An alternative nucleic acid based directing group that is fully compatible with SPPS protocols is PNA. PNA has been shown to be an excellent DNA mimic that exhibits enhanced selectivity and thermodynamic stability (*vide supra*) when complexed to oligonucleotides.

It is proposed that attaching PNA *via* some linker unit to a peptide will enable the hybrid to selectively anneal with a peptide bearing the complementary PNA strand. The peptide – PNA chimeras will be synthesised upon a solid support using standard Fmoc chemistry (Figure 1.44).

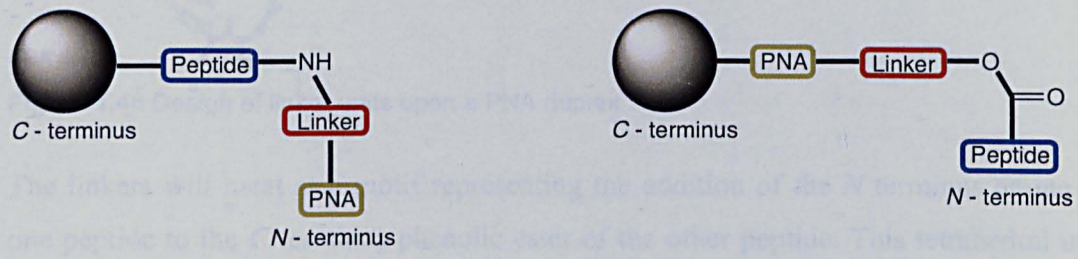


Figure 1.44 Solid supported synthesis of PNA-peptide chimeras

The annealed duplex PNA will hold the peptides termini in close sustained proximity driving a pseudo intramolecular acylation to form a native peptide bond. Tagging PNA strands on the *C* and *N* terminus of a peptide should allow the cascade ligation of peptides at both junctions; this would allow the simultaneous ligation of multiple peptides (Figure 1.45).

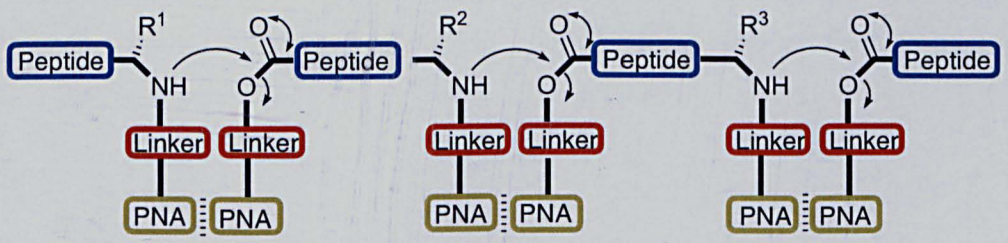


Figure 1.45 PNA directed ligation of peptides

The ligation reaction has been designed such that it will be proximity driven. The *C*-terminus peptide will be joined to the PNA *via* a phenolic ester linkage; it is believed

that this will spontaneously react with the *N*-terminus of a partner peptide when held close together to form the native peptide bond. The linker units will be designed with the assistance of computer modelling. A crystal structure of a PNA duplex will be used as a scaffold upon which the linkers will be 'grown' from either terminus toward a partially ligated peptide (Figure 1.46).[164]

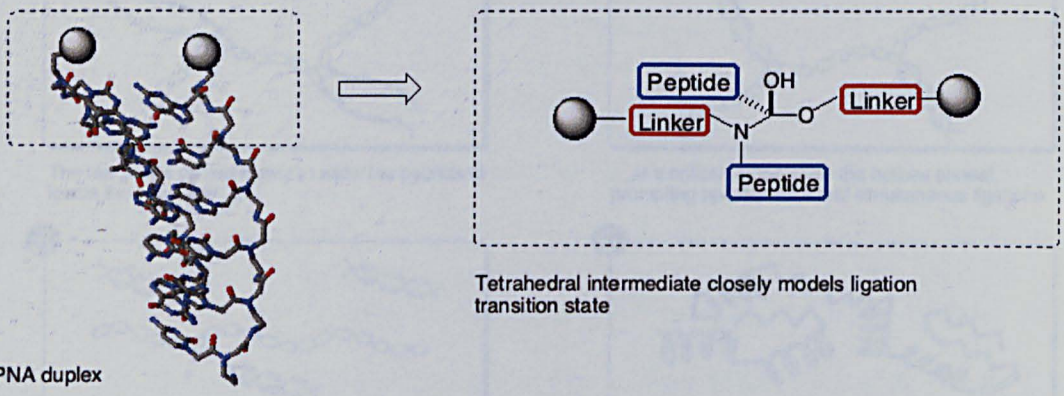


Figure 1.46 Design of linker units upon a PNA duplex

The linkers will meet at a motif representing the addition of the *N* terminus amine of one peptide to the *C* terminus phenolic ester of the other peptide. This tetrahedral unit models the transition state of the ligation. Designing the linkers to occupy the space between the PNA helix and this unit will maximise the probability of a proximity-driven ligation upon duplex formation; at equilibrium, the reacting centres will be at the transition-state reaction coordinate. Ultimately, a mixture of unprotected peptides tagged with the appropriate PNA directing groups will be spontaneously ordered and ligated to yield a target protein (Figure 1.47).

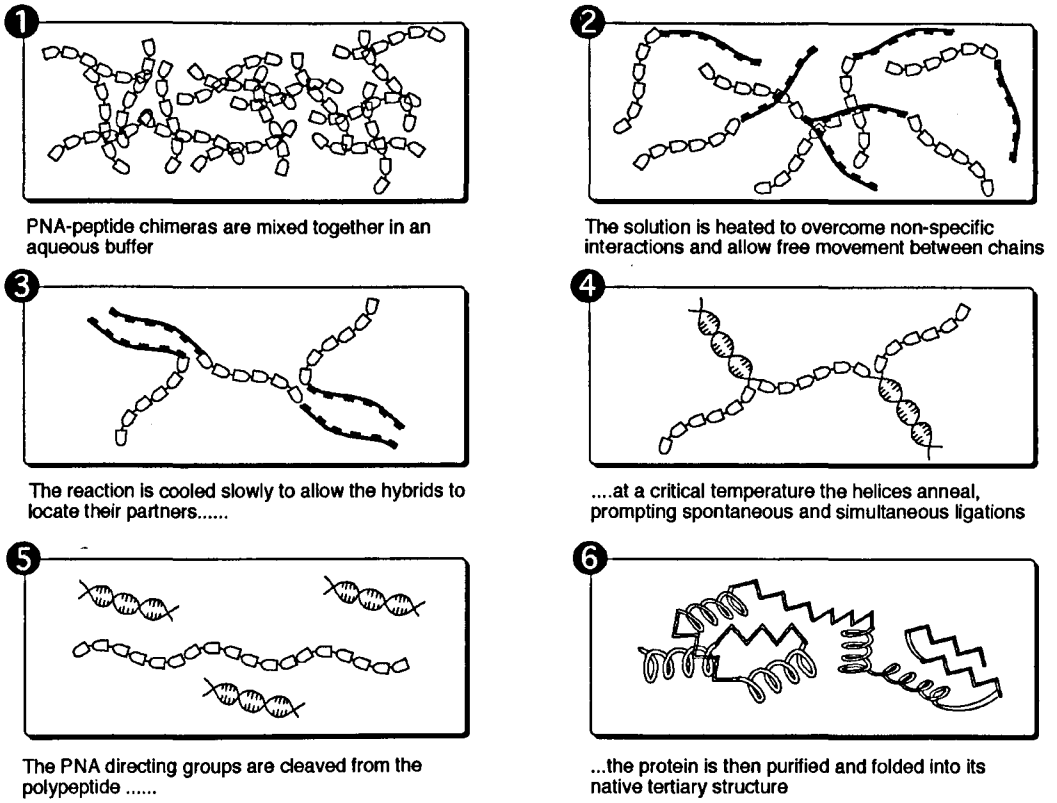


Figure 1.47 Chemical total protein synthesis directed by PNA

The proposed methodology provides a number of advantages compared to the prior art. No noxious and difficult thioester/thiol chemistry is required and the methodology will also deliver a much higher degree of convergence; potential targets could be much larger than those previously achieved. Furthermore there is broad scope for the introduction of unnatural components into a protein eg probes, non-genetic amino acids, carbohydrates

The success of DTS by Liu *et al.* and the nature of the proximity-driven acylation in ribosomal peptide synthesis supports strongly the hypothesis that PNA-directed peptide assembly is a fully viable strategy for protein synthesis.

1.5 References

[1] L. L. Cavalli-Sforza, *Nature Reviews Genetics* **2005**, *6*, 333.
[2] S. Jenni, N. Ban, *Current Opinion in Structural Biology* **2003**, *13*, 212.
[3] C. R. Harwood, A. Wipat, *Molecular Medical Microbiology* **2002**, *1*, 321.
[4] M. Mahoney Joseph, R. Shukla, R. A. Marshall, M. Beatty Alicia, J. Zajicek, D. Smith Bradley, *The Journal of organic chemistry* **2002**, *67*, 1436.
[5] A. Mateo-Alonso, M. Prato, *Tetrahedron* **2006**, *62*, 2003.
[6] D. N. Reinhoudt, F. De Jong, H. P. M. Tomassen, *Tetrahedron Lett.* **1979**, 2067.
[7] E. J. Kang, E. Lee, *Chem. Rev.* **2005**, *105*, 4348.
[8] A. Fuerstner, J. Ruiz-Caro, H. Prinz, H. Waldmann, *J. Org. Chem.* **2004**, *69*, 459.
[9] A. Fuerstner, M. Albert, J. Mlynarski, M. Matheu, E. DeClercq, *JACS* **2003**, *125*, 13132.
[10] C. A. Schalley, *J. Phys. Org. Chem.* **2004**, *17*, 967.
[11] J.-P. Collin, V. Heitz, J.-P. Sauvage, *Top. Curr. Chem.* **2005**, *262*, 29.
[12] J.-P. Collin, D. Jouvenot, M. Koizumi, J.-P. Sauvage, *Eur. J. Inorg. Chem.* **2005**, 1850.
[13] J.-P. Sauvage, *Chemical Communications (Cambridge, United Kingdom)* **2005**, 1507.
[14] C. O. Dietrich-Buchecker, J. P. Sauvage, J. M. Kern, *JACS* **1984**, *106*, 3043.
[15] C. Dietrich-Buchecker, J.-P. Sauvage, *Chemical Communications (Cambridge)* **1999**, 615.
[16] M. Weck, B. Mohr, J.-P. Sauvage, R. H. Grubbs, *J. Org. Chem.* **1999**, *64*, 5463.
[17] G. R. Desiraju, *Angew. Chem. Int. Ed.* **1995**, *34*, 2311.
[18] F. Arico, J. D. Badjic, S. J. Cantrill, A. H. Flood, K. C. F. Leung, Y. Liu, J. F. Stoddart, *Top. Curr. Chem.* **2005**, *249*, 203.
[19] F. Arico, T. Chang, S. J. Cantrill, S. I. Khan, J. F. Stoddart, *Chem. Eur. J.* **2005**, *11*, 4655.
[20] C. A. Schalley, *Angew. Chem. Int. Ed.* **2004**, *43*, 4399.
[21] M. M. Rozenman, D. R. Liu, *ChemBioChem* **2006**, *7*, 253.
[22] K. Sakurai, T. M. Snyder, D. R. Liu, *JACS* **2005**, *127*, 1660.
[23] C. T. Calderone, J. W. Puckett, Z. J. Gartner, D. R. Liu, *Angew. Chem. Int. Ed.* **2002**, *41*, 4104.
[24] S. Ficht, A. Mattes, O. Seitz, *JACS* **2004**, *126*, 9970.
[25] C. T. Calderone, D. R. Liu, *Angew. Chem. Int. Ed.* **2005**, *44*, 7383.
[26] M. W. Kanan, M. M. Rozenman, K. Sakurai, T. M. Snyder, D. R. Liu, *Nature (London, United Kingdom)* **2004**, *431*, 545.
[27] X. Li, D. R. Liu, *JACS* **2003**, *125*, 10188.
[28] D. Summerer, A. Marx, *Angew. Chem. Int. Ed.* **2002**, *41*, 89.
[29] S. M. Du, B. D. Stollar, N. C. Seeman, *JACS* **1995**, *117*, 1194.
[30] M. Koppitz, P. E. Nielsen, L. E. Orgel, *J. Am. Chem. Soc. Full Journal Title:Journal of the American Chemical Society* **1998**, *120*, 4563.
[31] J. G. Schmidt, L. Christensen, P. E. Nielsen, L. E. Orgel, *Nucleic Acids Res. Full Journal Title:Nucleic Acids Research* **1997**, *25*, 4792.
[32] T. Wu, L. E. Orgel, *J. Am. Chem. Soc. Full Journal Title:Journal of the American Chemical Society* **1992**, *114*, 5496.
[33] A. R. Hill, Jr., S. Kumar, N. J. Leonard, L. E. Orgel, *J. Mol. Evol. Full Journal Title:Journal of Molecular Evolution* **1988**, *27*, 91.

[34] T. Haertle, L. E. Orgel, *J. Mol. Biol. Full Journal Title:Journal of Molecular Biology* **1986**, *188*, 77.
[35] A. W. Schwartz, L. E. Orgel, *J. Mol. Evol. Full Journal Title:Journal of Molecular Evolution* **1985**, *21*, 299.
[36] A. Luther, R. Brandsch, G. von Kiedrowski, *Nature Full Journal Title:Nature* **1998**, *396*, 245.
[37] G. von Kiedrowski, *Nature Full Journal Title:Nature* **1996**, *381*, 20.
[38] D. Sievers, G. von Kiedrowski, *Nature Full Journal Title:Nature* **1994**, *369*, 221.
[39] J. C. Leitzel, D. G. Lynn, *Chem. Rec. Full Journal Title:Chemical Record* **2001**, *1*, 53.
[40] J. Ye, Y. Gat, D. G. Lynn, *Angew. Chem., Int. Ed. Full Journal Title:Angewandte Chemie, International Edition* **2000**, *39*, 3641.
[41] Y. Gat, D. G. Lynn, *Templated Org. Synth. Full Journal Title:Templated Organic Synthesis* **2000**, 133.
[42] P. Luo, J. C. Leitzel, Z.-Y. J. Zhan, D. G. Lynn, *J. Am. Chem. Soc. Full Journal Title:Journal of the American Chemical Society* **1998**, *120*, 3019.
[43] Z.-Y. J. Zhan, D. G. Lynn, *J. Am. Chem. Soc. Full Journal Title:Journal of the American Chemical Society* **1997**, *119*, 12420.
[44] Y. Xu, N. B. Karalkar, E. T. Kool, *Nat. Biotechnol.* **2001**, *19*, 148.
[45] H. Wallrabe, A. Periasamy, *Curr. Opin. Biotechnol.* **2005**, *16*, 19.
[46] I. Rasnik, S. A. McKinney, T. Ha, *Acc. Chem. Res.* **2005**, *38*, 542.
[47] K. Kikuchi, H. Takakusa, T. Nagano, *TrAC, Trends in Analytical Chemistry* **2004**, *23*, 407.
[48] E. A. Jares-Erijman, T. M. Jovin, *Nat. Biotechnol.* **2003**, *21*, 1387.
[49] A. Dietrich, V. Buschmann, C. Muller, M. Sauer, *Reviews in Molecular Biotechnology* **2002**, *82*, 211.
[50] S. Nampalli, W. Zhang, S. Rao, S. Kumar, (Amersham Biosciences Corp., USA). Application: US
US, **2004**, p. 22 pp.
[51] D. Belik, H.-D. Josel, R. Herrmann, B. Koenig, F. Mueller, (Roche Diagnostics Gmbh, Germany; F.Hoffmann-La Roche Ag). Application: WO
WO, **2002**, p. 42 pp.
[52] T. S. Rao, S. Nampalli, K. Lavrenov, W. Zhang, H. Xiao, J. Nelson, S. Kumar, *Nucleosides, Nucleotides & Nucleic Acids* **2001**, *20*, 673.
[53] Z. J. Gartner, R. Grubina, C. T. Calderone, D. R. Liu, *Angew. Chem. Int. Ed.* **2003**, *42*, 1370.
[54] X. Li, D. R. Liu, *Angew. Chem. Int. Ed.* **2004**, *43*, 4848.
[55] Z. J. Gartner, B. N. Tse, R. Grubina, J. B. Doyon, T. M. Snyder, D. R. Liu, *Science (Washington, DC, United States)* **2004**, *305*, 1601.
[56] X. Li, Z. J. Gartner, B. N. Tse, D. R. Liu, *JACS* **2004**, *126*, 5090.
[57] T. Calderone Christopher, R. Liu David, *Angewandte Chemie (International ed. in English)* **2005**, *44*, 7383.
[58] T. M. Snyder, D. R. Liu, *Angew. Chem. Int. Ed.* **2005**, *44*, 7379.
[59] M. Egholm, O. Buchardt, P. E. Nielsen, R. H. Berg, *JACS* **1992**, *114*, 1895.
[60] P. Wittung, P. Nielsen, B. Norden, *JACS* **1996**, *118*, 7049.
[61] P. E. Nielsen, M. Egholm, O. Buchardt, *Journal of Molecular Recognition* **1994**, *7*, 165.
[62] T. Bentin, H. J. Larsen, E. Nielsen Peter, *Biochemistry* **2003**, *42*, 13987.

[63] T. Ratilainen, A. Holmen, E. Tuite, P. E. Nielsen, B. Norden, *Biochemistry* **2000**, *39*, 7781.
[64] N. Sugimoto, N. Satoh, K. Yamamoto, *Nucleic Acids Symposium Series* **1999**, *42*, 93.
[65] F. P. Schwarz, S. Robinson, J. M. Butler, *Nucleic Acids Res.* **1999**, *27*, 4792.
[66] G. Haaiima, H. F. Hansen, L. Christensen, O. Dahl, P. E. Nielsen, *Nucleic Acids Res.* **1997**, *25*, 4639.
[67] A. Sen, E. Nielsen Peter, *Biophys. J.* **2006**, *90*, 1329.
[68] D. F. Doyle, D. A. Braasch, C. G. Simmons, B. A. Janowski, D. R. Corey, *Biochemistry* **2001**, *40*, 53.
[69] S. Sforza, R. Corradini, S. Ghirardi, A. Dossena, R. Marchelli, *Eur. J. Org. Chem.* **2000**, 2905.
[70] G. L. Igloi, *Proceedings of the National Academy of Sciences of the United States of America* **1998**, *95*, 8562.
[71] A. Abibi, E. Protozanova, V. V. Demidov, M. D. Frank-Kamenetskii, *Biophys. J.* **2004**, *86*, 3070.
[72] Z. Bajor, G. Sagi, Z. Tegyey, F. Kraicsovits, *Nucleosides, Nucleotides & Nucleic Acids* **2003**, *22*, 1963.
[73] M. Borgatti, A. Romanelli, M. Saviano, C. Pedone, I. Lampronti, L. Breda, C. Nastruzzi, N. Bianchi, C. Mischiati, R. Gambari, *Oncology Research* **2003**, *13*, 279.
[74] Z. Chu, K.-L. Liu, *Youji Huaxue* **2005**, *25*, 254.
[75] P. E. Nielsen, *Lett. Pept. Sci.* **2004**, *10*, 135.
[76] O. Seitz, *Highlights in Bioorganic Chemistry* **2004**, 327.
[77] X. Zhang, J. Han, *Zhongguo Shengwu Gongcheng Zazhi* **2003**, *23*, 30.
[78] A. Ray, B. Norden, *FASEB J.* **2000**, *14*, 1041.
[79] L. Tan, Y. Li, T. J. Drake, L. Moroz, K. Wang, J. Li, A. Munteanu, C. J. Yang, K. Martinez, W. Tan, *Analyst (Cambridge, United Kingdom)* **2005**, *130*, 1002.
[80] F. Wu, P. Della-Latta, S. Tyagi, F. R. Kramer, *Mol. Microbiol.* **2004**, 285.
[81] J. A. M. Vet, B. J. M. Van der Rijt, H. Blom, *Expert Review of Molecular Diagnostics* **2002**, *2*, 77.
[82] A. E. Marras Salvatore, S. Tyagi, R. Kramer Fred, *Clinica chimica acta; international journal of clinical chemistry* **2006**, *363*, 48.
[83] A. Tsourkas, G. Bao, *Briefings in functional genomics & proteomics* **2003**, *1*, 372.
[84] S. Kamihira, K. Tsukazaki, *Rinsho Byori Rebyu, Tokushugo* **2004**, *128*, 162.
[85] J. M. Levsky, R. H. Singer, *J. Cell Sci.* **2003**, *116*, 2833.
[86] A. K. Raap, *Mutation Research* **1998**, *400*, 287.
[87] B. D. Alexander, E. D. Ashley, L. B. Reller, S. D. Reed, *Diagnostic Microbiology and Infectious Disease* **2006**, *54*, 277.
[88] M. J. Lehtola, E. Torvinen, I. T. Miettinen, C. W. Keevil, *Applied and Environmental Microbiology* **2006**, *72*, 848.
[89] W. R. Teertstra, L. G. Lugones, H. A. B. Woesten, *Fungal Genetics and Biology* **2004**, *41*, 1099.
[90] H. Stender, *Expert Review of Molecular Diagnostics* **2003**, *3*, 649.
[91] M. Radwanska, S. Magez, H. Perry-O'Keefe, H. Stender, J. Coull, J. M. Sternberg, P. Buscher, J. J. Hyldig-Nielsen, *J. Clin. Microbiol.* **2002**, *40*, 4295.
[92] B. Williams, H. Stender, J. M. Coull, *Methods in Molecular Biology (Totowa, NJ, United States)* **2002**, *208*, 181.

[93] J.-D. Luo, E.-C. Chan, C.-L. Shih, T.-L. Chen, Y. Liang, T.-L. Hwang, C.-C. Chiou, *Nucleic Acids Res.* **2006**, *34*, e12/1.
[94] T. Kirishima, M. Minami, T. Okanoue, *Saibo* **2003**, *35*, 244.
[95] D. G. Murdock, D. C. Wallace, *Methods in Molecular Biology (Totowa, NJ, United States)* **2002**, *208*, 145.
[96] W. Cai, H. Aburatani, V. P. Stanton, Jr., D. E. Housman, Y.-K. Wang, D. C. Schwartz, *Proceedings of the National Academy of Sciences of the United States of America* **1995**, *92*, 5164.
[97] Y.-K. Wang, E. J. Huff, D. C. Schwartz, *Proceedings of the National Academy of Sciences of the United States of America* **1995**, *92*, 165.
[98] M. D. Frank-Kamenetskii, V. V. Demidov, *Peptide Nucleic Acids (2nd Edition)* **2004**, 187.
[99] A. G. Veselkov, V. V. Demidov, P. E. Nielsen, M. D. Frank-Kamenetskii, *Nucleic Acids Res.* **1996**, *24*, 2483.
[100] P. E. Nielsen, *Current Opinion in Molecular Therapeutics* **2000**, *2*, 282.
[101] L. Good, P. E. Nielsen, *Current Issues in Molecular Biology* **1999**, *1*, 111.
[102] P. E. Nielsen, *Current Opinion in Structural Biology* **1999**, *9*, 353.
[103] T. Shiraishi, *Seibutsu Kogaku Kaishi* **2004**, *82*, 152.
[104] K. Kaihatsu, A. Janowski Bethany, R. Corey David, *Chemistry & biology* **2004**, *11*, 749.
[105] J. Temsamani, G. S. Pari, P. Guinot, *Expert opinion on investigational drugs* **1997**, *6*, 1157.
[106] M. A. Polis, *Book of Abstracts, 213th ACS National Meeting, San Francisco, April 13-17* **1997**, CARB.
[107] K. Kilk, U. Langel, *Methods in Molecular Biology (Totowa, NJ, United States)* **2005**, *298*, 131.
[108] J. Oehlke, G. Wallukat, Y. Wolf, A. Ehrlich, B. Wiesner, H. Berger, M. Bienert, *Eur. J. Biochem.* **2004**, *271*, 3043.
[109] J. J. Turner, G. D. Ivanova, B. Verbeure, D. Williams, A. A. Arzumanov, S. Abes, B. Lebleu, M. J. Gait, *Nucleic Acids Res.* **2005**, *33*, 6837.
[110] N. Nekhotiaeva, S. K. Awasthi, P. E. Nielsen, L. Good, *Molecular Therapy* **2004**, *10*, 652.
[111] R. Villa, M. Folini, S. Lualdi, S. Veronese, M. G. Daidone, N. Zaffaroni, *FEBS Lett.* **2000**, *473*, 241.
[112] F. W. Rasmussen, N. Bendifallah, V. Zachar, T. Shiraishi, T. Fink, P. Ebbesen, P. E. Nielsen, U. Koppelhus, *Oligonucleotides* **2006**, *16*, 43.
[113] D. R. Corey, *Lett. Pept. Sci.* **2004**, *10*, 347.
[114] E. Kawada, S. Yano, H. Kojima, (National Institute of Advanced Industrial Science and Technology, Japan). Application: JP
JP, **2004**, p. 9 pp.
[115] D. A. Braasch, D. R. Corey, *Methods in Molecular Biology (Totowa, NJ, United States)* **2002**, *208*, 211.
[116] G. Zon, (Applera Corporation, USA). Application: US
US, **2006**, p. 22 pp.
[117] E. Richelson, B. M. Tyler, B. M. Cusack, C. L. Douglas, K. Jansen, (Mayo Foundation for Medical Education and Research, USA). Application: US
US, **2006**, pp. 52 pp
[118] A.-A. O. Mohammed, T. Yamazaki, M. Otsuka, (Mitsubishi Rayon Co., Ltd., Japan). Application: JP
JP, **2003**, p. 18 pp.

[119] K. Braun, I. Braun, J. Debus, R. Pipkorn, W. Waldeck, (Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts, Germany). Application: WO
WO, **2003**, p. 34 pp.
[120] J. P. Behr, P. Belguise-Valladier, M.-A. Zanta, (Universite Louis Pasteur de Strasbourg, Fr.). Application: EP
EP, **2000**, p. 30 pp.
[121] R. P. Stock, A. Olvera, R. Sanchez, A. Saralegui, S. Scarfi, R. Sanchez-Lopez, M. A. Ramos, L. C. Boffa, U. Benatti, A. Alagon, *Nat. Biotechnol.* **2001**, *19*, 231.
[122] B. L. Nilsson, M. B. Soellner, R. T. Raines, *Annu. Rev. Biophys. Biomol. Struct.* **2005**, *34*, 91.
[123] G. Cotton, *Speciality Chemicals Magazine* **2004**, *24*, 32.
[124] R. David, M. P. O. Richter, A. G. Beck-Sickinger, *Eur. J. Biochem.* **2004**, *271*, 663.
[125] G. Casi, D. Hilvert, *Current Opinion in Structural Biology* **2003**, *13*, 589.
[126] K. Josephson, M. C. T. Hartman, J. W. Szostak, *JACS* **2005**, *127*, 11727.
[127] M. Ailenberg, N. M. Goldenberg, M. Silverman, *BMC Biotechnology* **2005**, *5*, No pp given.
[128] A. N. Warrens, M. D. Jones, R. I. Lechler, *Gene* **1997**, *186*, 29.
[129] M. Selmer, C. M. Dunham, F. V. I. V. Murphy, A. Weixlbaumer, S. Petry, A. C. Kelley, J. R. Weir, V. Ramakrishnan, *Science (Washington, DC, United States)* **2006**, *313*, 1935.
[130] B. T. Wimberly, D. E. Brodersen, W. M. Clemons, Jr., R. J. Morgan-Warren, A. P. Carter, C. Vonrhein, T. Hartsch, V. Ramakrishnan, *Nature (London)* **2000**, *407*, 327.
[131] J. M. Ogle, A. P. Carter, V. Ramakrishnan, *Trends Biochem. Sci* **2003**, *28*, 259.
[132] J. Xie, P. G. Schultz, *Curr. Opin. Chem. Biol.* **2005**, *9*, 548.
[133] F. Tian, M.-L. Tsao, P. G. Schultz, *JACS* **2004**, *126*, 15962.
[134] J. C. Anderson, N. Wu, S. W. Santoro, V. Lakshman, D. S. King, P. G. Schultz, *Proceedings of the National Academy of Sciences of the United States of America* **2004**, *101*, 7566.
[135] T. Kimmerlin, D. Seebach, *Journal of Peptide Research* **2005**, *65*, 229.
[136] R. B. Merrifield, *JACS* **1963**, *85*, 2149.
[137] S. Kent, *J. Pept. Sci.* **2003**, *9*, 574.
[138] E. Falb, T. Yechezkel, Y. Salitra, C. Gilon, *Journal of Peptide Research* **1999**, *53*, 507.
[139] P. Raman, S. S. Stokes, Y. M. Angell, G. R. Flentke, D. H. Rich, *J. Org. Chem.* **1998**, *63*, 5734.
[140] K. Akaji, N. Kuriyama, T. Kimura, Y. Fujiwara, Y. Kiso, *Pept. 1992, Proc. Eur. Pept. Symp., 22nd* **1993**, 220.
[141] E. Frerot, J. Coste, A. Pantaloni, M. N. Dufour, P. Jouin, *Tetrahedron* **1991**, *47*, 259.
[142] C. Garcia-Echeverria, *Lett. Pept. Sci.* **1996**, *2*, 369.
[143] E. Oliveira, R. Marchetto, G. N. Jubilut, A. C. M. Paiva, C. R. Nakaie, *Pept.: Chem. Biol., Proc. Am. Pept. Symp., 12th* **1992**, 569.
[144] W. A. Klis, J. M. Stewart, *Pept.: Chem., Struct. Biol., Proc. Am. Pept. Symp., 11th* **1990**, 904.
[145] Z. Miao, J. P. Tam, *JACS* **2000**, *122*, 4253.
[146] J. P. Tam, Z. Miao, *JACS* **1999**, *121*, 9013.

[147] T. Wieland, E. Bokelmann, L. Bauer, H. U. Lang, H. Lau, W. Schafer, *Ann. Full Journal Title:* **1953**, *583*, 129.
[148] D. S. Kemp, N. G. Galakatos, S. Dranginis, C. Ashton, N. Fotouhi, T. P. Curran, *J. Org. Chem.* **1986**, *51*, 3320.
[149] D. S. Kemp, N. G. Galakatos, B. Bowen, K. Tan, *J. Org. Chem.* **1986**, *51*, 1829.
[150] D. S. Kemp, D. J. Kerkman, *Tetrahedron Lett.* **1981**, *22*, 185.
[151] D. S. Kemp, S.-L. Leung, D. J. Kerkman, *Tetrahedron Lett.* **1981**, *22*, 181.
[152] N. Fotouhi, N. G. Galakatos, D. S. Kemp, *J. Org. Chem. Full Journal Title:Journal of Organic Chemistry* **1989**, *54*, 2803.
[153] P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. Kent, *Science* **1994**, *266*, 776.
[154] G. G. Kochendoerfer, D. Salom, J. D. Lear, R. Wilk-Orescan, S. B. Kent, W. F. DeGrado, *Biochemistry* **1999**, *38*, 11905.
[155] W. Lu, M. A. Starovasnik, S. B. Kent, *FEBS Lett.* **1998**, *429*, 31.
[156] Y. Shin, K. A. Winans, B. J. Backes, S. B. H. Kent, J. A. Ellman, C. R. Bertozzi, *J. Am. Chem. Soc. Full Journal Title:Journal of the American Chemical Society* **1999**, *121*, 11684.
[157] B. J. Backes, J. A. Ellman, *J. Org. Chem. Full Journal Title:Journal of Organic Chemistry* **1999**, *64*, 2322.
[158] A. Nagalingam, *University of Leeds* **2007**.
[159] R. Xu, B. Ayers, D. Cowburn, T. W. Muir, *Proc Natl Acad Sci U S A Full Journal Title:Proceedings of the National Academy of Sciences of the United States of America* **1999**, *96*, 388.
[160] T. W. Muir, D. Sondhi, P. A. Cole, *Proc Natl Acad Sci U S A Full Journal Title:Proceedings of the National Academy of Sciences of the United States of America* **1998**, *95*, 6705.
[161] D. W. Low, M. G. Hill, M. R. Carrasco, S. B. Kent, P. Botti, *Proceedings of the National Academy of Sciences of the United States of America* **2001**, *98*, 6554.
[162] Y. Shin, K. A. Winans, B. J. Backes, S. B. H. Kent, J. A. Ellman, C. R. Bertozzi, *JACS* **1999**, *121*, 11684.
[163] V. Esposito, A. Randazzo, A. Messere, A. Galeone, L. Petraccone, C. Giancola, G. Piccialli, L. Mayol, *Eur. J. Org. Chem.* **2003**, 3364.
[164] H. Rasmussen, J. S. Kastrup, J. N. Nielsen, J. M. Nielsen, P. E. Nielsen, *Nat. Struct. Biol.* **1997**, *4*, 98.

Results and Discussion

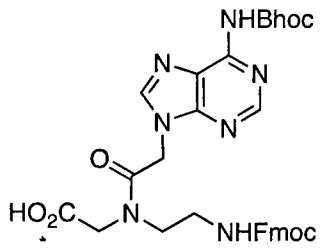
Chapter 2

Chapter 2 - Results and Discussion

The first objective for the project was to establish reliable synthetic routes for the synthesis of the four PNA monomers and prepare them in multigram quantities.

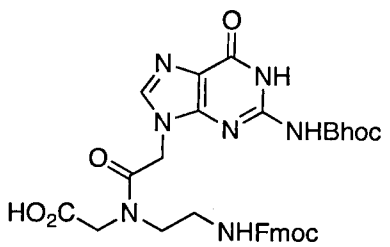
2.1 Synthesis of PNA monomers

PNA monomers are commercially available as the Bhoc protected nucleobase / Fmoc protected backbone monomers (Figure 2.1). This orthogonal protection Scheme allows the use of standard SPPS reagents and materials for the construction of PNA oligomers under relatively mild conditions. Fmoc protection (ie base labile) of the aminoethyl glycine (AEG) amino group is vital for the PNA to be useful in this project. Reagents used for PNA oligomerisation must not repeatedly or excessively expose the growing polymer to harsh acidic conditions; to do so would curtail the potential to involve other (relatively fragile) biological motifs, such as carbohydrates, within the scope of the methodology.



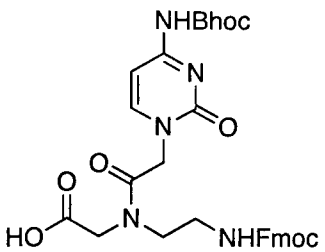
32

A-PNA monomer £239 per gram



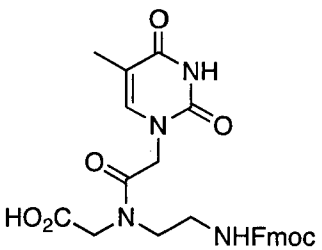
33

G-PNA monomer £266 per gram



34

C-PNA monomer £239 per gram



35

T-PNA monomer £211 per gram

Bhoc :

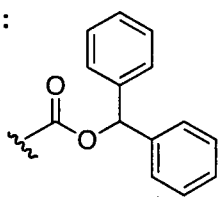
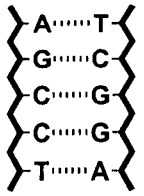


Fig 2.1 Commercially available PNA monomers

Multigram quantities of each monomer are required for this project and so the PNA must be prepared in house as the monomers are prohibitively expensive. This is more

readily appreciated if the cost of constructing a small PNA duplex is considered (Figure 2.2).



5 equivalents of monomer are used per coupling
To synthesise 0.05 mmol of the duplex, 0.25mmol of each monomer is required - 0.50 mmol (A) : 0.50 mmol (T) : 0.75 mmol (G) : 0.75 mmol (C)
Price per mmol of Monomer - (A) £173 : (T) £107 : (G) £197 : (C) £167
TOTAL COST OF PNA USED TO MAKE 50 μMol OF DUPLEX - £413

Figure 2.2 Cost of PNA monomers used in synthesis of a small PNA duplex

The example highlighted in Figure 2.2 is scaled for use in a single test reaction. If the reaction were scaled to preparative amounts (1 mmol or greater), the cost would increase 20 fold, or greater, per duplex.

Clearly, there was a significant financial cost that required that the PNA monomers to be synthesised in-house. However, it should be noted that synthesis of the linker units required the knowledge of how to synthesise PNA monomers because they (the linkers) are modified PNA monomers themselves. Without being able to synthesise PNA therefore, the linker units could not have been made and the project could not have been attempted. There was therefore a synthetic requirement that was as equally compelling as the financial premise for the synthesis of the PNA monomers.

Syntheses of the commercially available PNA monomers have not been published, however PNA monomers bearing different amine protecting groups have been described.[1-7]

Thompson *et al.* published the synthesis of PNA monomers that used a CBz (nucleobase)/Fmoc (backbone) protection Scheme and these were thought suitable for our needs.[6] These monomers retain the critical Fmoc protection for the AEG, but exchange the mild acidic conditions of Bhoc Cleavage (10-20% TFA in CH_2Cl_2) for neutral hydrogenation removal of the CBz group.

Thompson details full syntheses of the backbone and nucleobase acids however Dueholm has reported the synthesis of PNA monomers bearing a CBz (nucleobase)/BOC (backbone) protection Scheme.[3] The BOC protected backbone is unsuitable for this project but the reported syntheses of the CBz-protected nucleobases

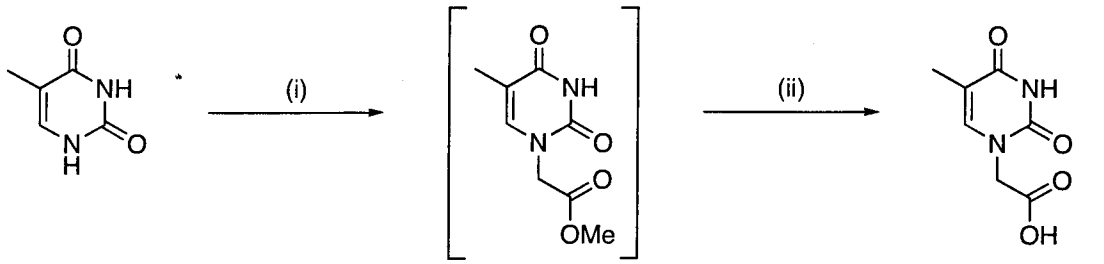
are more efficient than those of Thompson. The synthesis of the monomers was therefore comprised from a combination of the Thompson and Dueholm procedures to maximise the efficiency of the procedures.

The synthesis of PNA monomers was broken down into three parts – synthesis of each nucleobase acetic acid, synthesis of the backbone and coupling of each nucleobase to the backbone.

2.1.1 Synthesis of CBz Protected Nucleobase acetic acids.

2.1.1.1 Thyminyl Acetic Acid

Thymine is unique among the four nucleobases in that it lacks an exocyclic amino group and does not require a protecting group for inclusion with the Thyminyl PNA monomer. Thymine was alkylated with methyl bromoacetate in the presence of potassium carbonate to give the intermediate ester **37**. Hydrolysis of **36** under harsh alkaline conditions gave the target compound upon acidification (Scheme 2.1). The crude product was contaminated with inorganic salts that are removed by aqueous washes on a sinter funnel – this also removed some product from the isolated material.



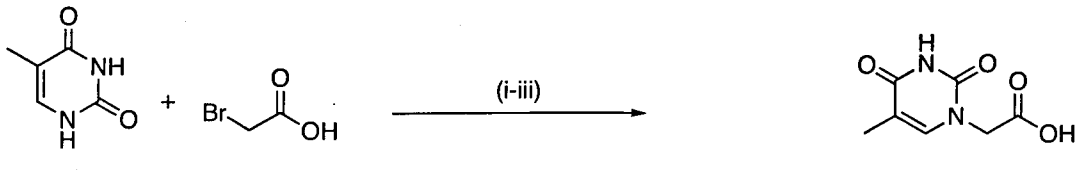
36 **37** **38**

Not isolated **55%** (62% Lit.)

(i) DMF, K_2CO_3, Methyl bromoacetate, RT, 18 h
(ii) 2M NaOH , 100°C, 10 min

Scheme 2.1 Thymine acetic acid synthesis

A more direct procedure by Finn that discarded the saponification step was investigated to improve the overall yield of **38** (Scheme 2.2).[8]



36 **39** **38**

65% (85% Lit.)

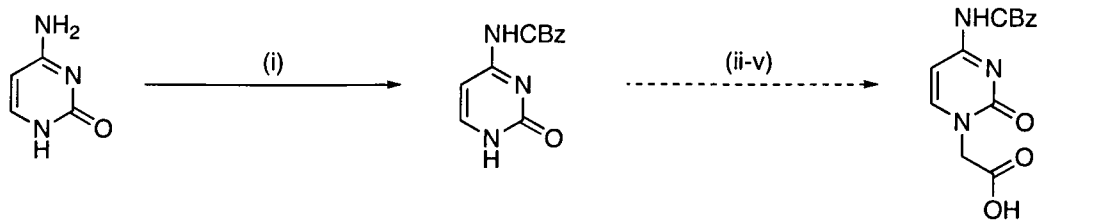
(i) H_2O, **36**, 2M KOH (Aq.), 40°C, 30 min
(ii) Cool to RT, 2 h
(iii) Conc. HCl, filter off **36** at pH 5.5; cool to -4°C add conc HCl until pH 0, 2h stirring

Scheme 2.2 Improved thymine acetic acid synthesis

36 was dissolved in aqueous potassium hydroxide and alkylated with **39**. **38** was isolated upon acidic work up without inorganic salt contaminants and in slightly better overall yield. The procedure however involves a difficult workup to remove unreacted **36** from the reaction and results in the loss of some product. Although this method gives a slightly greater overall yield, the simpler procedure of Dueholm was preferred and used as the method of choice.

2.1.1.2 CBz-protected Cytosyl Acetic Acid

40 was stirred in pyridine at 0°C and treated dropwise with CBz chloride to give the desired carbamate **41**. Alkylation of **41** with methyl bromacetate was then attempted (Scheme 2.3).



40 **41** **42**

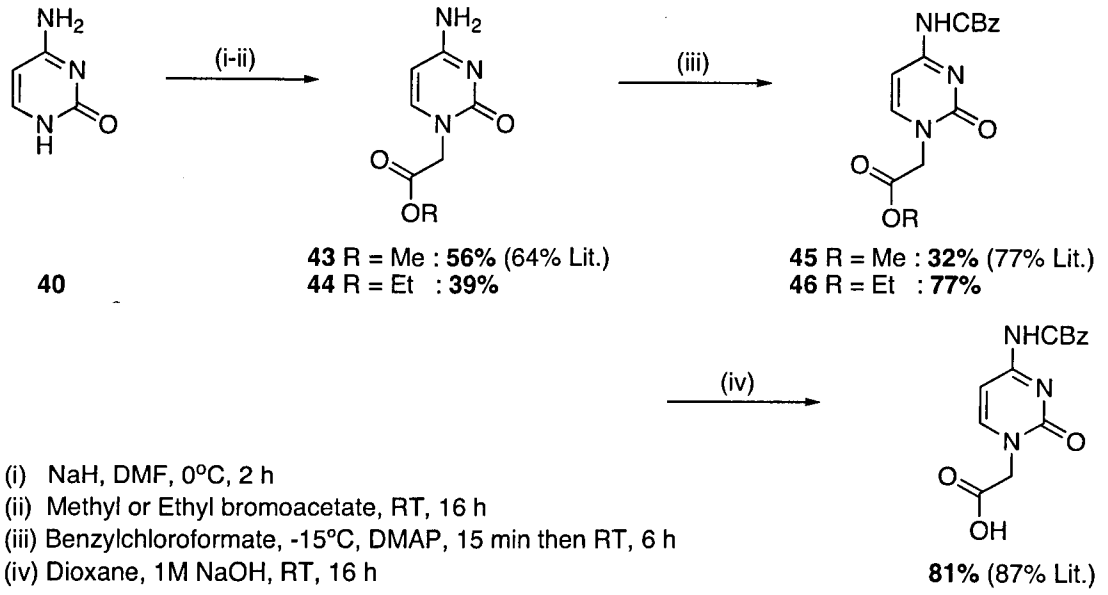
37% (56% Lit.) **Trace** (40% Lit.)

(i) Benzylchloroformate, pyridine, RT, 0°C, 24 h
(ii) Methyl bromoacetate, K_2CO_3, DMF, RT, 18 h
(iii) H_2O/4M HCl (80:3, v/v), 0°C, 15 min
(iv) H_2O/2M NaOH (6:1, v/v), RT, 30 min
(v) 4M HCl

Scheme 2.3 Dueholm CBz-cytosyl acetic acid synthesis

The results of Dueholm could not be reproduced. The reaction gave a complex mixture of products that were inseparable by chromatography.

An alternative synthesis of **42** (Scheme 2.4) by Schwergold *et al.* was then attempted.



40

43 R = Me : **56%** (64% Lit.)
44 R = Et : **39%**

45 R = Me : **32%** (77% Lit.)
46 R = Et : **77%**

(i) NaH, DMF, 0°C, 2 h
(ii) Methyl or Ethyl bromoacetate, RT, 16 h
(iii) Benzylchloroformate, -15°C, DMAP, 15 min then RT, 6 h
(iv) Dioxane, 1M NaOH, RT, 16 h

81% (87% Lit.)

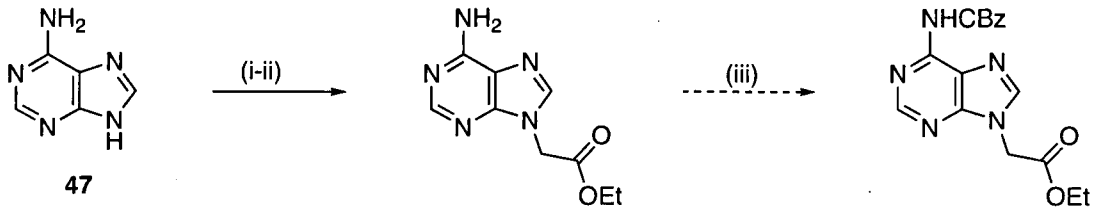
42

Scheme 2.4 Schwergold CBz-cytosyl acetic acid synthesis

Cytosine was deprotonated with sodium hydride to allow the selective alkylation of the *N*-1 nitrogen with bromomethylacetate. The crude product was recrystallised once to purity with aqueous methanol (1:1). DMAP catalysed acylation of **43** gave *N*-7 protected carbamate **45** as the major product. **45** was readily hydrolysed to give cytosyl acetic acid **45**. Acylation of the exocyclic amine did not proceed as efficiently as stated in the literature – it was thought that the low solubility of the intermediate **43** was responsible for the low yield. The synthesis of homologue **44** was attempted with ethyl bromoacetate. It was thought that **44** would be more soluble in DMF and therefore react more readily with the acylating agent. Alkylation of **40** proceeded less efficiently than when using methyl bromoacetate but the acylation of **44** did indeed proceed with much improved yield. Hydrolysis of the esters **45** & **46** was found to be equally efficient and acid **42** was recovered in ~80% yield irrespective of the starting material.

2.1.1.3 CBz-protected Adenyl Acetic Acid

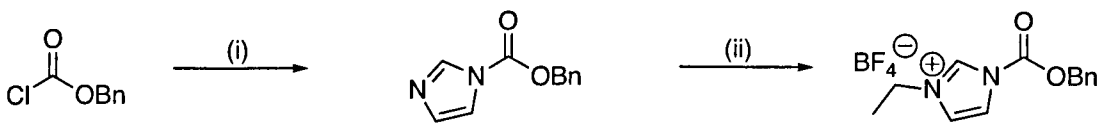
The procedure of Thompson was followed for the synthesis of CBz Adenyl acetic acid (Scheme 2.5).[6]



(i) NaH, DMF, RT, 3.5 h
(ii) Ethyl bromoacetate, <30°C, 19 h
(iii) Benzylchloroformate, DMAP, -15°C, 15 min then RT, 6h

Scheme 2.5 Attempted synthesis of Cbz-protected Adenyl acetic acid

Adenine was deprotonated at *N*-9 with sodium hydride and alkylated with bromoethylacetate to give **48**. The acylation of the exocyclic *N*-6 amino group is difficult; it is only weakly nucleophilic and was found to be unreactive toward benzylchloroformate at RT under DMAP-mediated catalysis. At elevated temperatures, the reaction produced an inseparable mixture of products. The procedure of Dueholm stated that protection of the relatively unreactive amino group had been achieved using Rapoport's reagent **52**. Thompsons route to **49** was exchanged for that of Dueholm and acylation of **48** was attempted with Rapoport's reagent. **52** is not commercially available, and was prepared according to the procedure of Watkins (Scheme 2.6).[9]

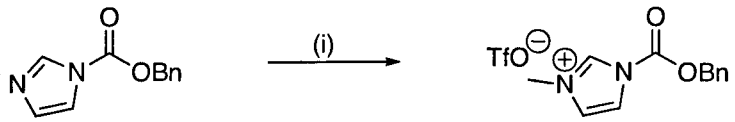


(i) Imidazole (2.0 equiv.), RT, 1 h
(ii) 1M Triethyloxonium tetrafluoroborate in CH_2Cl_2, 0°C, 2 h

Scheme 2.6 Synthesis of Rapoport's reagent

Intermediate **51** was synthesised as a colourless oil that could not be recrystallised from petrol to purity as stated in the literature; it was found to contain trace impurities in the

^{1}H NMR. Reagent **52** was prepared in-situ from crude **51** and combined immediately as a crude solution with a solution of **48** dissolved in DMF at 0°C. The reaction failed to give clean acylation as expected but rather gave an inseparable mixture. The high-reactivity and low stability of reagent **52** is believed to make this reaction capricious and unreliable. An alternative acylation reagent was sought, and the analogous salt **50** utilised by Haaima was prepared (Scheme 2.7) from the previously synthesised **51**.



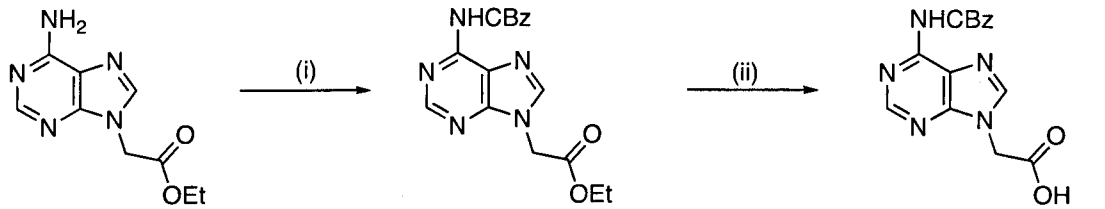
~ **99%** (Lit. Quant.)

51 **53**

(i) Methyl triflate, CH_2Cl_2, 0°C, 5 min

Scheme 2.7 Preparation of alternative acylating reagent

Unlike **52**, triflate salt **53** is an air stable solid that is easily handled and was used successfully to acylate **48** (Scheme 2.8). Hydrolysis of **49** under standard conditions gave the the CBz protected adenyl actic acid **54** in moderate yield.



48

70% (73% Lit.)
49

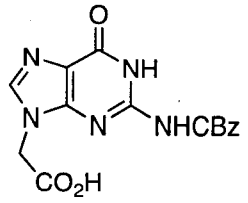
70% (Lit. 104%)
54

(i) **53**, DMF, 0°C,18 h
(ii) Dioxane, 1M NaOH, RT, 16 h

Scheme 2.8 Completed synthesis of Cbz-protected Adenyl acetic acid

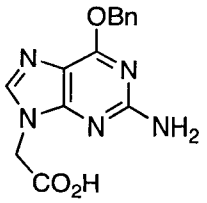
2.1.1.4 Oxybenzyl Guanyl acetic acid

Dueholm and Thompson did not prepare the same Guanyl acetic acid derivitive to synthesise their respective PNA monomers (Figure 2.3).



55

Thomson's guanyl acetic acid

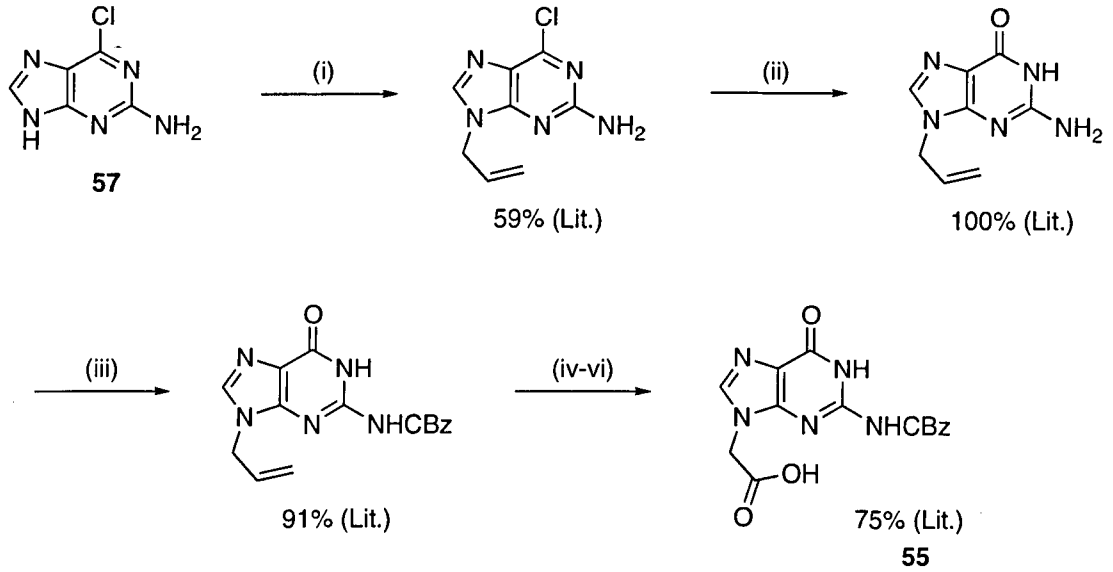


56

Dueholm's guanyl acetic acid

Figure 2.3 Different guanyl acetic derivitives

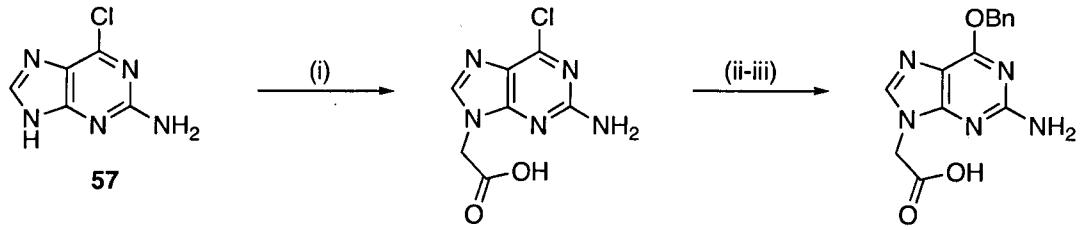
Thompson prepared CBz guanyl acetic acid over 4 steps (Scheme 2.9).



(i) Allylbromide, K_2CO_3, DMF
(ii) 1M HCl, reflux
(iii) CBz-Imidazole, 18-crown-6, KH, THF
(iv) O_3, CH_2Cl_2, MeOH
(v) Me_2S
(vi) $NaClO_2$, NaH_2PO_4, H_2O, THF, $(CH_3)_3OH$, 2-methyl-2-butene

Scheme 2.9 Thompson synthesis of guanyl acetic acid derivitive

56 was prepared by Dueholm in only two steps, albeit in less overall yield than **55** (Scheme 2.10).



57

49% (Lit.)

58

52% (Lit.)

56

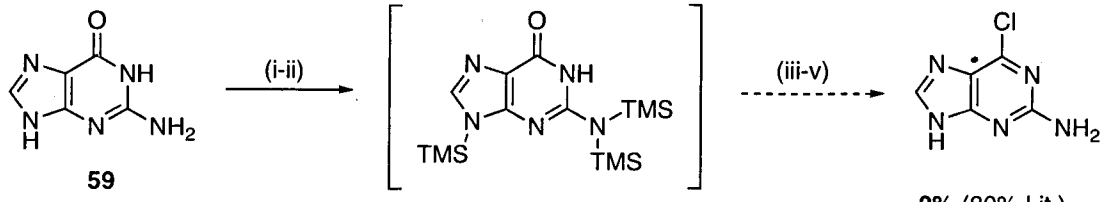
(i) Bromoacetic acid, K_2CO_3, DMF, RT, 2 h
(ii) Sodium, benzyl alcohol, 0°C, 1 h
(iii) Recrystallise from ethanol/water (3:1) at pH 2

Scheme 2.10 Dueholm synthesis of guanyl acetic acid derivitive

Dueholm stated that the exocyclic amine of guanine is unreactive toward solid phase chemistry reagents and does not need protecting. Thompson states that it is difficult to protect the amine but it is necessary to do so. Dueholm claimed that the oxybenzyl group, not present in **55**, is crucial to the solubility of the final monomer. It was decided that the synthesis of Dueholm would be attempted first as it was the shortest synthesis. If **56** or its derived G-PNA monomer were found to be unsuitable for PNA synthesis, then the approach of Thompson would be used instead.

Guanine cannot be selectively alkylated at the *N*-7 position; attempting to do so gives inseparable mixture of *N*-7 and *N*-9 isomers.[10] For this reason, the synthesis of any G-PNA monomer described in the literature begins with 2-amino 6-chloropurine **57**, a masked guanine equivalent that can only be alkylated at the desired N-7 position.
2-Amino 6-Chloropurine is commercially available but was not readily available in large quantities; **57** had to be synthesised in house. The synthesis of **57** is known (exclusively) in patent literature. Of the methods published, chlorination of guanine is the only suitable method to access multigram quantities of **57**.[11-14]

Initially the method of Reiner was attempted (Scheme 2.11) as this quoted the greatest yield.[14]



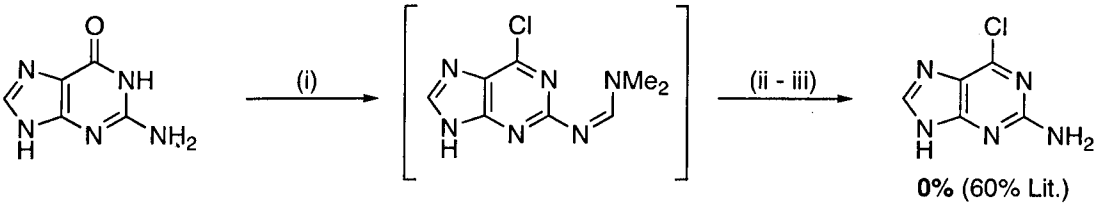
59 **60 - Not isolated** **0%** (80% Lit.) **57**

(i) DMF or DMA, HMDS, NH_4SO_4, reflux, 4 h
(ii) Distill off solvent and excess HMDS
(iii) $POCl_3$, TEBAC or TEBAB, MeCN, reflux 1 h
(iv) Distill off excess $POCl_3$, MeCN
(v) Neutralise with 2M NaOH then reflux until remaining organics distlled off

Scheme 2.11 Reiner synthesis of 2-amino 6-chloropurine

This reaction was attempted first using the phase transfer catalyst TEBAC in both DMA and DMF. The method returned an insoluble brown fibrous precipitate upon the hydrolysis of the crude reaction oil at neutral pH. Repeating the experiments using the alternative catalyst cited by Reiner (TEBAB) in both DMF and DMA gave the same result. The precipitate could not be characterised.

The less technically challenging but lower yielding procedures of Taketo were then investigated (Scheme 2.12).[13]



59 **61** **0%** (60% Lit.) **57**

(i) $POCl_3$, DMF, 100°C, 5 h then neutralisation with $NaHCO_3$
(ii) 35% HCl, RT, 20 h
(iii) 25% ammonia (aq), 60°C

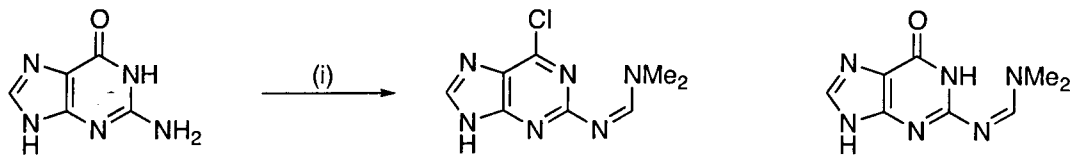
Scheme 2.12. Taketo synthesis of 2-amino 6-chloropurine

The reaction ultimately yields a dark brown solid that, despite attempts to purify, could not be characterised.

The reaction was repeated without external heating (however the reaction is exothermic and the temperature was elevated to 30°C). A pale yellow solid was recovered that was identified as 2-formaminidineamino purine **62** (Scheme 2.13).

This result demonstrated that the exocyclic amino group is readily reactive with the Vilsmeier reagent and rapidly generated an exocyclic formamidine.

However, the guanyl oxygen did not react with the chlorinating agent until the reaction was heated to 95°C. Desired intermediate **61** was isolated as pale brown crystals *strictly* when the reaction was run above 95°C yet below 100°C; above 100°C the reaction yielded an unidentifiable and insoluble brown precipitate. It was also found that the use of fresh dry $POCl_3$ is critical to the success of the reaction.



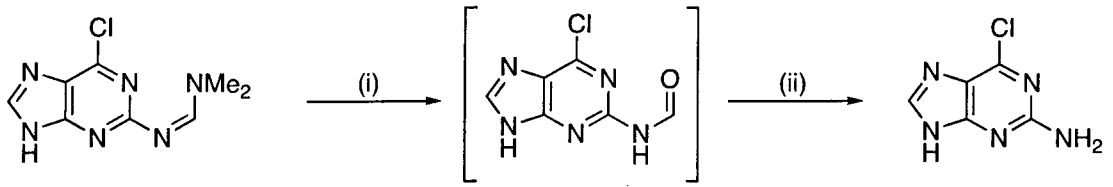
59 **61** **62**

(i) $POCl_3$, DMF, 95°C, 5 h then neutralisation with solid K_2CO_3 at 0°C

Scheme 2.13 Synthesis of key intermediate **61**

Intermediate **61** was treated with harsh hydrolysis conditions (as used earlier, Scheme 2.12) and was destroyed. Slightly milder hydrolysis conditions (2M NaOH, 40°C) fully hydrolysed **61** into **59**.

Milder conditions for the hydrolysis of **61** were given in Taketo's patent (Scheme 2.14) and were used successfully to convert **61** into **63** and then **57**.



61 **63** **57**

(i) 12% acetic acid, 70°C, 3 h

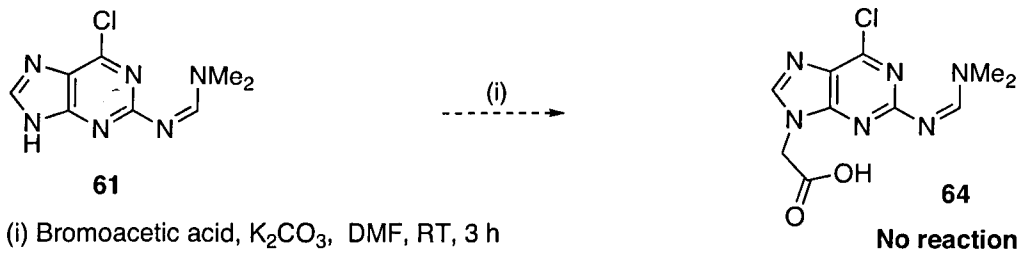
(ii) 10% sodium hydroxide, RT, 2 h then conc. HCl

Scheme 2.14 Mild hydrolysis of 2-formaminidineamino purine

Thorough washing and drying of the intermediate **63** was found to be critical to the success of this reaction; acetic acid interfered with hydrolysis of the formyl group. **59** and **57** are very difficult to difficult to dry (and therefore handle) on a large scale due to

agglutination problems. It was found that briefly heating **57** in water at 85°C and then rapidly chilling to RT solved the agglutination problem.

With precious **57** in hand, a guanyl acetic acid derivative for the construction of a G-PNA monomer could be synthesised. Initially the method of Dueholm (Scheme 2.11) was followed to gain access to target **58**, however this work could not be reproduced. Alkylation of **64** was attempted to determine if the acetic acid could be introduced instead during the synthesis of **57** (to give **64** instead), but **64** could not be isolated (Scheme 2.15).



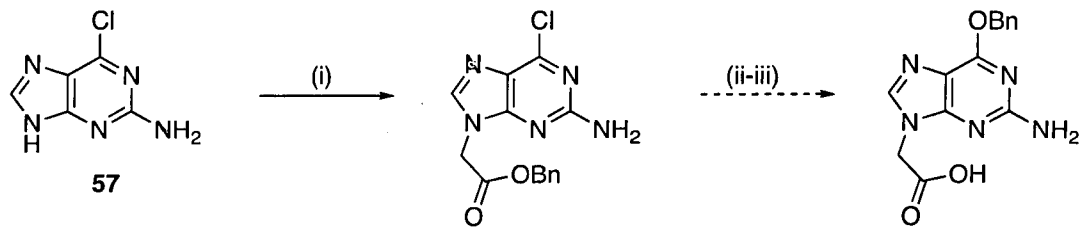
61 **64**

(i) Bromoacetic acid, K_2CO_3, DMF, RT, 3 h

No reaction

Scheme 2.15 Attempt to introduce acetic acid during the synthesis of **54**

Kofoed and co workers reported an improved synthesis of **57** in response to difficulties they encountered with the procedure of Dueholm (Scheme 2.16).[4]



57

47% (76% Lit.)
65

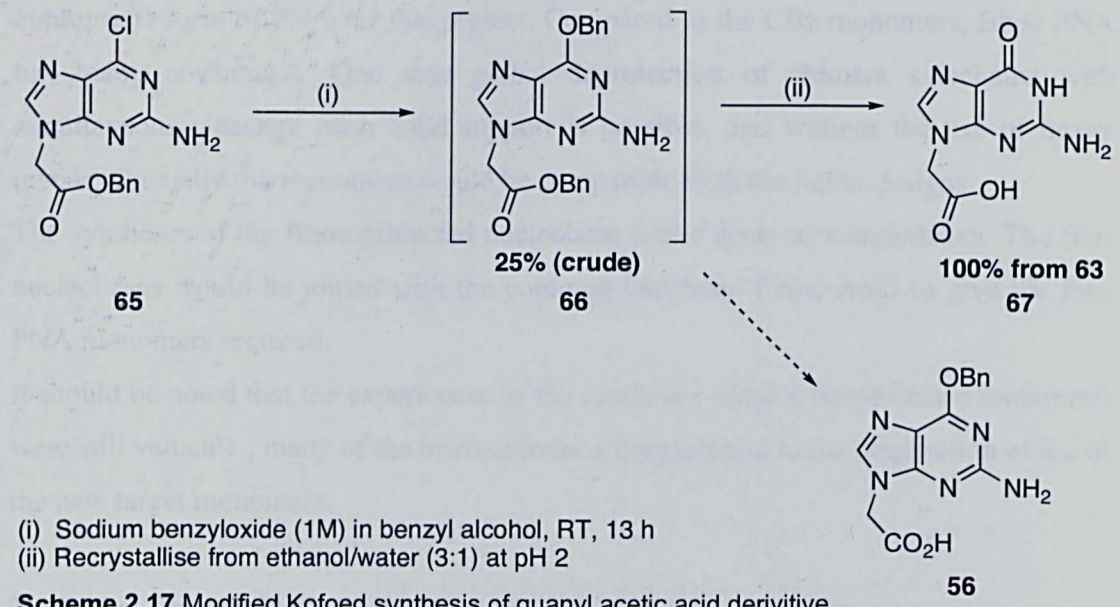
N/A (83% Lit.)
66

(i) Benzyl bromoacetate, K_2CO_3, DMF, RT, 2 h
(ii) Sodium, benzyl alcohol, 0°C, 1 h
(iii) Recrystallise from ethanol/water (3:1) at pH 2

Scheme 2.16 Kofoed synthesis of guanyl acetic acid derivitive

65 was obtained in lower yield than that reported. It was then considered if the lower yield was due to **57** being of poor quality. To verify the purity of **57**, 1.0 g of commercially sourced 2-amino 6-chloropurine (Aldrich) was converted into **65**. The yield obtained was identical, affirming that **57** was not the cause of the poor yields encountered.

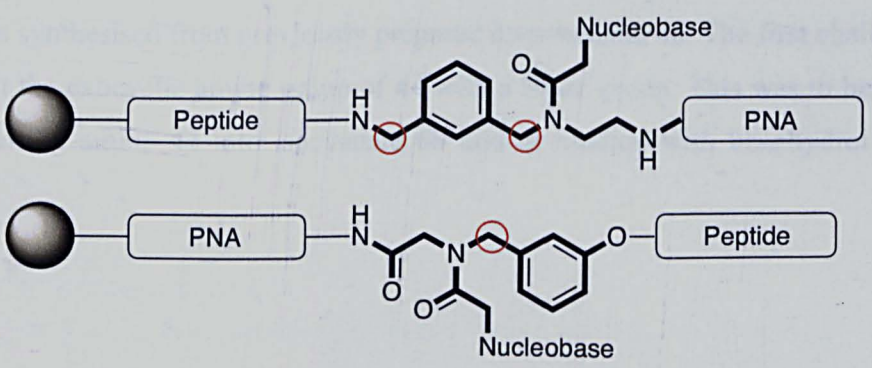
The conversion of **65** into **56** was then attempted (Scheme 2.17). The reaction yielded a small quantity of crude material that contained intermediate **66** (as evidenced by crude NMR and MS analysis). However, after acidic recrystallisation, **66** was fully hydrolysed into **67** instead of giving **56**.



(i) Sodium benzyloxide (1M) in benzyl alcohol, RT, 13 h
(ii) Recrystallise from ethanol/water (3:1) at pH 2

Scheme 2.17 Modified Kofoed synthesis of guanyl acetic acid derivitive

Before the synthesis of **56** could be completed, the linker unit designs were generated and they were found to be incompatible with the CBz protected PNA monomer targets. Both of the linker unit designs were based on an aryl core with benzylic linkages. It was thought that the hydrogenlytic cleavage conditions required for removing the PNA Cbz protecting groups would almost certainly cleave the chimeras apart at these sensitive linker junctions (Figure 2.4).



Linkers shown in bold - bonds sensitive to hydrogenolysis circled in red

Figure 2.4 Linker designs - incompatible with Cbz protected PNA

A new set of PNA monomers that were compatible with the linkers were needed therefore to replace the redundant CBz protected monomers.

A previously undiscovered patent by Coull detailing the synthesis of the commercially available PNA monomers **32-35** was found.[7] Bhoc/Fmoc protected PNA is the most appropriate form of PNA for this project. Compared to the CBz monomers, Bhoc PNA has many advantages. One step global deprotection of chimera sidechains with simultaneous cleavage from solid support is possible, and without the use of heavy metals. Crucially the monomers would be compatible with the linker designs.
The syntheses of the Bhoc protected nucleobase acetic acids was undertaken. The four nucleobases would be joined with the common backbone Fmoc AEG to give the four PNA monomers required.
It should be noted that the experiences of the syntheses of the CBz-protected monomers were still valuable.; many of the intermediates were common to the preparation of the of the new target monomers.

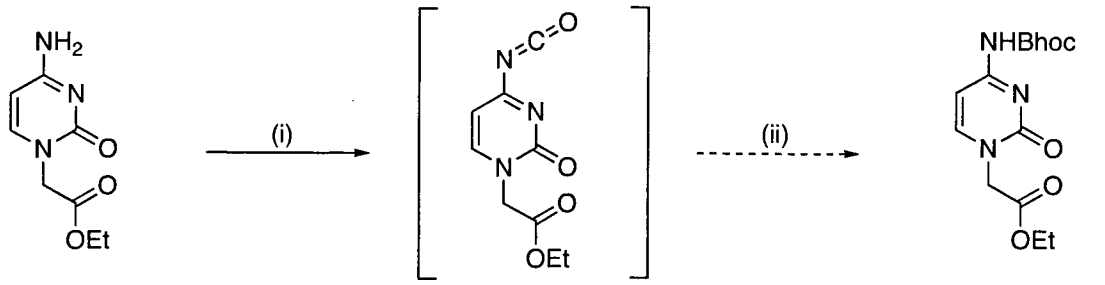
2.1.2 Bhoc-Protected Nucleobase Acetic Acid Synthesis

2.1.2.1 Thyminyl acetic acid

The T-PNA monomer described by Coull does not require protection and the thyminyl acetic acid **37** synthesised earlier using the method of Dueholm is suitable for its construction.

2.1.2.2 Bhoc Cytosyl Acetic Acid

69 can be synthesised from previously prepared intermediate **44**. The first challenge was to protect the exocyclic amine group of **44** with a Bhoc group. This was to be achieved by converting amine **44** into isocyanate **68** and quenching with benzhydrol (Scheme 2.18).



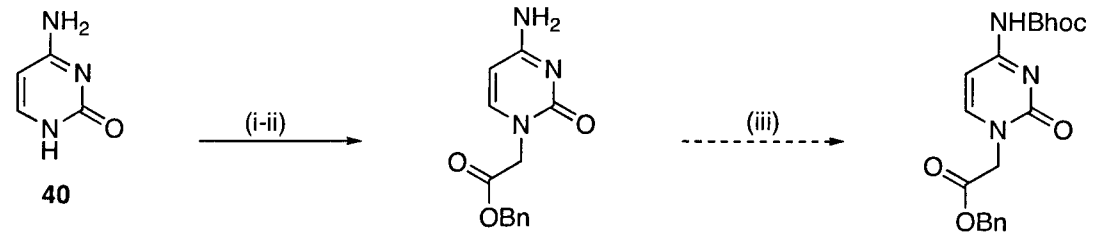
44 **Not isolated** **68** **0%** (77% Lit.) **69**

(i) Carbonyl Diimidazole, DMF, RT, 1.5 h
(ii) Benzyhydrol, 60°C, 6 h

Scheme 2.18 Attempted Bhoc protection of **44**

The formation of intermediate **68** was followed by tlc analysis of methanol quenched reaction aliquots and this appeared to be complete after 1h. Recrystallisation of the crude orange oil from ethanol returned only starting material **44**.

A modified procedure by Winssinger was attempted wherein the benzyl cytosylacetate **70** was used instead of **44** (Scheme 2.19).[15] It was thought that the benzyl ester would be much more soluble that **44,** and thus its derived isocyanate would be present in solution in much greater quantity than **68**. It was hoped that increased solubility would allow for benzhydrol to react with the isocyanate.and give the desired product. Ester **70** was observed to have fully reacted with the carbonyldiimidazole after 2h. The isocyanate was then reacted with benzhydrol for 6h at 60°C. Quenching the reaction with methanol and recrystallising from ethanol only yielded starting material **70**.

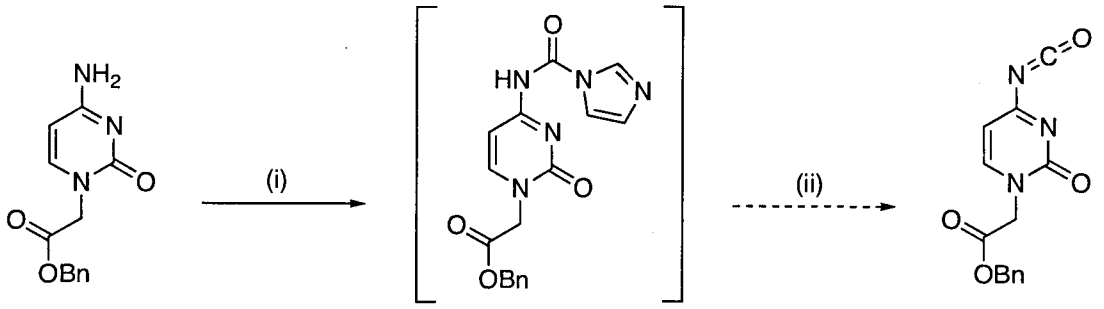


40 **80%** (91% Lit.) **70** **0%** (76% Lit.) **71**

(i) ^{t}BuOK, DMF, 100°C, 2 h
(ii) Benzyl bromoacetate, 10°C, 3 h
(iii) Coull conditions (Scheme 2.18)

Scheme 2.19 Winnsinger modification for Bhoc cytosine PNA monomer synthesis

Recovery of the starting material led us to believe that the isocyanate was not being formed.



70 72 73

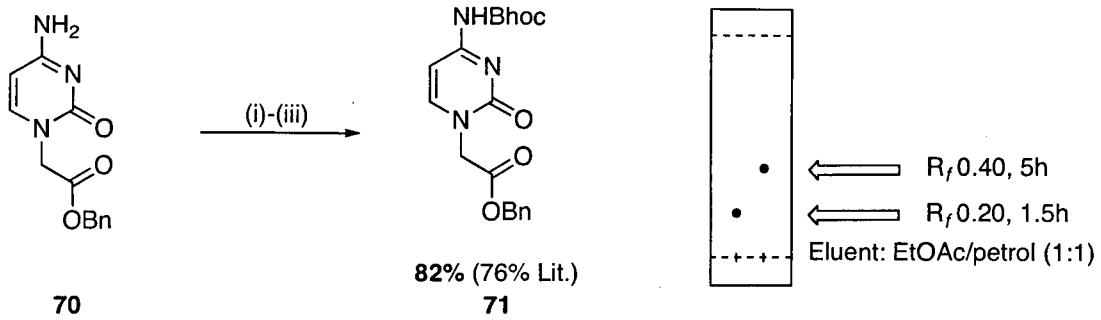
(i) Carbonyl Diimidazole, DMF, RT, 1.5 h
(ii) Longer reaction time and/or heat ?

Figure 2.20 Isocyanate formation

It was speculated that after the 1.5 h reaction of **70** with carbonyl diimidazole at RT that the starting material had been fully converted into intermediate **72**, and not **73** as required (Scheme 2.20). To explain the recovery of starting material, **72** was assumed to be (effectively) unreactive at RT or 60°C and was hydrolysed back to **70** during aqueous recrystallisation at the end of the reaction. It was therefore required to determine a method of converting **72** into **73**.

It was suspected that a longer initial reaction time would allow **73** to form in useful quantities and that **73** would react with benzhydrol as desired. The reaction was repeated with an initial aim of allowing 24 h to preform **73**.

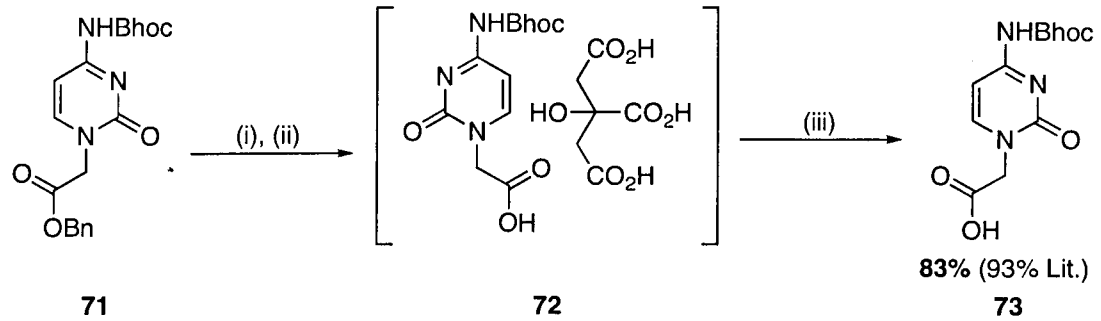
However after 5 h stirring at RT, the reaction rapidly solidified to become an immobile colourless mass (Scheme 2.21). The R_f (ethyl acetate) of the original (single component) reaction mixture after 1.5 hr was 0.20. After 5h that spot had disappeared and one new component was then observed at an R_f of 0.40. Under the assumption that **73** had now formed, benzhydrol was added and the temperature was raised to 60°C. The colourless solid melted between 40 and 55°C and after 6h of continuous heating,**71** was obtained in good yield.



(i) Carbonyl Diimidazole, DMF, RT, 5h
(ii) Benzyhydrol, 60°C, 6h
(iii) Methanol quench, recrystallization ($MeOH/H_2O$, 3:1)

Scheme 2.21 Bhoc protection of benzyl cytosylacetate X

The procedure of Winssinger was followed further to carefully hydrolyse **71** and complete the synthesis of **73**, although this was not without complications (Scheme 2.21).[15] The rapid hydrolysis was monitored regularly for the total consumption of **71** and quenched immediately with 20% citric acid solution (to prevent the hydrolysis of the Bhoc group). Bhoc cytosyl acid **73** was recovered initially as the citrate adduct **72**. The citric acid would interfere with the final step of the monomer synthesis and had to be removed.



(i) 0°C, LiOH solution, pH 14, ~ 6 min
(ii) 20% citric acid solution
(iii) Saturated brine 24h

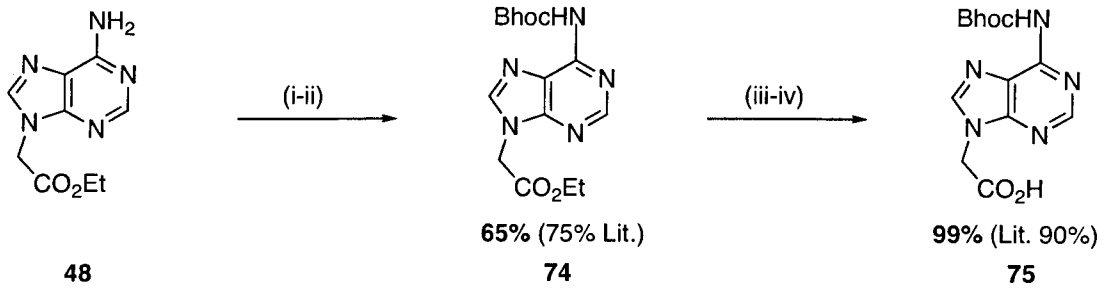
Scheme 2.22 Benzyl ester hydrolysis

Unfortunately the hydrolysis reaction cannot be quenched by any other acid than citric acid. It was found that other acids (acetic acid, hydrochloric acid) in varying concentrations (0.01M-1M) would partially or fully remove the Bhoc group below pH 5 and **73** would not precipiate above pH 3. Citric acid has a high affinity for **73** and could not be washed out with water or weakly basic solutions (Na_2CO_3, NH_4Cl). Strong base solutions (Eg KOH, NaOH, 25% NH_3(aq)) dissolve the salt and destroy the Bhoc group.

Hydrophobic **73** was slurried with a minimal amount of saturated brine and stirred at RT for 24 h. After 6 h most of the citric acid had been sequestered by the sodium chloride solution (as evidenced by NMR). After 24 h **73** was recovered by filtration and no traces of citric acid were detectable by NMR.

2.1.2.3 Bhoc Adenyl acetic acid

Using intermediate **48** (synthesised earlier), the target **75** was made according to the method of Coull, albeit with some modifications (Scheme 2.23).[7] Intermediate **74** did not crystallise upon quenching of the reaction with water as stated in the literature. The crude sticky yellow oily gum that formed instead was intractable and found to be soluble only in DMF and DMSO. This material could not be purified by means of recrystallisation or chromatography. With some effort, it was found that quenching the reaction with a solution of water/acetone (1:1) effected the crystallisation of the crude compound.



48 **74** **75**

65% (75% Lit.) **99%** (Lit. 90%)

(i) Carbonyl diimidazole, DMF, 105°C, 2 h
(ii) 105°C -> 95°C then benzhydrol, cool to RT over 16 h; H_2O/acetone (1:1) quench
(iii) LiOH, 0°C, 6min then 20% citric acid, 2 h
(iv) Saturated brine, 18 h

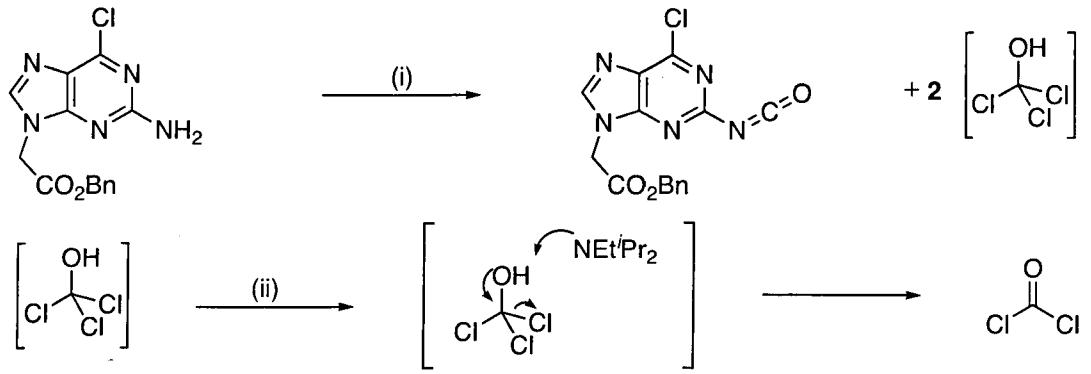
Figure 2.23. Coull synthesis of Bhoc adenyl acetic acid

The impurities formed an oil in this solvent mixture and upon standing were partitioned as the less dense layer and decanted from the solid product. The crude amorphous product was recrystallised to purity from ethanol to give **75**.

2.1.2.4 Bhoc Guanyl acetic acid

The procedures of Coull and Winssinger both use triphosgene as the reagent of choice for the formation of guanyl isocyanate formation.[7, 15] Triphosgene is a crystalline

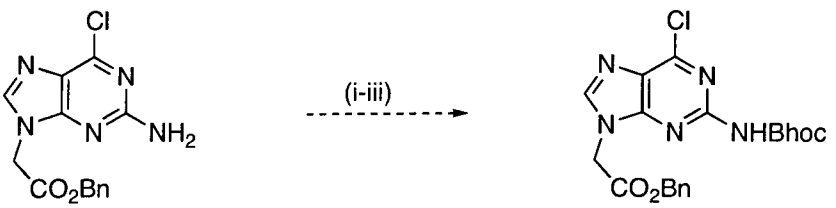
reagent that is, as the name suggests, a phosgene equivalent. Under standard reaction conditions, triphosgene will react with an amine to form an isocyanate and in doing so generates two equivalents of phosgene *in-situ* (Scheme 2.24).



(i) $CO_3(CCl_3)_2$, dimethylformamide, 0°C, 30 min
(ii) Diisopropylethylamine

Scheme 2.24 Isocyanate formation with triphosgene

Although triphosgene is easier and safer to handle than phosgene solution, it still requires very careful use and represents a significant potential hazard for use in large scale chemistry. The literature does not mention the use of any other reagents or methods for converting **65** to **76**. It was thought worthwhile to attempt the transformation using the conditions for making **76**; these would be safer and scalable conditions (Scheme 2.25).



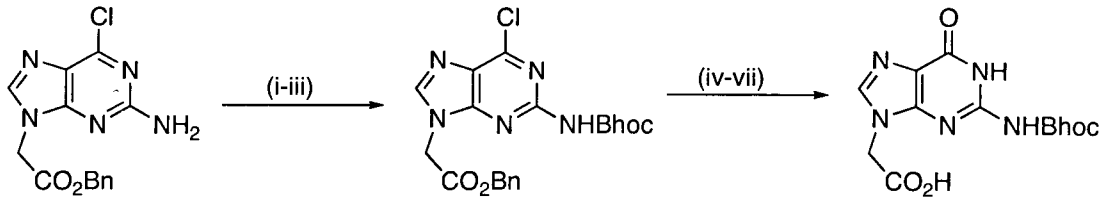
65 **76**

(i) Carbonyl diimidazole, DMF, 105°C, 2 h
(ii) 105°C -> 95°C then benzhydrol, cool to RT over 16 h
(iii) Ethanol quench

Scheme 2.25. Investigating the use of carbonyl diimidazole for isocyanate formation

The reaction was not successful, demonstating that the relatively unreactive *C*-2 amine group of guanine required the more reactive triphosgene to undergo isocyanate formation. The method of Winssinger was used, with some modification, to obtain **76**

(Scheme 2.26).[15] **76** was purified by recrystallisation from methanol in the literature procedure, but this was found to be difficult and unreliable. Changing the recrystallisation solvent to acetonitrile was found to improve the reliability of the purification, albeit in lower yield (35%) than the cited literature recovery of **76** (63%). The mother liquors can be concentrated and recrystallised for further recovery of up to 15%. It was found that direct purification of the crude reaction mixture by chromatography reliably gives an acceptable yield of 55% of **76**; this was deemed the most suitable method for purification of **76**.



65 **76** **55%** (63% Lit.) **77** **82%** (Lit. 86%)

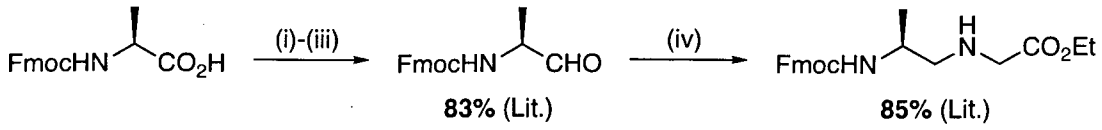
(i) Triphosgene, THF, 0°C, 1 h
(ii) Diisopropylethylamine, 0°C, 30 min
(iii) Benzhydrol then warm to RT over 16 h
(iv) NaH, 3-hydroxypropionitrile,THF, -78°C then add **76**, warm to 0°C and stir 2.5 h
(v) Benzhydrol then allow to warm to RT over 16 h
(vi) Quench with 10% citric acid (aq.)
(vii) Saturated brine solution, 24 h

Scheme 2.26 Modified Winssinger synthesis of Bhoc guanyl acetic acid

76 was transformed into the target guanyl acetic acid **77** by reaction with 3-sodiumhydroxypropionitrile at low temperature as according to the procedure of Winssinger.[15]

2.1.3 Fmoc protected amino ethyl glycine

The common backbone upon which the nucleobase acetic acids are grafted onto is **82**. A number of different approaches to **82** have been reported in the literature. Initially a high yielding procedure by Xu *et al.* (Scheme 2.27) for the synthesis of PNA monomers with chiral backbones was thought to be suitable for modification for the synthesis of **82**.

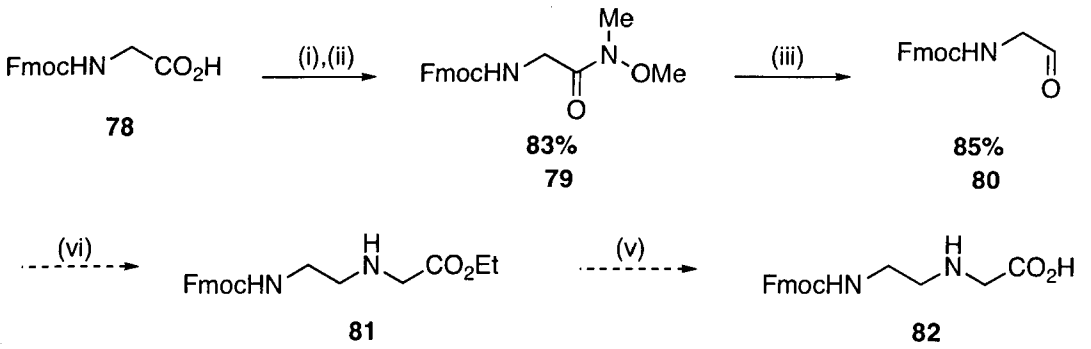


83% (Lit.) **85%** (Lit.)

(i) Isobutylchloroformate, NMM, THF, -20°C, 15min
(ii) NEt_3, HCl.NHMe(OMe), DMF, -20°C, 30min
(iii) Lithium Aluminium hydride, THF, 0°C, 30min
(iv) Glycine ethyl ester hydrochloride, 10% Pd/C, hydrogen, MeOH, RT, 24h

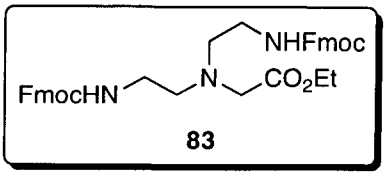
Scheme 2.27 Wu & Xu synthesis of chiral PNA backbone

In place of alanine, Fmoc glycine would be used to allow rapid access to **81**. Wu cites conditions for hydrolysis of the ethyl ester with minimal loss of the Fmoc group (Scheme 2.28).



78 **83%** **79** **85%** **80** **81** **82**

(i) Isobutylchloroformate, NMM, THF, -20°C, 15min
(ii) NEt_3, HCl.NHMe(OMe), DMF, -20°C, 30min
(iii) Lithium Aluminium hydride, THF, 0°C
(iv) Glycine ethyl ester hydrochloride, 10% Pd/C, MeOH, RT hydrogen, 0°C, 24h
(v) 2.5M NaOH, H_2O/THF, 0°C, 5min



83

Scheme 2.28 Modified Wu & Xu synthesis of PNA backbone

The method of Wu was used to convert **78** into **80.** Reductive amination of aldehyde **80** with glycine ethyl ester did not give the desired product **81** using the literature conditions. The substance isolated was found to be the doubly-reductively aminated product **83**. The secondary amine product **81** is much more reactive than the starting material (glycine ethyl ester) and thus cannot be isolated in the presence of the very reactive aldehyde **80**. It is most likely that without substitution at the alpha carbon, **80** cannot be prevented from further reacting with **81**(Figure. 2.5).

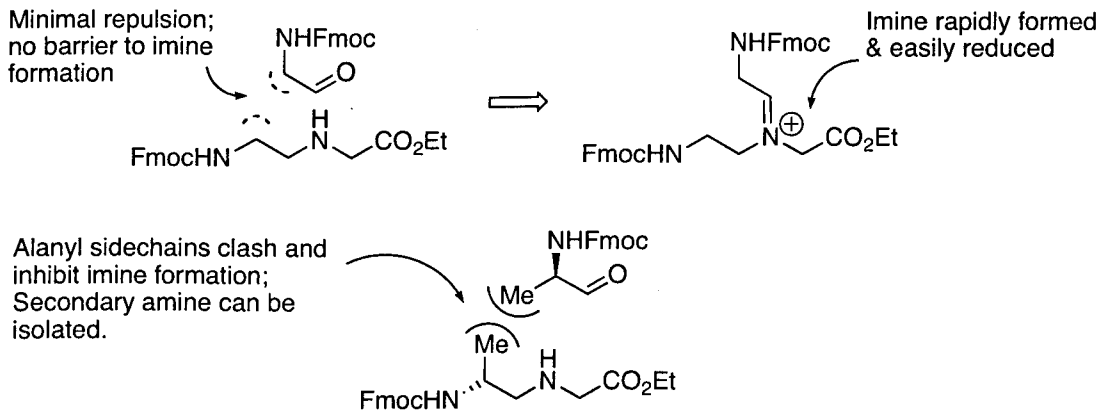
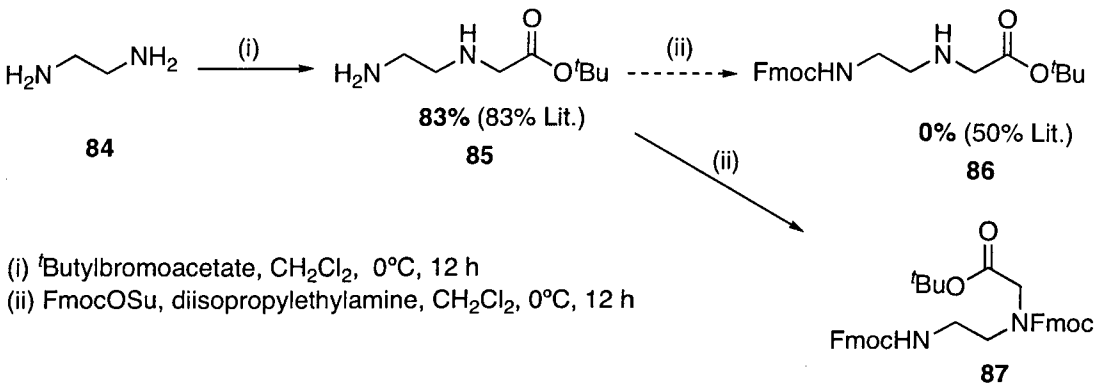


Figure 2.5 Differences in reactivity between alanine and glycine derived aldehydes

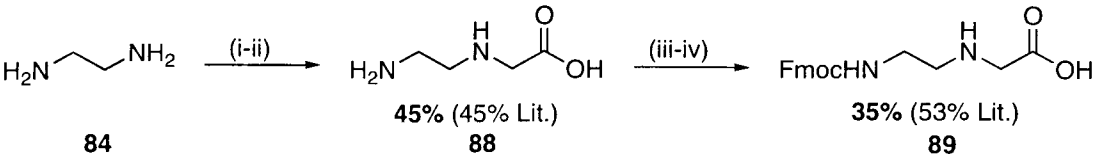
Thompson detailed a procedure for the synthesis of the 'butyl-ester **86** in his synthesis of Cbz protected PNA monomers. Protolysis of **86** with a strong acid (such as TFA) would give the desired backbone **82** (Scheme 2.29).[6]



Scheme 2.29 Thompson synthesis of Fmoc AEG; PNA backbone

Low temperature addition of 'butylbromoacetate to a concentrated solution of **84** gave the intermediate **85** as reported. Selective protection of the primary amine of **84** (ahead of the neighbouring secondary amine) was attempted using the method described in the literature. That a primary amine was selectively protected over a (typically) more nucleophilic secondary amine was surprising. It was known from the synthesis of **83** that the secondary amine of **85** would be very reactive. It was thought that the selectivity reported by Thompson must be due to a sensitivity of the protecting reagent toward steric bulk; that the 'butyl ester sterically shields the secondary amine from the FmocOSu. However, the mono-protection reported by Thompson could not be repeated using the conditions stated. Doubly protected ester **87** was the only product recovered. The procedure of Coull was then attempted. This was designed to overcome the

problem encountered with the previous synthesis; Coull masked the secondary amine of **88** with a trimethylsilyl group before reacting with FmocOSu (Scheme 2.30).



84 **88** **89**

45% (45% Lit.) **35%** (53% Lit.)

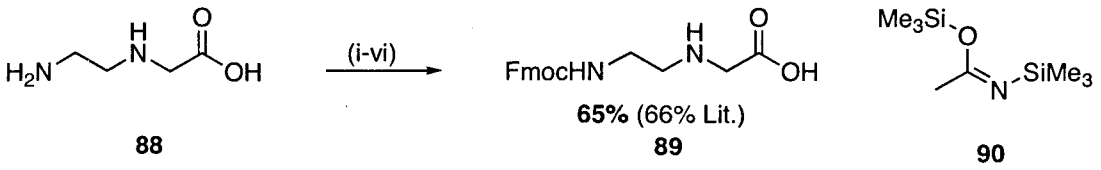
(i) Chloroacetic acid, 5°C, 18 h
(ii) Evaporate off excess ethylene diamine, trituration with DMSO
(iii) Trimethylsilylchloride, DMF, 40 min
(iv) 0°C, FmocOSu then NMM, 2 h

Scheme 2.30 Coull synthesis of Fmoc AEG; PNA backbone

Alkylation of **84** proceeded initially in higher yield than reported – 65% of colourless crystals were recovered that were characterised to be pure by NMR, MS, LC-MS. However it was found that if this material is not recrystallised (ethanol/water, 2:1) to (at the expense of the higher than lit. yield), it would not react with silylating reagents to give **89**.

Treating **88** with trimethylsilylchloride selectively silated the secondary amine and carboxylic acid ahead of the primary amine. Treatment of the masked intermediate with FmocOSu resulted in only the free primary amine being protected. Methanolysis of the TMS groups afforded a colourless amorphous solid that was very insoluble in organic solvents and was difficult to characterise. The mass spectrum of the material was observed to correspond to the mass of **89**, however the NMR was very noisy and difficult to assign. Furthermore it was found that this material reacted very sluggishly with the nucleobase acetic acids in the final coupling experiments. Poor solubility of this material meant it was not suitable for the synthesis of the PNA monomers. It was not possible to find the direct cause for the insolubility, although it was suspected that byproducts from the silating agent were contaminating the sample and retarding its solubility.

An analogous reaction by Breipohl that used bistrimetylsilylacetamide **90** in place of TMSCl as the silylating agent was attempted.[16] Repeating the method of Breipohl gave **89** as a colourless amorphous solid (Scheme 2.31). Characterisation of **89** (made using the method of Breipohl) was straightforward (the material was readily soluble in organic solvents) and proved the material to be pure **89**.



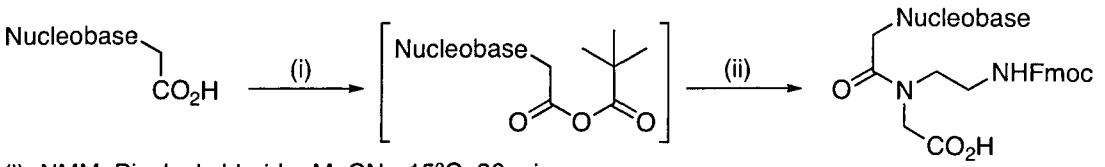
88 89 90

(i) *N,O*-bistrimethylsilylacetamide **90**, DMF, RT, 40 min
(ii) FmocOSu, DMF, RT, 3 h
(iii) MeOH, H_2O, 5°C, 10 min
(iv) 6M HCl, 0°C, 10 min
(v) Evaporate to dryness then 0°C, triturate with CH_2Cl_2
(vi) MeOH, sodium hydroxide, H_2O, 0°C

Scheme 2.31 Breipohl synthesis of Fmoc AEG; PNA backbone

2.2 Nucleobase coupling onto the backbone – PNA monomer synthesis

The final synthetic step for making each PNA monomer is the same. The nucleobase acetic acid is activated as a mixed anhydride and coupled to backbone **89** (Figure 2.6).



(i) NMM, Pivaloyl chloride, MeCN, -15°C, 30 min
(ii) Triethylamine, backbone, MeCN/H_2O, RT, 1-18 h

Figure 2.6 General method for coupling nucleobase acetic acids to Fmoc AEG

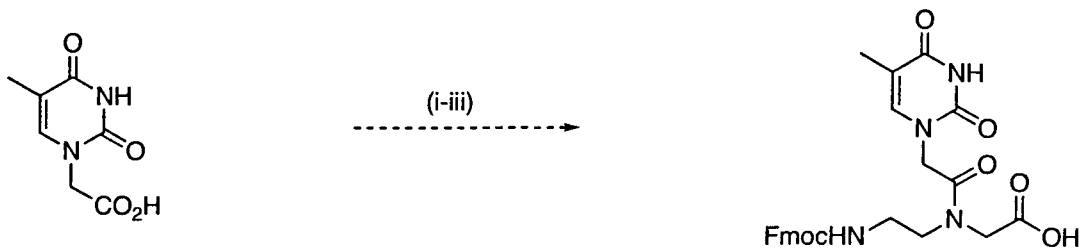
The coupling reaction is not 100% efficient and the crude reaction mixture contained small amounts of starting materials in addition to amine salts and other byproducts. A PNA monomer and its two primary starting materials are carboxylic acids – purification of the monomers from byproducts and starting materials is thus not trivial. PNA monomers are very hydrophobic and sensitive to both basic and acidic pH conditions (by virtue of their protecting groups) and are only appreciably soluble in DMF and DMSO at RT; purification at neutral pH, given the poor solubilities, is not trivial. Each monomer required its own optimised set of conditions for purification.

2.2.1 Thyminyl PNA monomer

The four nucleobase acetic acids are coupled with the backbone using essentially the same reaction conditions (Figure 2.6). The three Bhoc protected nucleobase derivatives (**73,75,77**) required significantly more labour to prepare than thyminyl acetic acid **38**. It

was decided to use the coupling of thyminyl acetic acid to **89** as a model system to establish good working conditions for the general coupling of nucleobase acetic acids onto **89**. The reliable and reproducible synthesis of T-PNA would, in theory, minimise any potential loss of precious **73, 75** and **77** when the synthesis of the A,C and G PNA monomers was attempted.

The procedure of Coull was followed (Scheme 2.32).[7]



38

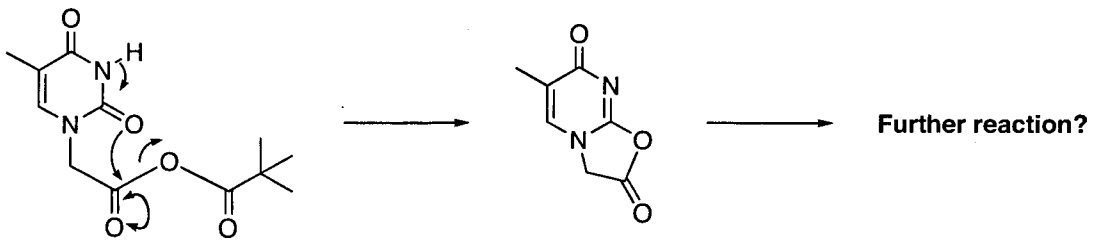
0% (65% Lit.)
35

(i) Pivaloyl Chloride, NMM, MeCN, 0°C, 25 min
(ii) **89**, NEt_3, H_2O/MeCN (3:4), RT, 30 min
(iii) Adjust to pH 3 with 3M HCl at 0°C then stir at RT, 18 h

Scheme 2.32 Attempted T-PNA synthesis

The reaction did not give a colourless precipitate as stated. Instead a brown intractable sticky gum formed. Analysis of the crude reaction mixture indicated that a small amount of product had formed but it was not possible to recrystallise the gum from any common solvent. The reaction was thought to be failing for one or more reasons; Activation of the thyminyl acetic acid, to form the mixed anhydride, was not going to completion; unreacted pivaloyl chloride was affecting the reaction in some adverse way; the activated acetic acid was degrading into one or more byproduct(s) before **89** could react with it; backbone **89** free nitrogen could not be acylated with the anhydride.

Activation of the thyminyl acetic acid was achieved by dropwise addition of pivaloyl chloride at 0°C. The reaction turned gradually yellow and then brown over the 25 minute activation period. It was known that the T-2 oxygen can react with the activated acid **91** to give an unstable intermediate (Figure 2.7).



91

Figure 2.7 Possible degradation pathway of activated thyminyl acetic acid

It was postulated that the yellow colour of the reaction was caused by the presence of the anhydride in the mixture and that a side reaction of the anhydride generated a brown byproduct. The time required for the activation of **38** was examined to establish an optimal time for anhydride formation (Table 2.1). For all of the experiments **35** failed to precipitate during the reaction workup (as expected from the method given by Coull). Evaporation of each crude reaction gave a crude residue that was inspected for **35** by mass spectrometry.

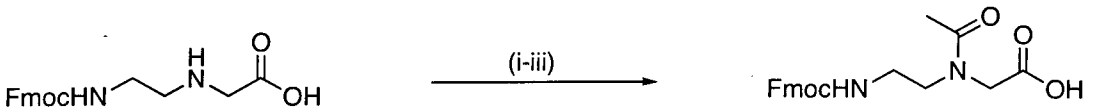
Reaction N°	Activation time (min)	Reaction colour on addition of backbone	Product detectable in mass spectrum of crude reaction ?
1	5	No colour	No
2	10	Pale yellow	No
3	15	Bright yellow	Yes
4	20	Brown/yellow	Yes
5	25	Brown	Yes

Table 2.1 Effect of activation time at 0°C on anhydride formation

Entry 3 in table 2.1 was thought to be the optimal reaction time. Analysis of crude reaction 3 by NMR estimated the product and starting materials to be approximately equimolar; no side products were apparent. Based on the evidence it appeared that the anhydride decomposed as the reaction exceeded 50% completion. Suppression of the side reaction was achieved with temperature control. The reaction was performed at -15°C (ice/salt, 3:1) and took on a bright yellow colour after ten minutes, and remained so indefinitely at that temperature. After 30 minutes **89** was added to the reaction then and allowed to stir while reaching RT over 1hr. The reaction was worked up, however after 48h the product had failed to precipitate. NMR analysis of the crude reaction mixture showed equimolar amounts of product and starting materials. Anhydride

formation should have been closer to completion given the greater reaction time, implying that another variable was affecting the yield of **35**.

It appeared that amine **89** was reacting sluggishly with the anhydride, possibly because the amine did not react well with, nor was easily acylated by, anhydrides. Additionally, the anhydride might be sufficiently sterically hindered to the point where the rate of amine acylation is retarded to be comparable with the rate of hydrolysis.
An attempt was made to acylate **89** using acetic anhydride under the same reaction conditions (Scheme 2.33).

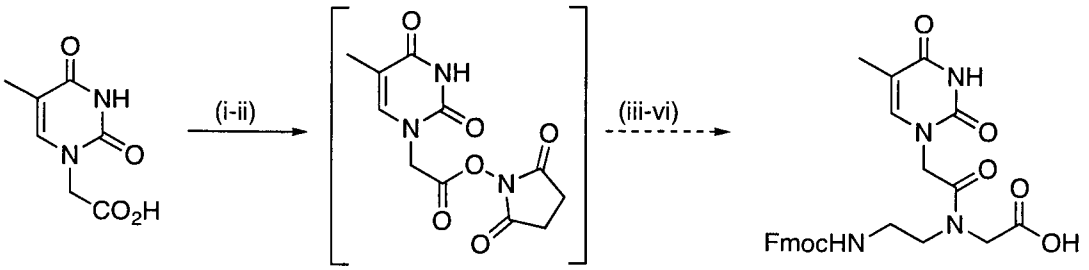


89 **92** ~ 60%

(i) Acetic anhydride, triethylamine, H_2O/MeCN (3:4), -15°C, 25 min
(ii) RT, 30 min
(iii) Adjust to pH 3 with 3M HCl at 0°C then RT, 18 h

Scheme 2.33 Acylation of backbone with acetic anhydride

92 was obtained in approximately 3:2 ratio with unreacted **89**. Given that acetic anhydride is both small and very reactive, this result suggests that steric congestion of **91** was not to blame for incomplete acylation of **89**. It is probable that the basic aqueous solvent necessary for dissolution of **89** is partially hydrolysing **91**. *N*-Hydroxysuccimyl esters are known to be relatively stable in aqueous conditions and competent electrophiles for amines; the nucleophilicity of **89** could be better gauged by reaction with such an ester. Thyminyl acetic acid was converted to **93** and treated with the standard aqueous solution of **89** (Scheme 2.34).



38 93 35

(i) N-hydroxysuccinimide, NMM,CH_2Cl_2, 0°C, 5 min
(ii) Dicyclohexylcarbodiimide, 30 min, RT
(iii) **89**, H_2O/MeCN(3:4), NEt_3, 30 min
(iv) Adjust to pH 3 with 3M HCl at 0°C then stir at RT, 18 h

Scheme 2.34 Attempted T-PNA monomer synthesis via N-hydroxysuccinimyl Thymine acetate

The reaction returned only starting materials. Intermediate **93** was not isolated but the precipitation of insoluble dicyclohexylurea *before* addition of **89** inferred its presence in the reaction. In hindsight, the use of CH_2Cl_2 as the solvent was probably to blame for the failure of the reaction. Backbone **89** is not soluble in CH_2Cl_2 and the reactants may have partitioned between the aqueous and chlorinated organic solvents in the reaction.

It was decided to investigate the purification of **38** by chromatography. Chromatographic purification of **35** was not trivial and after a great deal of effort, **35** was recovered pure in 10% yield.

Synthesis of the other monomers was then attempted - revealing an interesting experimental detail, noted during the activation of Bhoc cytosyl acetic acid. Activation of the pyrimidine **73** was accomplished with *inverse-addition* of pivaloyl chloride; **73** was added to pivaloyl chloride. An apparently innocuous difference at first glance, inverse-addition results in C-PNA precipitating readily from solution (vide infra) under the (nearly identical) reaction conditions used for T-PNA synthesis. Given the structural similarities between cytosine and thymine (and the urgent need for a better method to isolate T-PNA), the coupling of **38** and **89** was attempted with this method of activation in an otherwise unaltered procedure. The experiment was successful and the T-PNA monomer precipitated from the crude worked up reaction mixture and was recovered in 45% yield. The reaction was repeated twice on large scale and the results were reproducible and reliable.

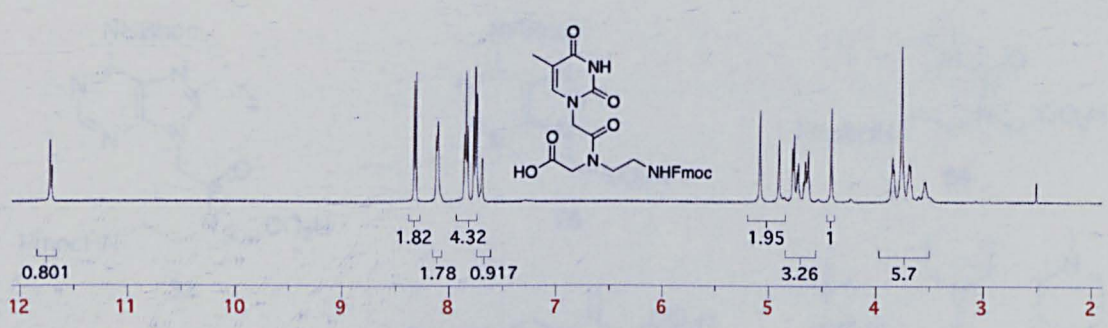


Figure 2.8. ^{1}H NMR spectrum of C-PNA purified by trituration

2.2.2 Adenyl PNA monomer

Initially the method of Coull was followed.[7] Activation of **75** at 0°C gave similar problems encountered with the activation of **38**; the crude anhydride became brown after 10-15 min and no product could be retrieved. Applying the modifications learnt during the synthesis of **35**, a crude mixture consisting mostly of A-PNA was generated (Figure 2.9).

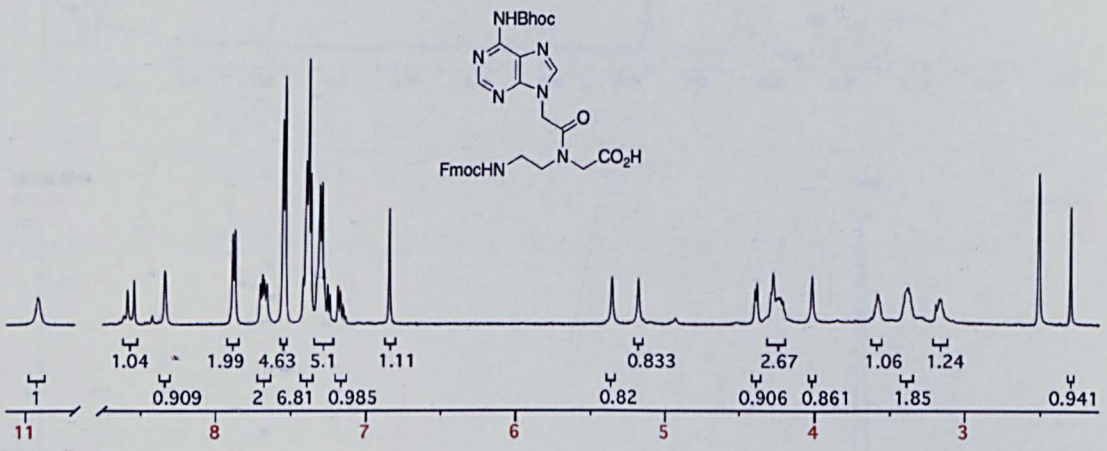


Figure 2.9. ^{1}H NMR spectrum of impure A-PNA

The purification method of Coull did not work. The crude reaction residue was dissolved in a minimum amount of hot methanol and added drop wise to stirred ice-water. Recrystallisation of the recovered colourless precipitate was attempted from ether/dioxane; the recrystallisation caused the solid to oil and would not crystallise.

It was not trivial to separate **32** from the other components of the crude reaction mixture under the constraints imposed by the labile protecting groups (Figures 2.10 and 2.11).

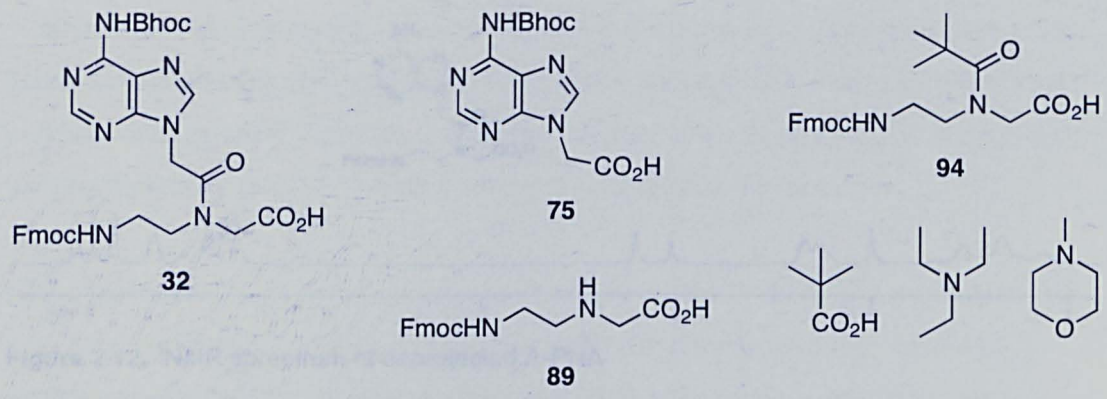


Figure 2.10 Crude reaction mixture components in A-PNA monomer synthesis

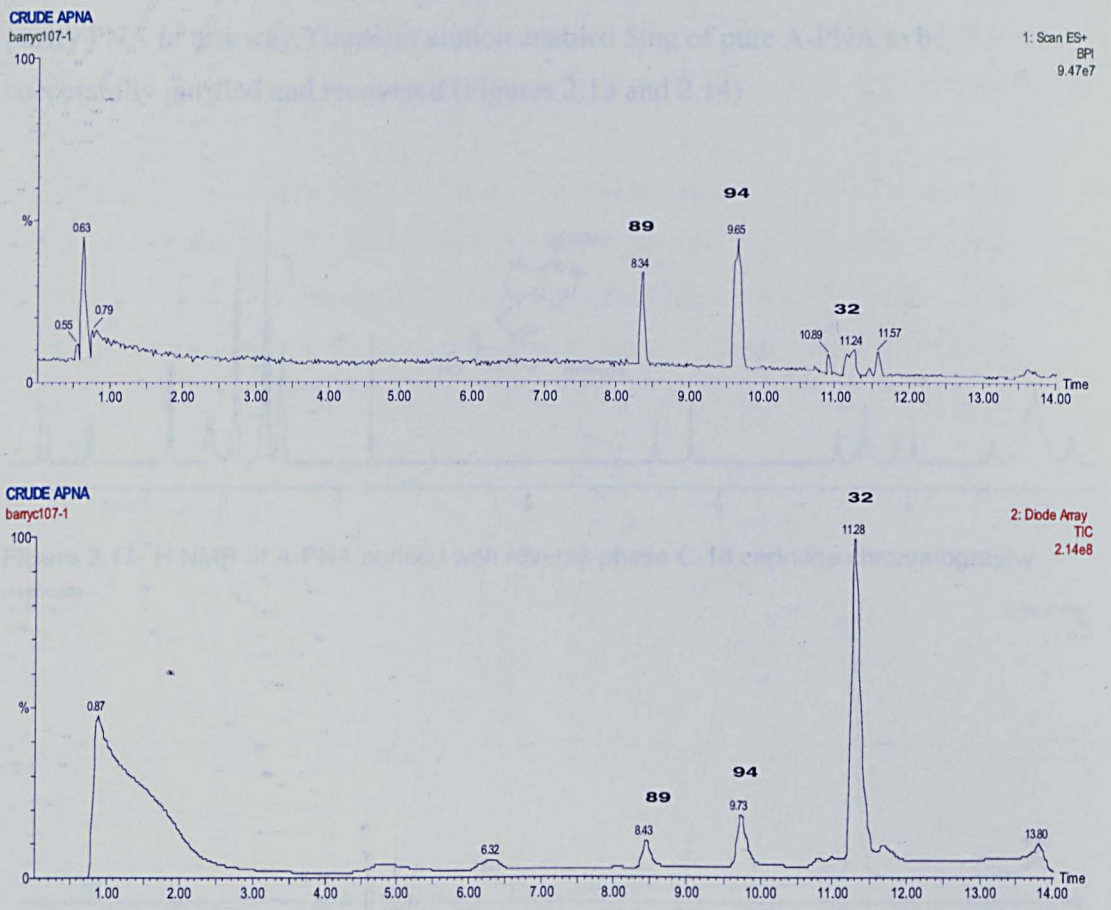


Figure 2.11 LC-MS traces of crude A-PNA

The crude reaction was dissolved in hot methanol and adsorbed onto silica to examine if the Bhoc group would survive flash chromatography. Despite rapid chromatograpghy, it was found that the Bhoc group was cleaved (Figure 2.12).

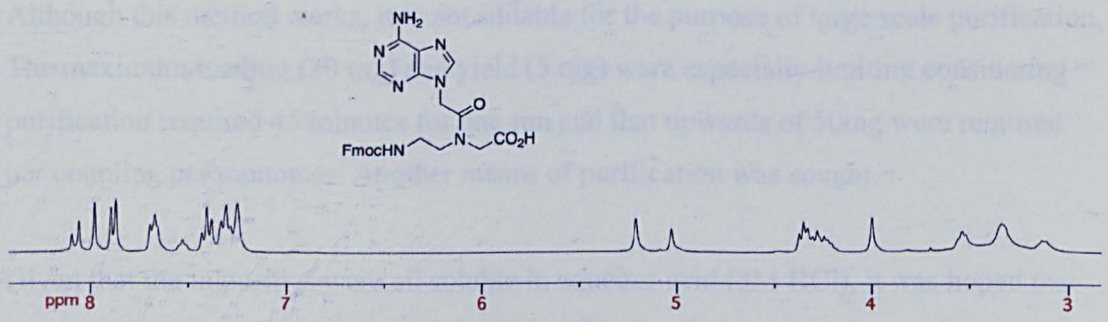


Figure 2.12. ^{1}NMR sprectrum of deprotected A-PNA

Reverse phase chromatography was then investigated to determine if it was viable to purify PNA in this way. Gradient elution enabled 5mg of pure A-PNA to be successfully purified and recovered (Figures 2.13 and 2.14)

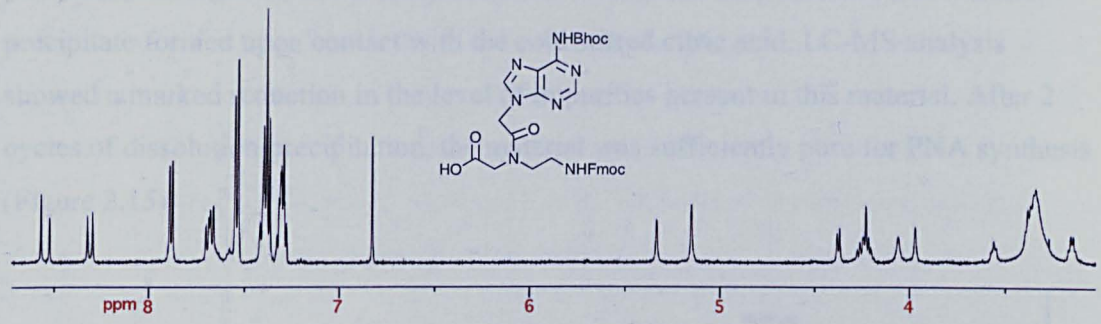


Figure 2.13 ^{1}H NMR of A-PNA purified with reverse phase C-18 cartridge chromatography

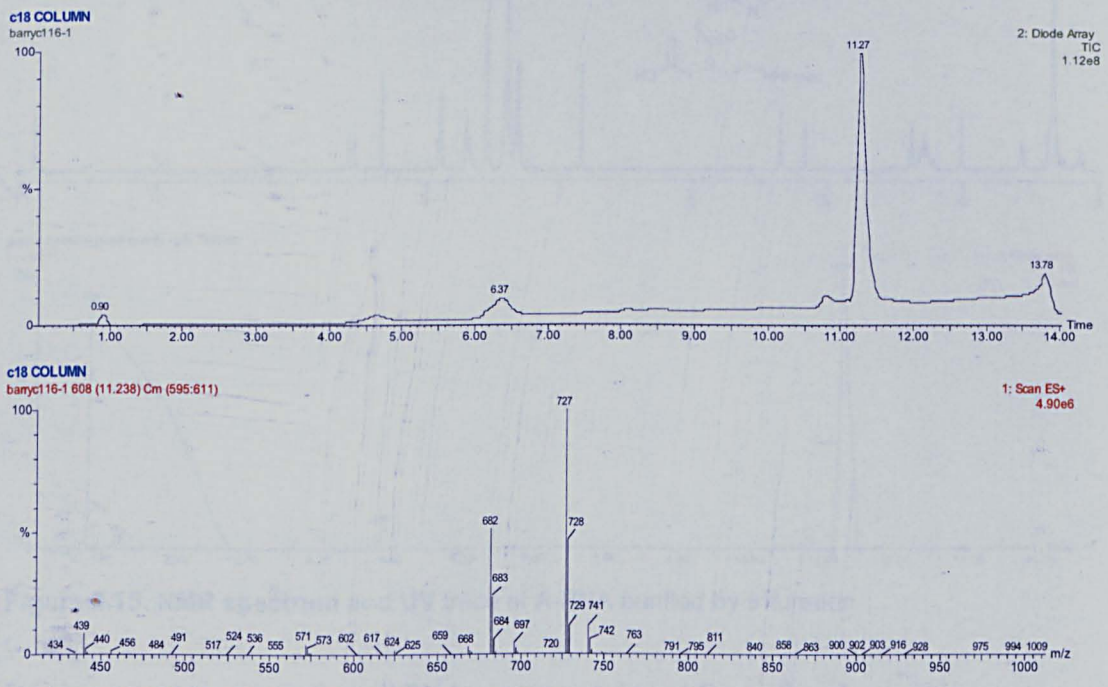


Figure 2.14 LC-MS data for A-PNA purified with C-18 reverse phase chromatography

Although this method works, it is not suitable for the purpose of large scale purification. The maximum loading (20 mg) and yield (5 mg) were especially limiting considering purification required 45 minutes for one run and that upwards of 50mg were required per coupling per monomer. Another means of purification was sought.

Given that the impurities were all soluble in aqueous acid (3M HCl), it was hoped that they could removed from the reaction mixture by dissolution in weakly acidic 10% citric acid solution. It was known that **32** was insoluble in citric acid solution and that the Bhoc group is stable for short periods of time in that media. The method devised was a modification of the Coull procedure whereby the crude reaction mixture was dissolved in a minimum amount of hot methanol but added to 10% citric acid instead of water.

A methanolic solution of the crude monomer was added dropwise and a colourless precipitate formed upon contact with the cold stirred citric acid. LC-MS analysis showed a marked reduction in the level of impurities present in this material. After 2 cycles of dissolution/precipitation, the material was sufficiently pure for PNA synthesis (Figure 2.15).

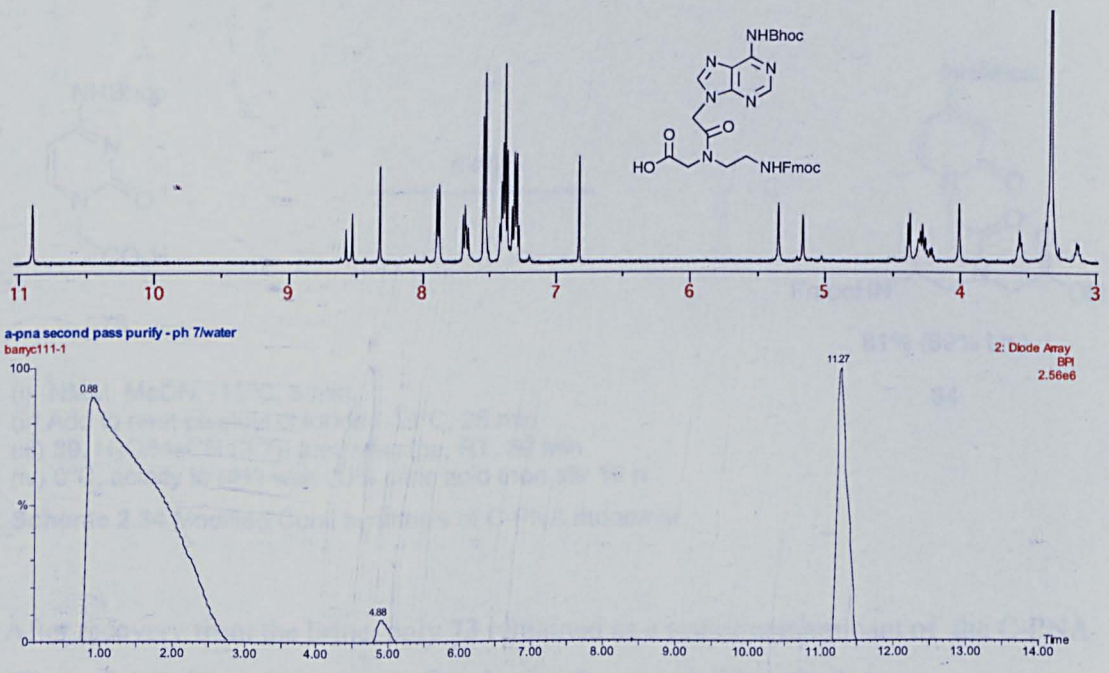


Figure 2.15. NMR spectrum and UV trace of A-PNA purified by trituration

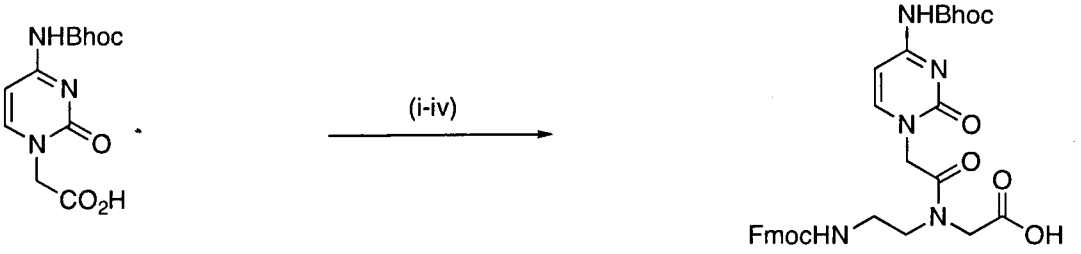
It is important to note is that all PNA monomers do not fly well under positive ionisation mass spectrometry conditions and are not readily observed. They are however

(like the starting materials they are made from) UV active. A UV trace that only contains UV activity from the target monomer (and no signal from the materials used to make the monomer) is the most useful indication that the monomer does not contain any starting-material contaminants.

Figure 2.15 showed that after two cycles of purification the material is essentially complete. There is a significant loss of material using this approach however; approximately 50% was lost before the desired level of purity is achieved. The process is scalable however and multigram quantities of pure A-PNA have been isolated in this manner.

2.2.3 Cytosyl PNA Monomer

In the same manner as all other PNA monomer syntheses, the method of Coull was found to work when the reaction temperature was lowered to -15°C. The reaction used inverse addition of the acid **73** to neat pivaloyl chloride to form the anhydride (Scheme 2.34). The crude material required an additional 24h stirring as a slurry in saturated brine to remove all the citric acid from the mixture.



73

81% (89% Lit.)

34

(i) NMM, MeCN, -15°C, 5 min,
(ii) Add to neat pivaloyl chloride, -15°C, 25 min
(iii) **89**, H_2O/MeCN (3:7), triethylamine, RT, 30 min
(iv) 0°C, acidify to pH3 with 20% citric acid then stir 18 h

Scheme 2.34 Modified Coull synthesis of C-PNA monomer

After recovery from the brine, only **73** remained as a minor contaminant of the C-PNA (Figure 2.16). **73** could not be removed using the citric acid method.

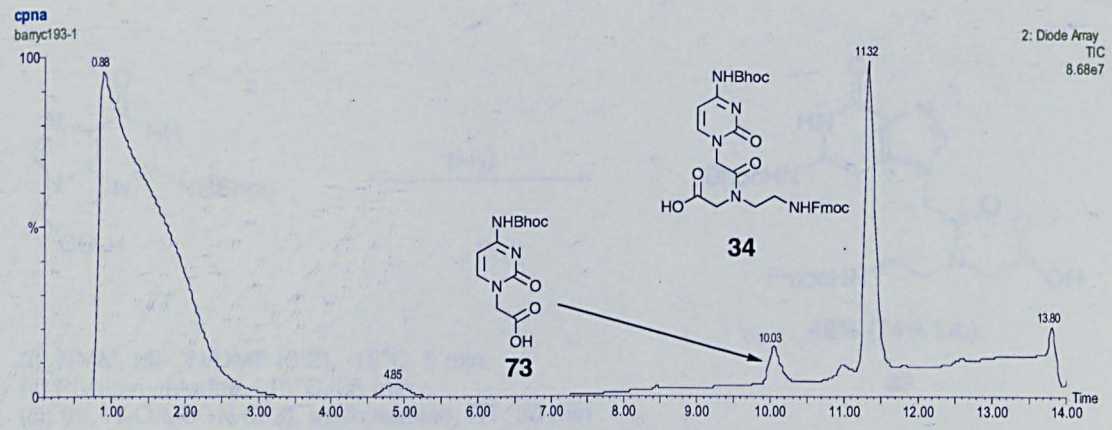


Figure 2.16 LC-MS data of impure C-PNA

Acid **73** is sparingly soluble in all common solvents and difficult to remove from **34**. It was found that warming the crude mixture in a minimal amount of DMF dissolved only **34** and allowed the **73** to be removed by filtration. **34** could then be precipitated from the filtrate by dropwise addition to stirred ice water and was recovered with minimal loss of material.

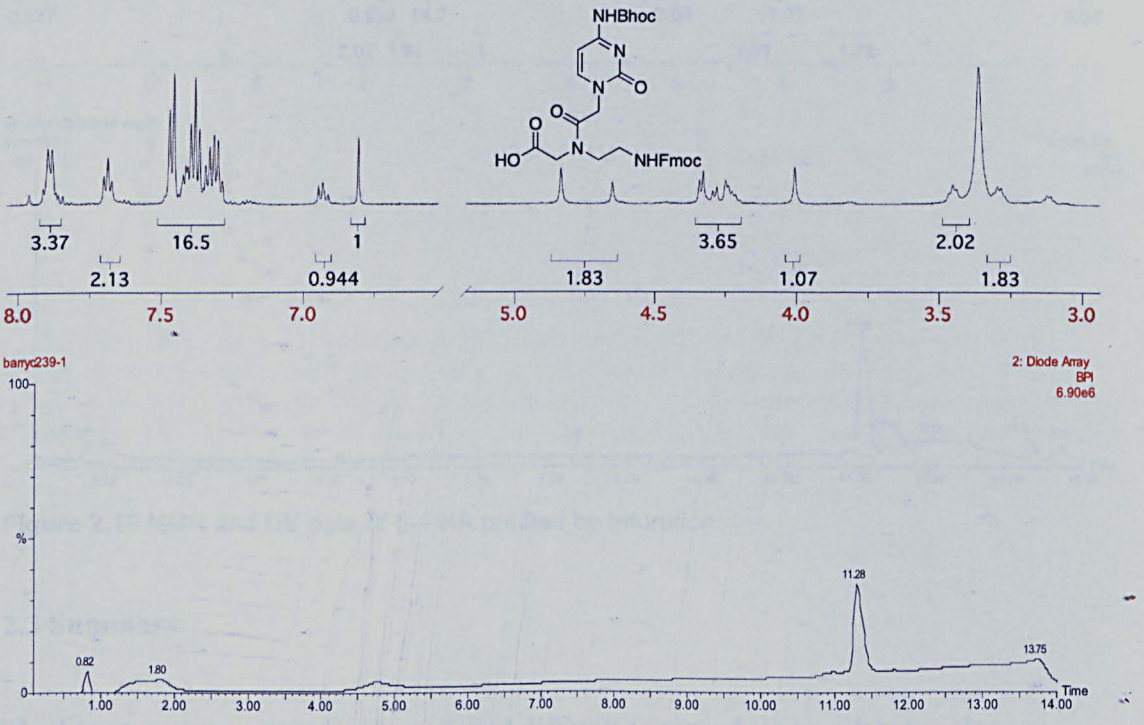
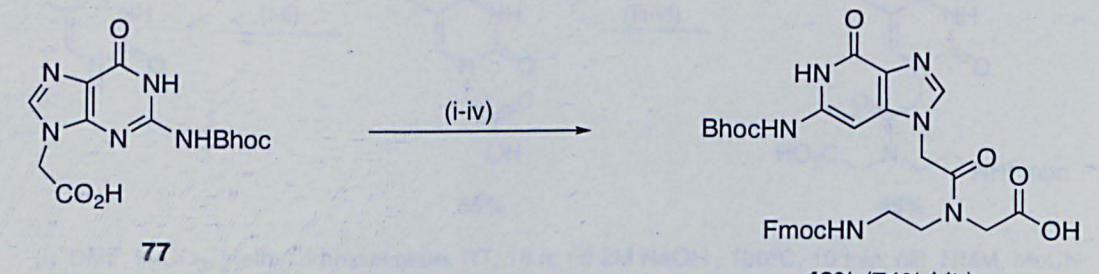


Figure 2.17 NMR and UV data of C-PNA purified by trituration

2.2.4 Guanyl PNA monomer

Following the procedure of Coull, enhanced with the same modifications used to purify **32** and **34**, the G-PNA monomer was obtained (Scheme 2.35, Figure 2.18).



77

49% (74% Lit.)

33

(i) NMM, MeCN/DMF (3:2), -15°C, 5 min,
(ii) Pivaloyl chloride, -15°C, 25 min
(iii) **89,** H_2O/MeCN (2:3), triethylamine, RT, 30 min
(iv) 0°C, acidify to pH4 with 20% citric acid then stir 18 h

Scheme 2.35 Modified Coull synthesis of G-PNA monomer

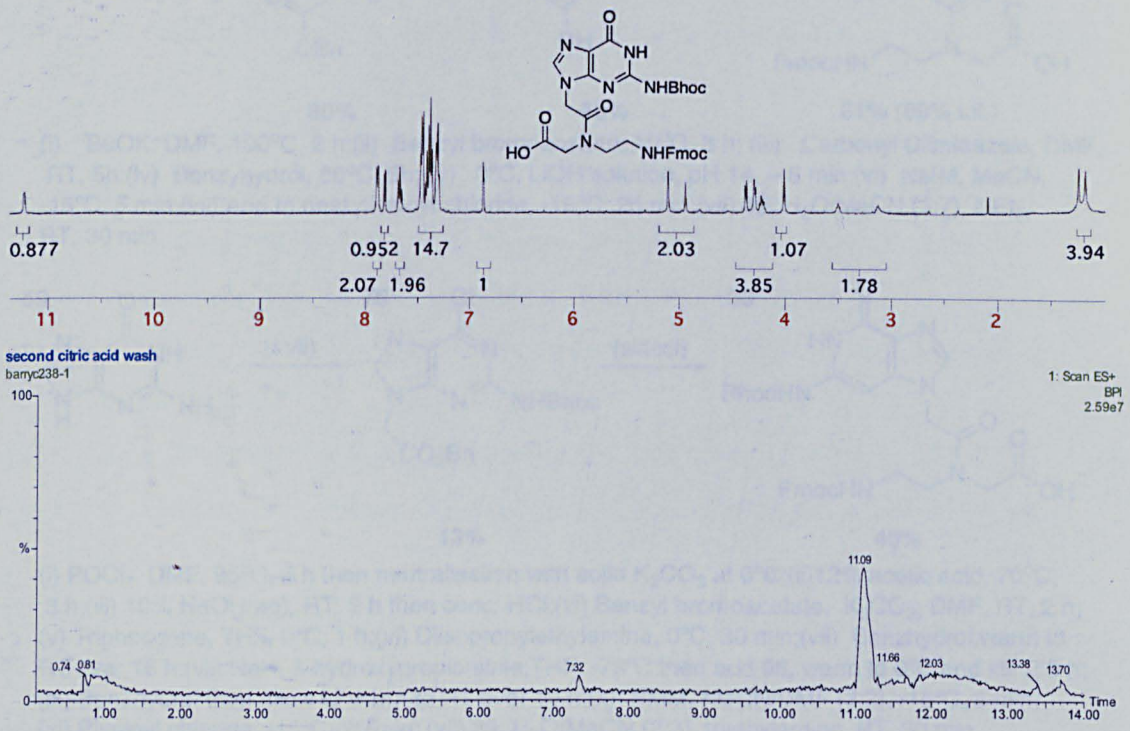
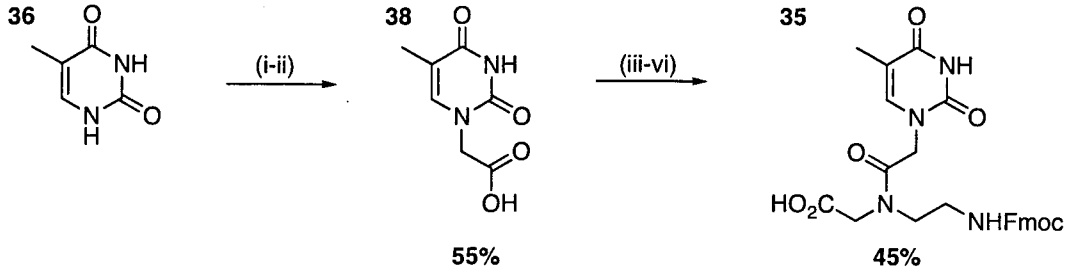


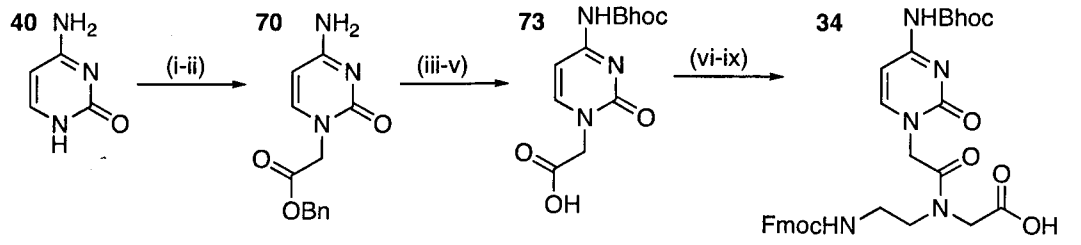
Figure 2.18 NMR and UV data of G-PNA purified by trituration

2.3 Summary

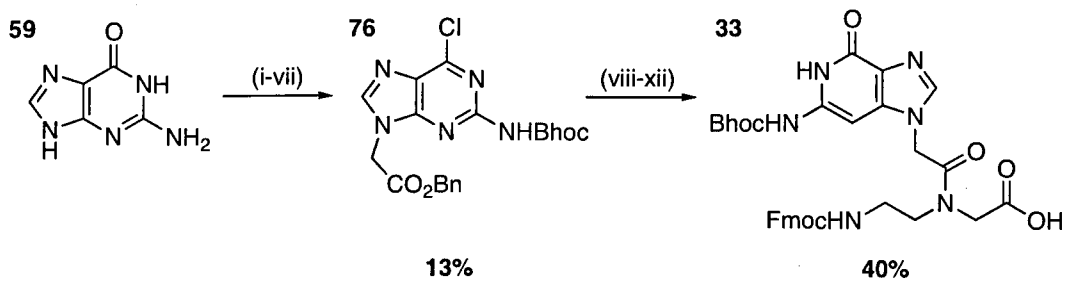
The literature procedures had been found difficult to reproduce on a large scale. Extensive optimisation was required. Modified routes (Scheme 2.36) that are able to produce all four PNA monomers on a multigram scale,were developed. Summaries of the routes are shown in Scheme 2.36



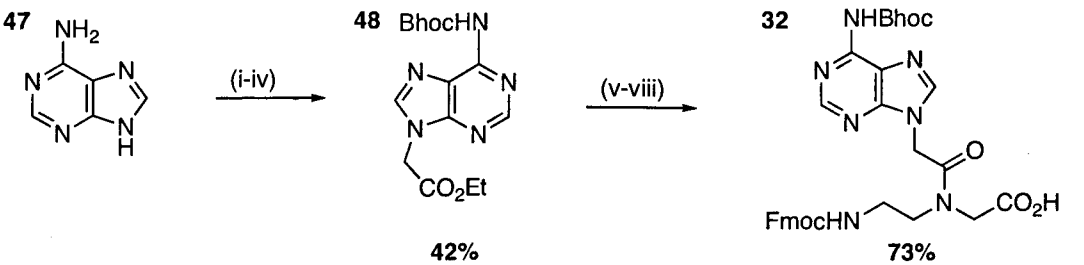
(i) DMF, K_2CO_3, Methyl bromoacetate, RT, 18 h; (ii) 2M NaOH , 100°C, 10 min; (iii) NMM, MeCN , -15°C, 5 min;(iv) Add to neat pivaloyl chloride, -15°C, 25 min;(v) **89,** H_2O/MeCN (3:7), NEt_3, RT, 30 min;



(i) ^{t}BuOK, DMF, 100°C, 2 h;(ii) Benzyl bromoacetate, 10°C, 3 h; (iii) Carbonyl Diimidazole, DMF, RT, 5h;(iv) Benzyhydrol, 60°C, 6h; (v) 0°C, LiOH solution, pH 14, ~ 6 min;(vi) NMM, MeCN, -15°C, 5 min;(vii) Add to neat pivaloyl chloride, -15°C, 25 min;(vii) **89,** H_2O/MeCN (3:7), NEt_3, RT, 30 min



(i) $POCl_3$, DMF, 95°C, 5 h then neutralisation with solid K_2CO_3 at 0°C;(ii)12% acetic acid, 70°C, 3 h;(iii) 10% NaOH(aq), RT, 2 h then conc. HCl;(vi) Benzyl bromoacetate, K_2CO_3, DMF, RT, 2 h; (v) Triphosgene, THF, 0°C, 1 h;(vi) Diisopropylethylamine, 0°C, 30 min;(vii) Benzhydrol,warm to RT over 16 h;(vii) NaH, 3-hydroxypropionitrile,THF, -78°C then add **65,** warm to 0°C and stir 2.5 h; (ix) Benzhydrol then allow to warm to RT over 16 h;(x) NMM, MeCN/DMF (3:2), -15°C, 5 min; (xi) Pivaloyl chloride, -15°C, 25 min; (xii) **89,** H_2O/MeCN (2:3), triethylamine, RT, 30 min



(i) NaH, DMF, RT, 3.5 h;(ii) Ethyl bromoacetate, <30°C, 19 h;(iii) Carbonyl diimidazole, DMF, 105°C, 2 h;(iv) 105°C -> 95°C then benzhydrol, cool to RT over 16 h;(v) LiOH, 0°C, 6min;(vi) NMM, MeCN, -15°C, 5 min;(vii) Pivaloyl chloride, -15°C, 25 min; (viii) **89,** H_2O/MeCN (1:1), NEt_3, RT, 30 min

Scheme 2.36 Summary of PNA monomer synthesis routes

With the four PNA monomers in hand the synthesis of the linker units was undertaken.

2.4 References

[1] G. Breipohl, J. Knolle, D. Langner, G. O'Malley, E. Uhlmann, *Bioorg. Med. Chem. Lett.* **1996**, *6*, 665.

[2] D. W. Will, G. Breipohl, D. Langner, J. Knolle, E. Uhlmann, *Tetrahedron* **1995**, *51*, 12069.

[3] K. L. Dueholm, M. Egholm, C. Behrens, L. Christensen, H. F. Hansen, T. Vulpius, K. H. Petersen, R. H. Berg, P. E. Nielsen, O. Buchardt, *J. Org. Chem.* **1994**, *59*, 5767.

[4] T. Kofoed, H. F. Hansen, H. Orum, T. Koch, *J. Pept. Sci.* **2001**, *7*, 402.

[5] G. Aldrian-Herrada, A. Rabie, R. Wintersteiger, J. Brugidou, *J. Pept. Sci.* **1998**, *4*, 266.

[6] S. A. Thomson, J. A. Josey, R. Cadilla, M. D. Gaul, C. F. Hassman, M. J. Luzzio, A. J. Pipe, K. L. Reed, D. J. Ricca, et al., *Tetrahedron* **1995**, *51*, 6179.

[7] J. M. Coull, M. Egholm, R. P. Hodge, M. Ismail, S. B. Rajur, (Perseptive Biosystems, Inc., USA). Application: WO, **1996**, p. 138 pp.

[8] P. J. Finn, N. J. Gibson, R. Fallon, A. Hamilton, T. Brown, *Nucleic Acids Res.* **1996**, *24*, 3357.

[9] B. E. Watkins, J. S. Kiely, H. Rapoport, *J. Am. Chem. Soc* **1982**, *104*, 5702.

[10] G. R. Geen, T. J. Grinter, P. M. Kincey, J. R.L., *Tetrahedron Lett.* **1990**, *46*, 6903.

[11] H. Hagen, P. Raatz, (BASF A.-G., Germany). Application: DE, **1993**, p. 4 pp.

[12] R. W. Balsiger, J. A. Montgomery, *J. Org. Chem.* **1960**, *25*, 1573.

[13] T. Hayashi, K. Nishiwaki, (Sumika Fine Chemicals Co., Ltd., Japan). Application: EP, **1999**, p. 10 pp.

[14] R. Ramert, R. Sobotta, (Boehringer Ingelheim KG, Germany). Application: DE, **1995**, p. 3 pp.

[15] F. Debaene, N. Winssinger, *Org. Lett.* **2003**, *5*, 4445.

[16] G. D. Breipohl, E. D. Uhlmann, J. D. Knolle, (Hoechst A.-G., Germany). Application: EP, **1995**, p. 31 pp.

Chapter 3

Chapter 3 - Linker Synthesis

Before the chimeras could be assembled, a pair of linker structures needed to be designed and synthesised. These linker units would be used to join PNA and polypeptides together so that PNA-directed ligation would be possible. The aim was to design and synthesise these linker units on a multigram scale.

3.1 SPROUT aided design of PNA linkers

To enable PNA to direct the ligation of peptides, the PNA cannot be attached directly to the peptides, intervening spacer units must be placed between each PNA strand and its associated peptide. These linker units perform two critical roles. The first function of the linker units is to bridge the inter-strand region and allow the peptides to meet and ligate. Each linker also acts as a temporary tether to allow attachment of the PNA directing groups to the peptides; post-ligation the linker can be induced to detach from the polypeptide, allowing the isolation of a natural peptide target.

Design of the linker units is critical to the success of the ligation reaction as their structure and conformation will have a strong bearing on the ligation reaction. Computer aided design of the linker units was used to expediate the design of a pair of linker units that would promote the ligation reaction. The molecular docking suite SPROUT was modified to accommodate this task.[1] It was thought that the quickest way to generate a pair of suitable linkers would be to modify the terminal PNA base-pair from a PNA helix (Figure 3.1). By removing the ends of the PNA backbone but retaining the base pair a useful scaffold for growing the linkers would be generated rapidly. Furthermore, as these linkers are derived from a PNA helix, disruption to helix formation during a ligation experiment is minimised.

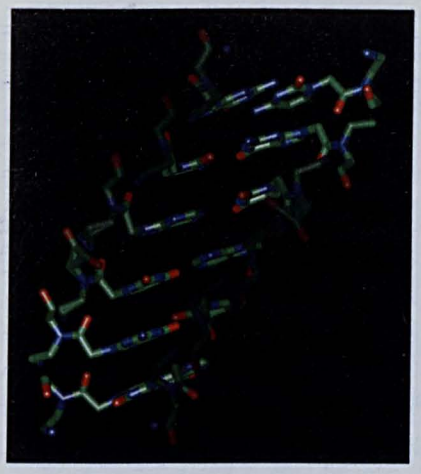


Figure 3.1 PNA duplex crystal structure

A crystal structure of a short PNA helix (Figure 3.1) was uploaded into the SPROUT program and the *C* and *N* terminus from one end of the helix were trimmed back and tagged as the site of attachment (node) for each linker (Figure 3.2).

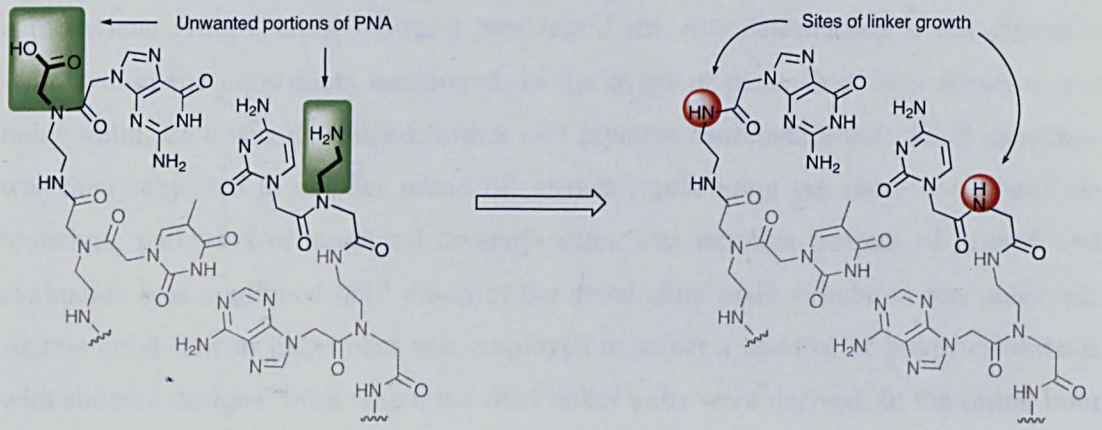


Figure 3.2 Modification of PNA duplex into partial linker units

SPROUT was instructed to 'grow' the linker units simultaneously from both nodes and join together in the middle. Growth was subject to the restrictions that fragment **95** has to appear once and only once in the final bridging structure and that growth may not occur outside a well defined boundary (Figure 3.3).

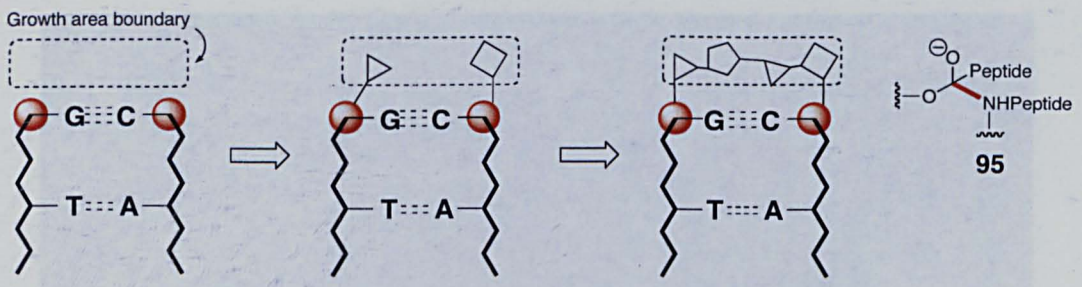


Figure 3.3 Growth of linkers

Fragment **95** represents one of the two tetrahedral intermediates that could be formed by the addition of one peptide *N*-terminus to the activated *C*-terminus of another (bond formed shown in bold red in Figure 3.3); effectively the transition state structure. Inclusion of this fragment optimises the linker design to promote formation of the transition state structure and thus promotes the ligation reaction. Both enantiomers of **95** were created and tested for each complete candidate structure generated by SPROUT.

Growth is accomplished by SPROUT attaching a chemical fragment to a node (such as a methylene group), chosen from a preselected list, and determining if the choice is valid within the constraints mentioned. In the event of more than one chemical unit being valid, each was developed further as a separate candidate linker. Each candidate was then subjected to another round of 'growth', generating yet more candidates *via* branching pathways of structural diversification. The iterative process of growth and evaluation was continued until union of the developing node structures was achieved. At this point human judgement was employed to select a handful of linker candidates with suitable designs, from which the final linker units were derived. In the initial bout of candidate design, two sets of related linker candidates were deemed promising (Figures 3.4, 3.5). These linkers were chosen as they included an aryl core; this feature was desired to allow for the opportunity of post-ligation cleavage of the linker/PNA from the peptide by protolysis of the benzylic linkages (Figure 3.6).

The two candidate structures were identical except that the aryl core of the PNA to peptide linker has a 1,4 (para) substitution pattern (Figure 3.4) on one structure while the other has 1,3 (meta) substitution (Figure 3.5). The 1,4 linker was rejected on the grounds that the overall bridging structure formed during ligation was sufficiently more crowded than that made using the competing 1,3 linker.

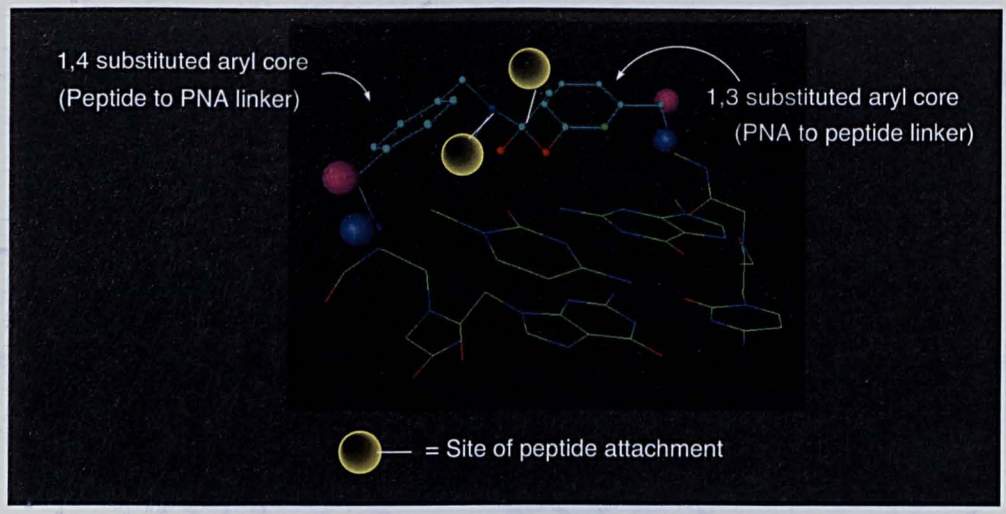


Figure 3.4 Sprout modelled ligation transition-state using rejected candidate linker structures

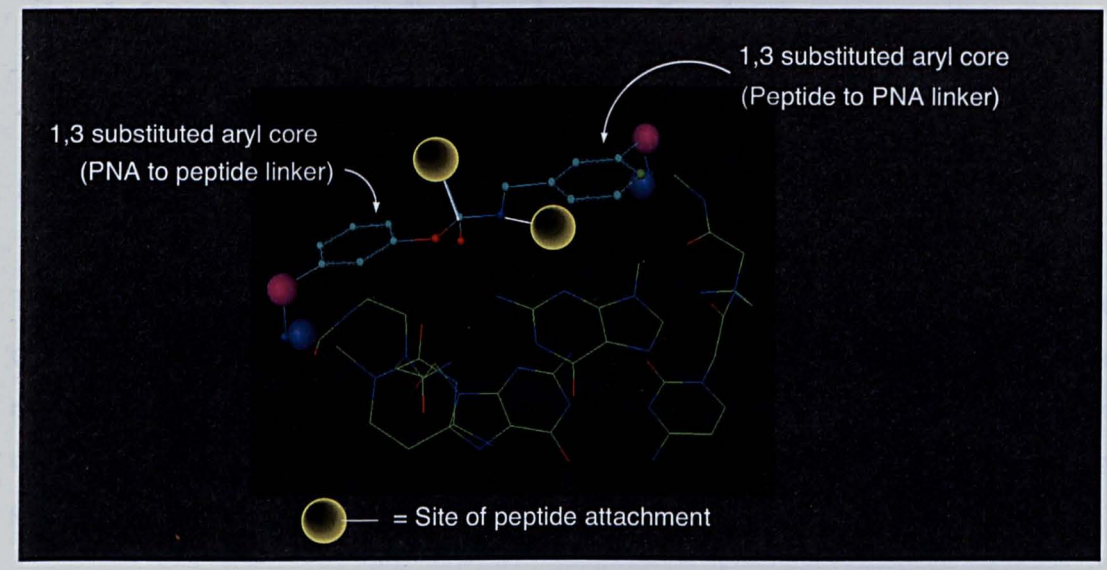
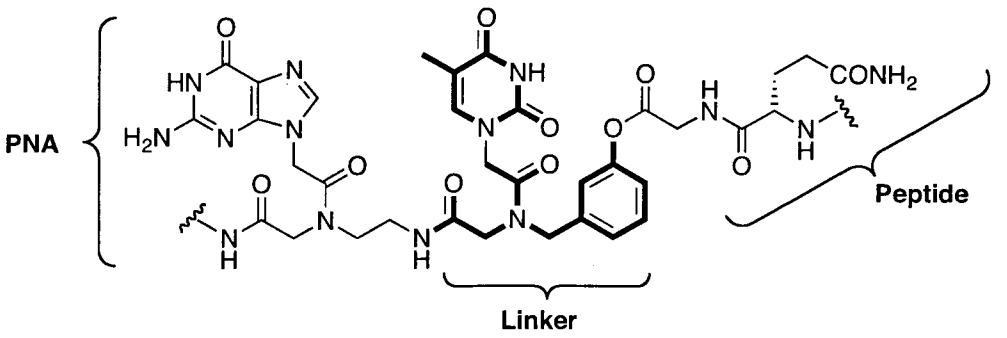


Figure 3.5 Sprout modelled ligation transition-state using best-candidate linker structures

Studying the best-candidate structure generated by SPROUT and regressing to a period pre-ligation revealed the linkages that would be required to join the PNA and peptides together (Figure 3.6).



Acid labile bond shown in bold red; bond is activated to cleavage after ligation has occurred

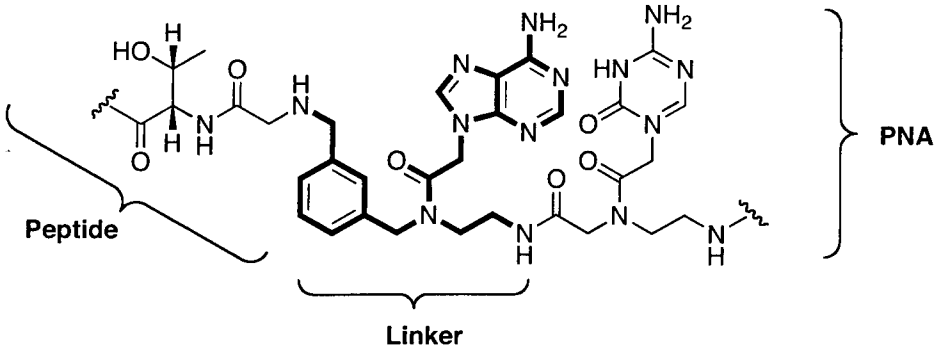


Figure 3.6 Diagram to show how the linkers are incorporated into the chimeras

To incorporate these linkages (Figure 3.6, bold bonds) into the final PNA-peptide chimeras, linker units that would be suitable for use with fmoc-based SPPS were required. The final linker designs from the SPROUT design for use in the construction of the PNA-peptide chimeras were therefore derived using a bhoc/fmoc protection strategy are shown in Figure 3.7.

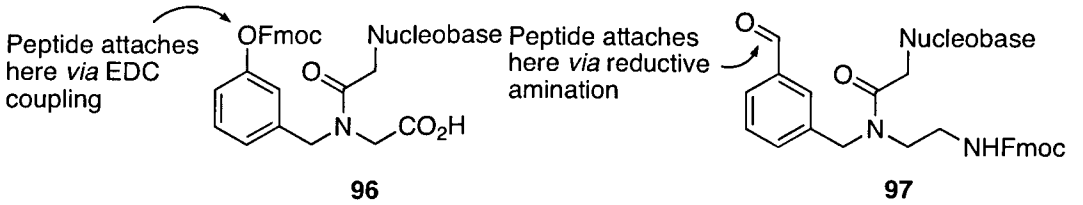


Figure 3.7 SPROUT designed linkers

With the final linker targets in hand, planning and execution of the linker syntheses was undertaken.

3.2 PNA to peptide linker : Linker 1

The SPROUT designed linker **96** was retrosynthesised as shown in Figure 3.8.

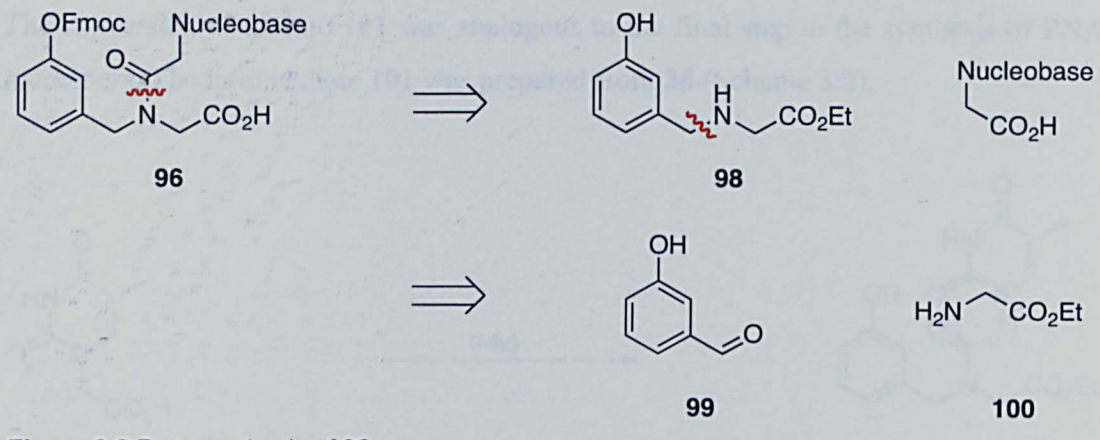
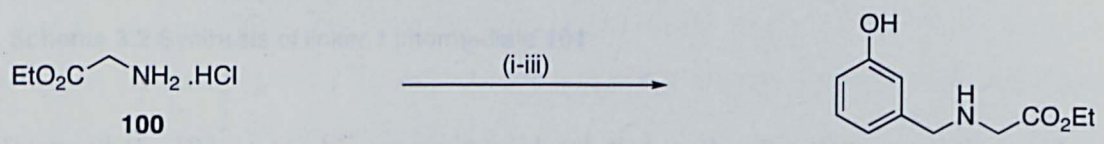


Figure 3.8 Retrosynthesis of **96**

The first intermediate **98** was prepared using the reductive amination procedure of Gebert (Scheme 3.1).[2]



(i) EtOH, potassium hydroxide, 80°C, 10 min
(ii) 3-hydroxybenzaldehyde, RT, 90 min
(iii) $NaBH_4$, RT, 30 min

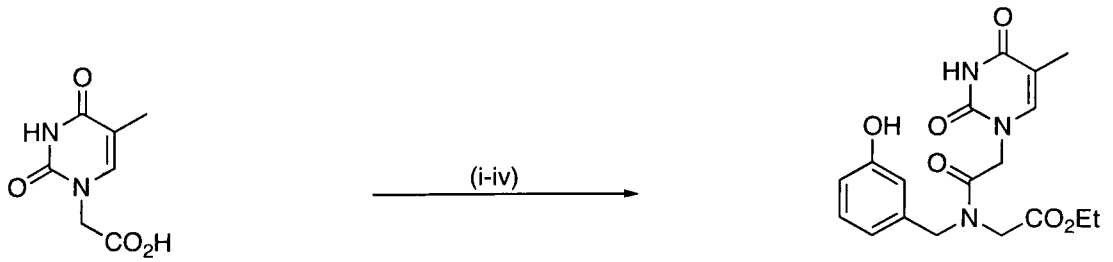
Scheme 3.1 Synthesis of **98** via reductive amination

Condensation was not promoted by addition of dessicating agents ($MgSO_4$, 3Å molecular sieves) or increasing the time allowed for reduction. Indeed, the conditions of Gebert were found to be optimal, and deviation resulted in a more challenging chromatographic purification and lower yield.

Amine **98** was then to be acylated with one of the four nucleobase acetic acids (**38, 73, 75** or **77**). It was decided to use thyminyl acetic acid **38** for the initial synthesis of **96**, given that it is the most readily prepared of the four nucleobase acetic acids. However, to enable a full and robust development of the project methodology, it will be necessary that the linker **96** can be adorned with any of **38, 73, 75** or **77**. The chemistry required to convert intermediate **98** into the final linker unit therefore had to be tailored to tolerate

the acid sensitivity of **73, 75** and **77**. It was planned that all four possible permutations of **96** would be synthesised after conditions for PNA-directed ligation had been developed.

The conversion of **98** into **101** was analogous to the final step in the synthesis of PNA monomers. The intermediate **101** was prepared from **38** (Scheme 3.2).



38

24% (iv-a)
86% (iv-b)

101

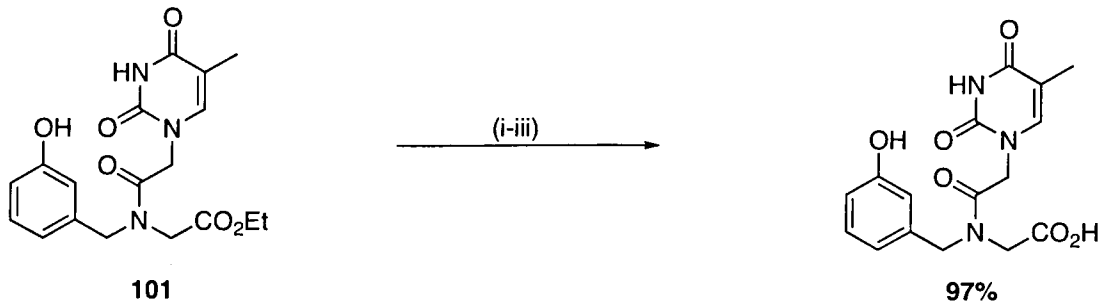
(i) MeCN, NMM, RT, 5 min
(ii) add to pivaloyl chloride, -15°C, 25 min
(iii) **98**, NEt_3, $MeCN/H_2O$ (1:1), 16 h, RT
(iv-a) 20% citric acid, extract with EtOAc
(iv-b) Evaporate to dryness then triturate with 6M HCl

Scheme 3.2 Synthesis of linker 1 intermediate **101**

Intermediate **101** is soluble in citric acid solution and exhaustive extraction with a variety of organic solvents (toluene, ether, CH_2Cl_2, ethyl acetate) gave an optimal yield of 24% using ethyl acetate. These conditions are suitable for the synthesis of the intermediate **96** when any nucleobase is present. However, since **101** did not bear an acid sensitive protecting group the workup was modified to use stronger acids (as per the T-PNA monomer synthesis) and a dramatic increase in yield was achieved. The crude reaction was evaporated to dryness and triturated with a minimal amount of 6M HCl solution to obtain **101** as a slurry. Filtration and simple washing of the slurry precipitate gave **101** in much improved yield. Although untested, it was predicted that the Bhoc group would not survive this procedure.

Hydrolysis of ester **101** is the penultimate step in the synthesis of **96** and the method for this transformation is derived from the prior experience with PNA monomer synthesis (Scheme 3.3). The hydrolysis conditions that were selected had previously been used to

convert **71** and **74** to their corresponding acids, and were used (without modification) to effect near quantitative conversion of **101** into **102**. (Scheme 3.3).

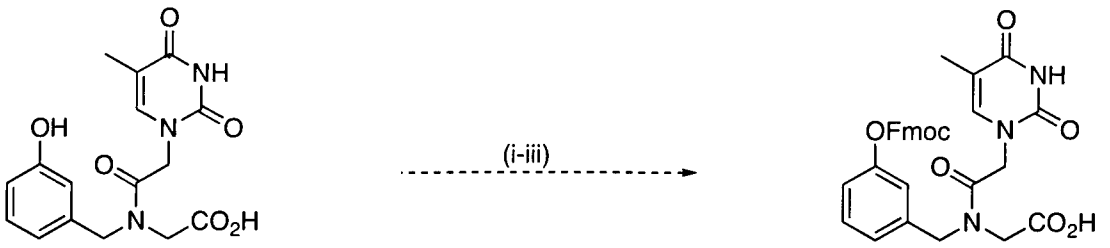


101 **97%** **102**

(i) H_2O, lithium hydroxide, 0°C, 5-10 min
(ii) 20% citric acid solution, 0°C, 1 h
(iii) Saturated brine solution, RT, 24 h

Scheme 3.3 Synthesis of linker 1 intermediate **102**

The final step in the synthesis of **98** was the protection of the phenol group. It was attempted to form the Fmoc carbonate under weakly basic conditions (Scheme 3.4). The Fmoc protection of phenol **102** was considered to be potentially challenging as the phenol nucleophilicity was predicted to be relatively poor. To minimise sluggish reaction kinetics, the most electrophilic Fmoc transfer reagent that was available, Fmoc chloride, was used for the synthesis of **96**.



102 **No reaction** **96**

(i) Sodium bicarbonate solution, DMF, RT, 5min
(ii) Fmoc Chloroformate, DMF, 48h
(iii) Neutralise with 1M $NaHSO_4$

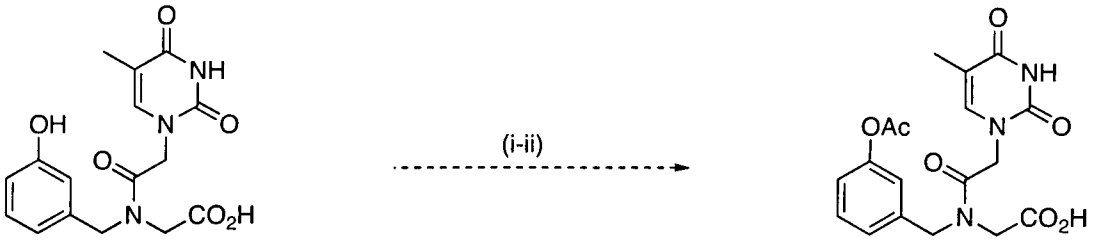
Scheme 3.4. Attempted Fmoc protection of phenol **102**

The crude reaction mixture was purified by chromatography to recover both starting materials. Under the basic conditions, the carboxylate group should not have interfered with the reaction, however to ensure that this was a valid assumption, the reaction was repeated with the ester **101**. The result was identical and only starting materials were

recovered. The phenol **101** is not sufficiently nucleophilic to react with Fmoc chloroformate, and so it was decided that a new protecting group was required.

The new protecting group had to fulfil the same criteria as the Fmoc group. Both protection and deprotection conditions must not affect the acid labile Bhoc group and the protecting group must be stable under standard SPPS conditions.

The immediate and obvious choice for protection of the phenol, partly inspired by solid phase carbohydrate chemistry, was the acetate group. An efficient procedure by Phukan in which phenol is acylated quantitatively seemed a promising approach for the protection of **102** (Scheme 3.5), however the reaction failed to acetylate **102**. [3]



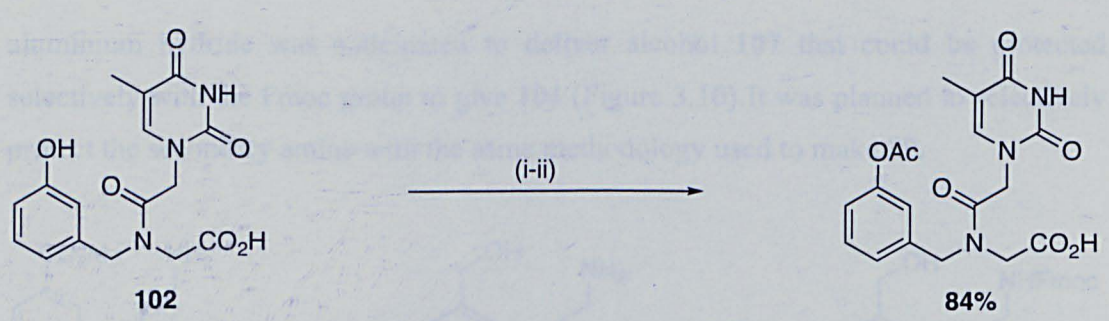
102 **No reaction** **103**

(i) Neat acetyl chloride, 10 mol % iodine, 1 min
(ii) Quench with saturated sodium thiosulfate solution

Scheme 3.5 Attempted acetyl chloride mediated phenol protection

Why acetyl chloride is unable to convert **102** into **103** is not clear. It is counterintuitive to suggest that the carboxylate anion is more nucleophilic than the phenol or that the presence of thymine should interfere with this reaction.

Using the Schotten-Bauman conditions of Chattaway, **102** was successfully acylated to give the first target linker **103** (Scheme 3.6).[4]



102 **84%** **103**

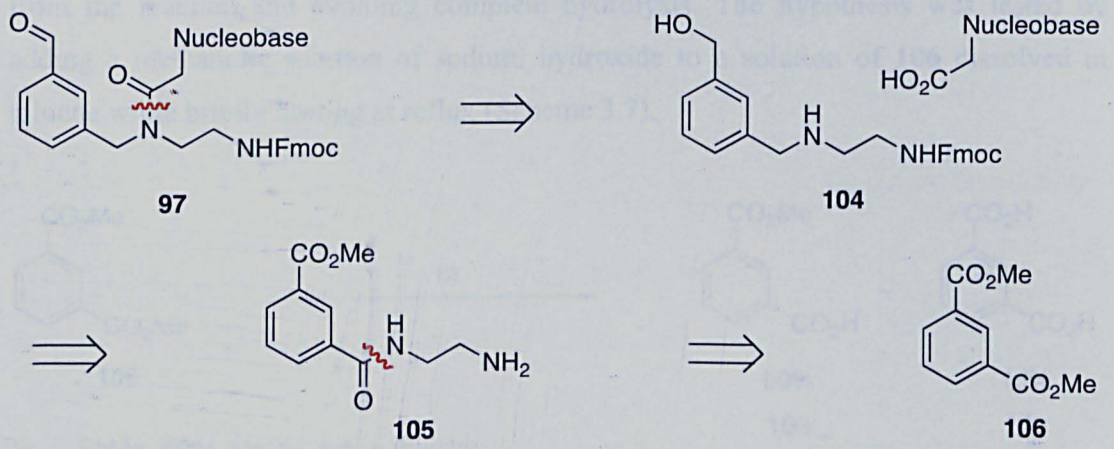
(i) Acetic anhydride, 10% NaOH solution, ice, 1 min
(ii) Quench with 2M hydrochloric acid solution, extract with EtOAc

Scheme 3.6 Schotten-Bauman type acylation of phenol **102**

Interestingly the final product was not initially recovered in good yield – **103** has a high affinity for magnesium sulfate and using this drying agent dramatically lowered the yield to ~ 18%. Changing the dying agent to the less acidic sodium sulfite increased the yield to 65%. To recover the optimal yield for **103** (Scheme 3.6) it was found that drying by azeotropic evaporation of water with toluene is required

3.3 Peptide to PNA linker : Linker 2

SPROUT designed linker **97** was retrosynthesised as according to Figure 3.9.

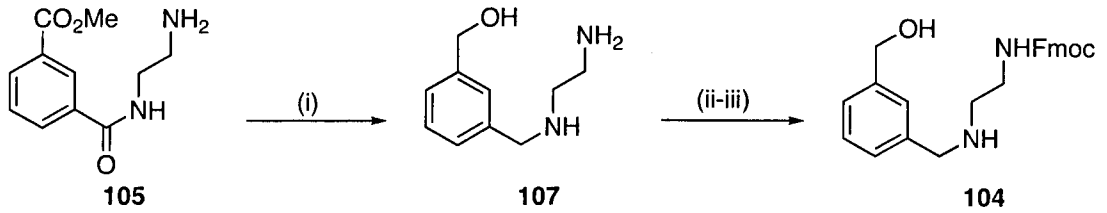


97 **104** **105** **106**

Figure 3.9 Retrosynthesis of **97**

Starting from commercially available **106** it was planned to desymmetrise **106** into the mono acid **107**. Activation of this carboxylic acid and subsequent reaction with a excess of ethylenediamine would give intermediate **105**. A global reduction of **105** with lithium

aluminium hydride was anticipated to deliver alcohol **107** that could be protected selectively with the Fmoc group to give **104** (Figure 3.10) It was planned to selectively protect the secondary amine with the same methodology used to make **89.**



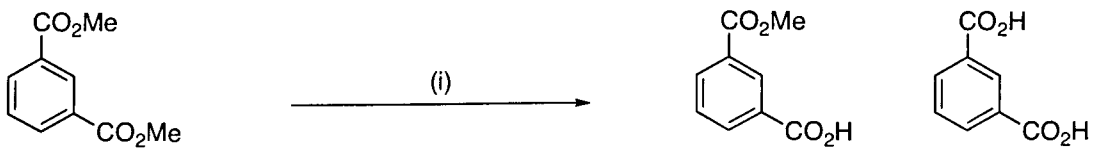
105 **107** **104**

(i) Reduction
(ii) Protection of secondary amine
(iii) Fmoc transfer reagent; unmask secondary amine

Figure 3.10 Proposed synthetic route to **104**

The synthesis of **97** would then be completed with the coupling of **104** with an activated nucleobase acetic acid and a final oxidation of the alcohol into the aldehyde **97.**

Initially a set of conditions that would allow the selective partial hydrolysis of **106** were required. Isolation of **108** was thought to be achievable by performing the reaction in a solvent in which the starting ester **106** is soluble but that the product **108** would be insoluble. In effect, the product salt would precipitate as it is formed, removing itself from the reaction and avoiding complete hydrolysis. The hypothesis was tested by adding a methanolic solution of sodium hydroxide to a solution of **106** dissolved in toluene while briefly heating at reflux (Scheme 3.7).



106 **69%** **16%**

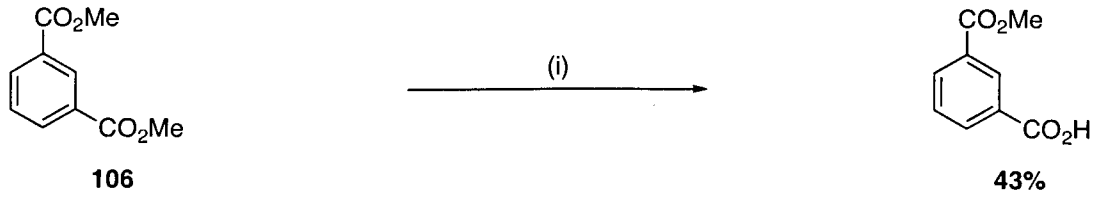
108 **109**

(i) PhMe, KOH, MeOH, reflux, 20 min

Scheme 3.7 Partial hydrolysis of dimethylisophthalate

108 did precipitate (very rapidly) from solution as planned but a significant quantity of the diacid **109** was also formed and found to be inseparable from **108.** A number of

modifications were made to eliminate the formation of the diacid **109** from the reaction (Scheme 3.8).



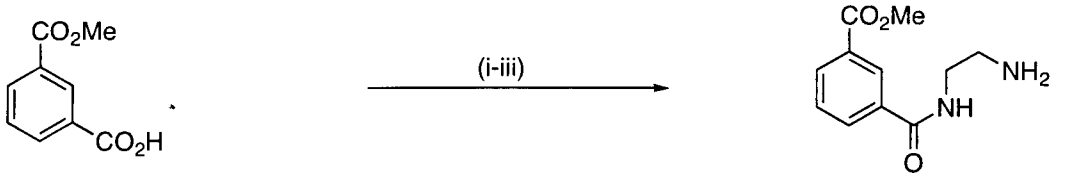
106 **43%** **108**

(i) PhMe, NaOH, MeOH, RT, 16 h

Scheme 3.8 Improved synthesis of monomethylisophthalate

Employing sodium hydroxide as the base and performing the reaction at RT gave a lower yield but there was no contaminating diacid. Unreacted starting material is recovered easily and can be recycled.

Slow addition of a dilute solution of activated **108** to neat ethylene diamine was attempted to gain direct access to **105**. Activation of **108** was attempted by conversion of **108** into the acyl chloride and was followed by *in situ* reaction with ethylene diamine (Scheme 3.9).



108 **Could not be purified**
105

(i) PhMe, $SOCl_2$, reflux, 90 min
(ii) Evaporate to dryness then dissolve in 25 volumes of CH_2Cl_2
(iii) Add dropwise to stirred ethylene diamine, RT

Scheme 3.9 Attempted mono-acylation of ethylene diamine with **108**

Even though the concentration of acyl chloride was low and EDA was high, the diacylated side-product **110** was formed in greater yield than the desired product **105**, as determined by LC-MS (Figure 3.11).

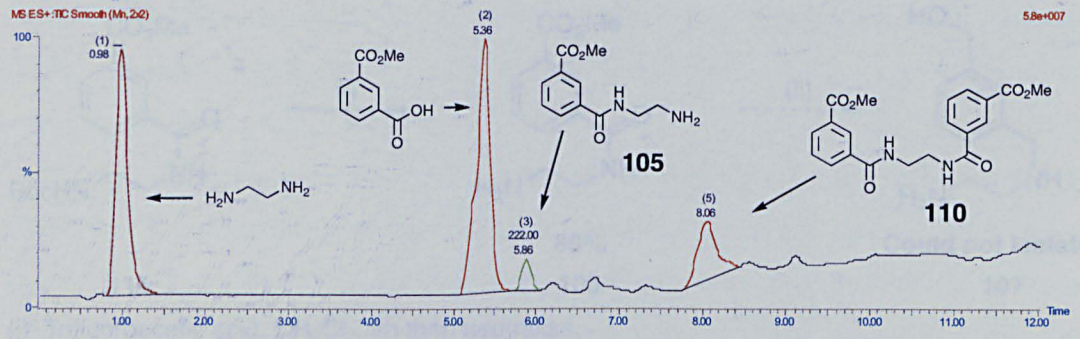
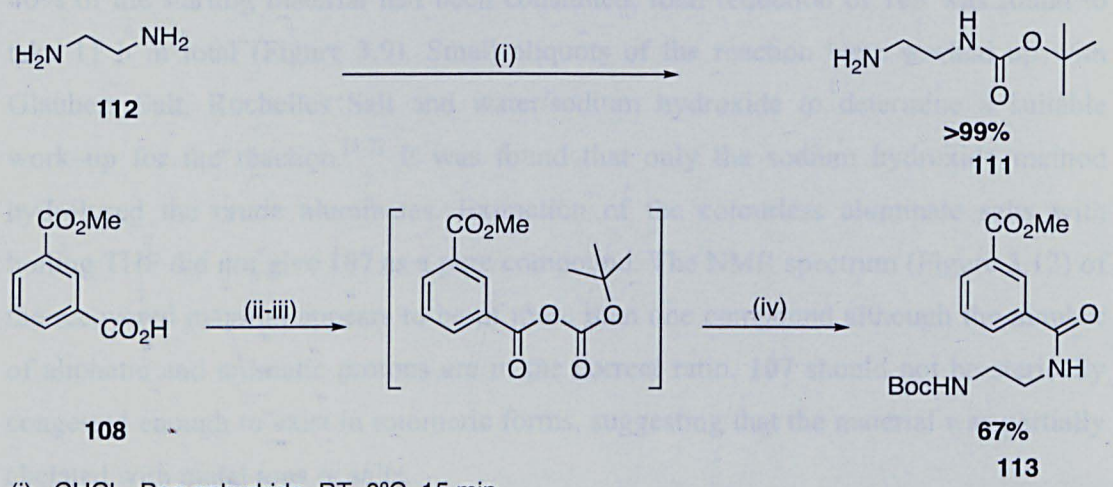


Figure 3.11 Attempted monoacylation of EDA with thionyl chloride-activated **108**

Ultimately **105** could not be separated from the excess ethylene diamine present in the reaction.

It was decided to monoprotect EDA to overcome the diacylation problem.

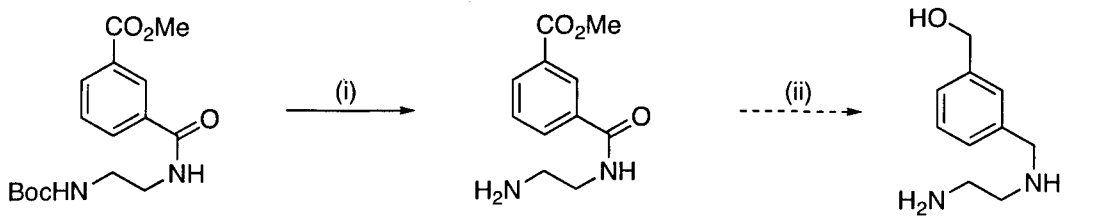
The procedure of Eisenführ was followed to make monoamine **111** and that was acylated using the mixed anhydride methodology to give **112** (Scheme 3.10)



(i) $CHCl_3$, Boc anhydride, RT, 0°C, 15 min
(ii) MeCN, NMM, RT, 5 min
(iii) Pivaloyl chloride, -15°C, 30 min
(iv) **111**, RT, 25 min

Scheme 3.10 Synthesis of Boc protected intermediate **113**

With **113** in hand, the Boc deprotection and global reduction was attempted (Scheme 3.11).



86%

Could not isolate

113 **105** **107**

(i) Trifluoroacetic acid, CH_2Cl_2, 2h then neutralise
(ii) THF, Lithium Aluminium Hydride (6 eq.), reflux, 11h

Scheme 3.11. Attempted total reduction of amine **105**

Deprotection of **113** was essentially quantitative however it was not possible to completely extract it from the neutralised aqueous reaction.

Reduction of **105** was monitored by LC-MS. The reaction is sluggish and after 1h only 40% of the starting material had been consumed; total reduction of **105** was found to take 11 h in total (Figure 3.9). Small aliquots of the reaction were worked up with Glaubers Salt, Rochelles Salt and water/sodium hydroxide to determine a suitable work–up for the reaction.[5-7] It was found that only the sodium hydroxide method hydrolysed the crude aluminates. Extraction of the colourless aluminate salts with boiling THF did not give **107** as a pure compound. The NMR spectrum (Figure 3.12) of the recovered material appears to be of more than one compound although the number of aliphatic and aromatic protons are in the correct ratio. **107** should not be sterically congested enough to exist in rotomeric forms, suggesting that the material was partially chelated with metal ions or salts.

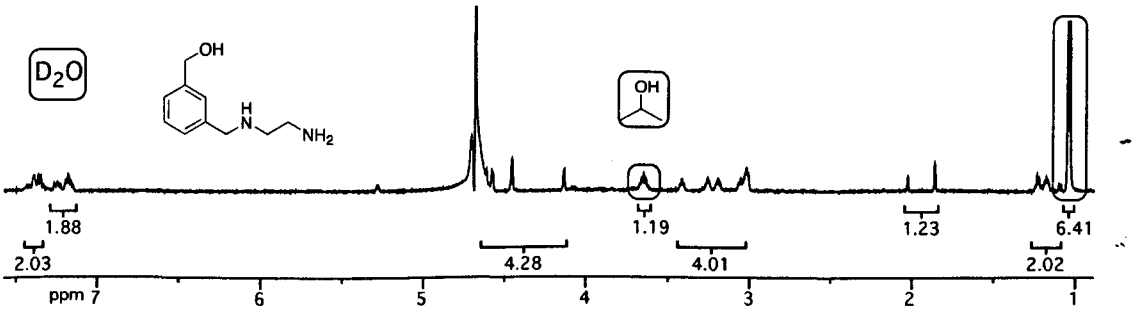


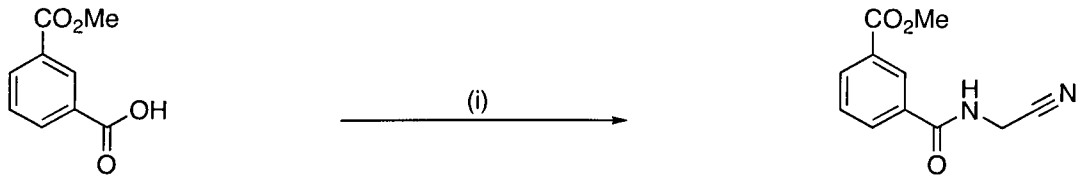
Figure 3.12 Lithium aluminium hydride reduction of 105

105 is potentially an excellent metal cation ligand. The recovery of **107** from the metal salts (generated post-reduction) is made difficult probably due to this multi-dentate

capacity. Attempts to sequester the aluminium using alternative work-up conditions (triethanolamine) failed.

There was concern that the free amino group of **105** was affecting the reduction in some adverse way. It was predicted that a reduction in which the amine was in a masked form *before* the reduction and unmasked during workup of the reaction would circumvent this problem. It was thought that this could be accomplished if the oxidation level of the amine was raised to that of a nitrile.; the amine would only be liberated during post-reductive reaction work-up and would therefore not be available to interfere with the course of reaction.

Nitrile **114** was prepared from isophthalate (Scheme **3.12**).



108 72%

114

(i) NEt_3, HOBt, DMF, RT, 5 min
(ii) EDC, RT, 1h
(iii) Aminoacetonitrile, RT, 16 h

Scheme 3.12 Synthesis of nitrile reduction precursor **114**

The use of alternative activating agent DCC gave a yield of 48% when used under otherwise identical reaction conditions.

Initially a partial reduction of **114** was attempted using a solution of borane dimethylsulphide complex. The borane was anticipated to reduce the nitrile and amide functionality, leaving the ester functionality intact as an extraction/purification handle. In practice the intermediate **115** could not be isolated as a pure compound and the reaction yielded an inseparable mixture. The reaction was heated at reflux for 3 days in aqueous acetic acid, but this failed to generate a major or isolable compound.

Giannis[8] reported that TMSCl and Lithium borohydride are combined to give a potent reducing cocktail that is suitable for substrates that struggle to fully reduce with

standard THF borane solutions. [8] **114** was treated with a solution of the reduction cocktail at RT for 2h, as according to the procedure of Giannis. A mixture of the starting material **114** and compounds **116** and **117** were observed in addition to **115** within the reaction mixture (Figure 3.13).

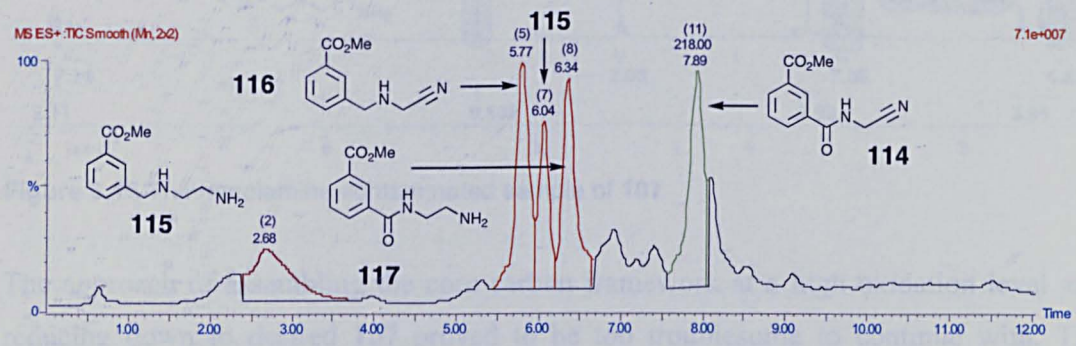


Figure 3.13 LC-MS trace of attempted partial reduction of **114** using $LiBH_4$

After a further 3h the LC-MS profile of the crude reaction mixture had not changed and so the reaction was heated at reflux for 24 h in an effort to complete the reduction. After this 24 h period the starting material was still observed in the reaction mixture however the number of side products had increased (Figure 3.14).

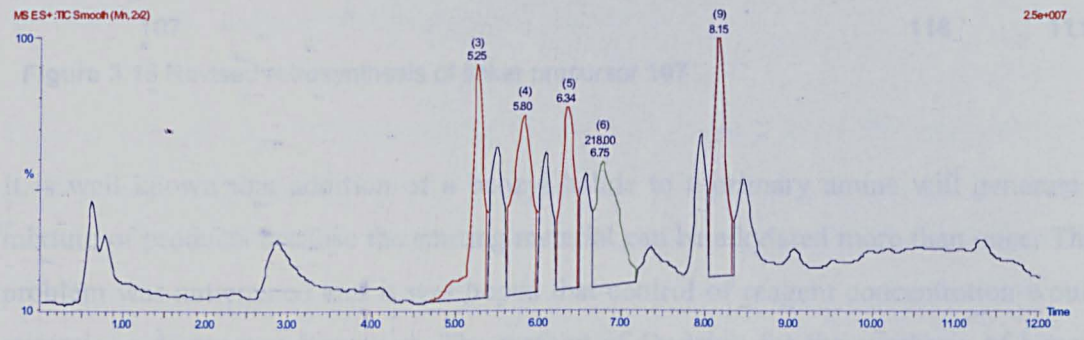


Figure 3.14 Extended reaction time in attempt for partial reduction of **114** by $LiBH_4$

It was decided to forgo the partial reduction approach and attempt total global reduction instead. **114** was treated with lithium aluminium hydride in THF at reflux and worked up with aqueous sodium hydroxide. Extraction of the crude metal salts with refluxing THF failed to give any organic extract. The reaction was repeated and worked up with triethanolamine over 16 h to give a colourless gel of metal salts that were filtered and

washed repeatedly with CH_2Cl_2. Evaporation of the combined crude organics gave an inseparable mixture of product **107** and triethanolamine (Figure 3.15).

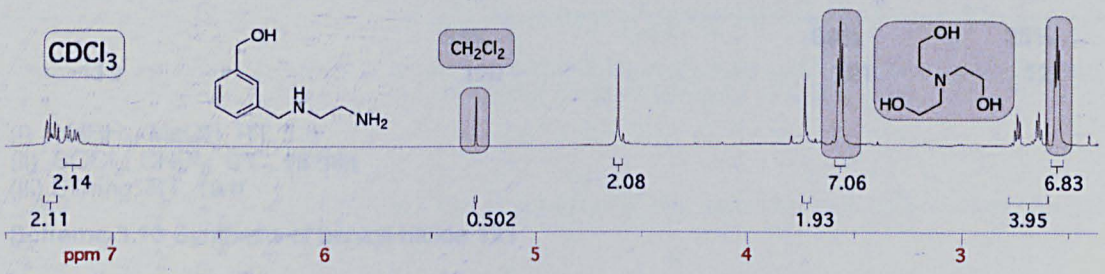
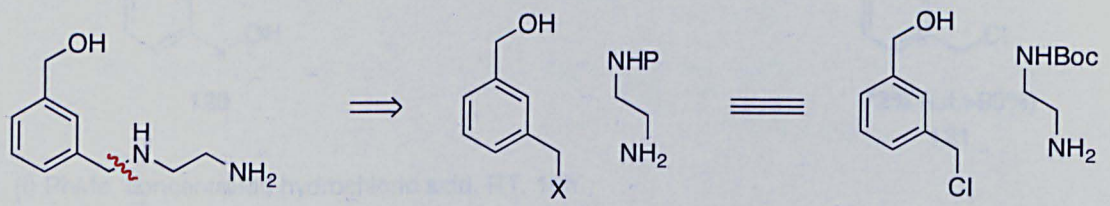


Figure 3.15 Triethanolamine contaminated sample of **107**

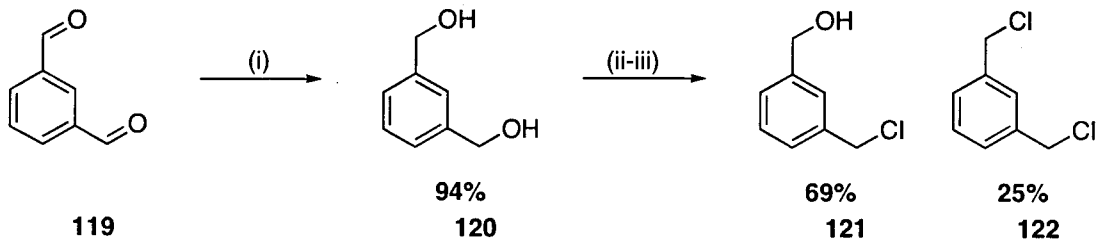
The approach of assembling the core carbon framework at a high oxidation level and reducing down to desired **107** proved to be too troublesome to continue with. The obvious change to the strategy was to assemble **107** using fragments already at the correct oxidation state. The revised approach to **107** is shown in Figure 3.16



107 **118** **111**

Figure 3.16 Revised retrosynthesis of linker precursor **107**

It is well known that addition of a benzyl halide to a primary amine will generate a mixture of products because the starting material can be alkylated more than once. This problem was anticipated and it was hoped that control of reagent concentration would minimise side-product formation. The method of Dodziuk for the synthesis of benzyl halide **121** was followed to effect the conversion of **120** to **121** (Scheme 3.13); diol **120** was prepared from commercially available **119** under standard conditions. [9]



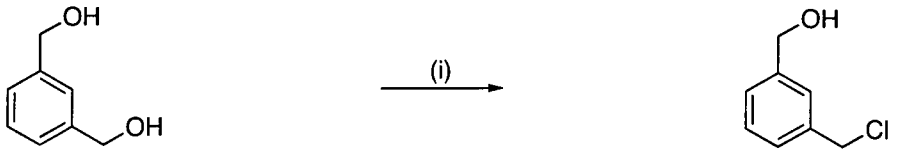
119 **94%** **120** **69%** **121** **25%** **122**

(i) $NaBH_4$, MeOH, RT, 1 h
(ii) $SOCl_2$, $CHCl_3$, 0°C, 15 min
(iii) Stirring, RT, 18 h

Scheme 3.13 Synthesis of benzyl halide **121**

121 was isolated after difficult chromatography.

A slightly modified method of Naghipour was used to prepare **121** as a single product without chromatography (Scheme 3.14).

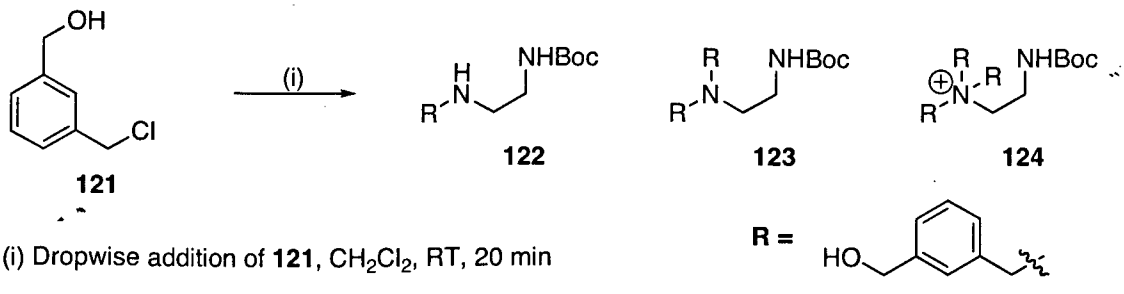


120 **72%** (Lit.>99%) **121**

(i) PhMe, concentrated hydrochloric acid, RT, 18h

Scheme 3.14 Alternative synthesis of benzyl halide **121**

For the approach to **122** depicted in Figure 3.17 to be useful, the monoalkylation product must effectively be the only polar product; it was known from prior experience that the amine **111** and product **122** were inseparable using silica chromatography. Careful monitoring and adjustment of the reaction conditions were used in an attempt to obtain **122** as the only product generated by addition of **121** to neat **111**(Table 3.1).



121 **122** **123** **124**

(i) Dropwise addition of **121**, CH_2Cl_2, RT, 20 min

Figure 3.17 Experiment for synthesis of intermediate **122**

[BenzylCl]	111	120	122	123	124
0.500M	0	0	1	4	5
0.250M	0	0	1	1	0
0.125M	200	200	1	1	0

Table 3.1 Approximate ratios of crude reaction components, relative to monoalkylation product **122**, for reactions run under different concentrations of benzyl chloride **121**

Even at (relatively) high dilution, the formation of the side product **123** is unavoidable (Figure 3.18).

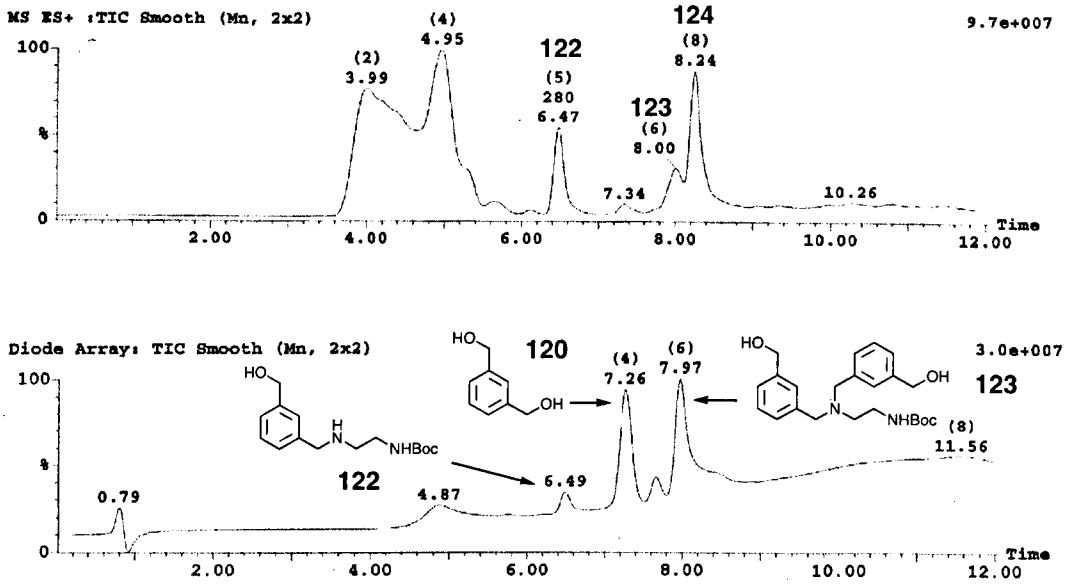


Figure 3.18 LC-MS traces for attempted monoacylation of Boc EDA with 0.125M **121**

The starting materials were clearly seen to be the dominant peaks in the mass-ion trace of the 0.125M reaction, and small equimolar amounts of the product and **123** (in addition to **124**) are also visible. The UV trace revealed that the actual amounts of **122** & **124** formed are negligible and that the two amines are present on the mass-ion trace because they easily ionise, not because they are present in any meaningful quantities. The data strongly suggests that with a 1:1 stoichiometry of starting materials that the reaction cannot be run under concentration control to generate solely the desired product. The enhanced nucleophilicity of the product relative to the starting material overcomes the high dilution effects and dialkylation side product **123** cannot be avoided.

A final attempt to directly synthesise **107** was attempted by slow addition of neat **121** to neat ethylene diamine . Without the carbamate protecting group both of the EDA amino groups potentially could have been acylated by **121** additional side products. However it was hoped that at low temperature, with EDA in excess to **121**, that only **107** could form; As **107** was formed its concentration would remain insignificant to that of EDA and therefore it would have been statistically very unlikely to encounter a second equivalent of **121** and react further to give byproducts.

It was found that adding neat **121** by syringe pump (1.0 mL per hour) to a tenfold molar excess of neat EDA at low temperature (<0°C) gave exclusively **107**. Separation of **107** from the excess EDA was difficult but was accomplished with 5 h of reduced-pressure evaporation at 60°C (Figure 3.19).

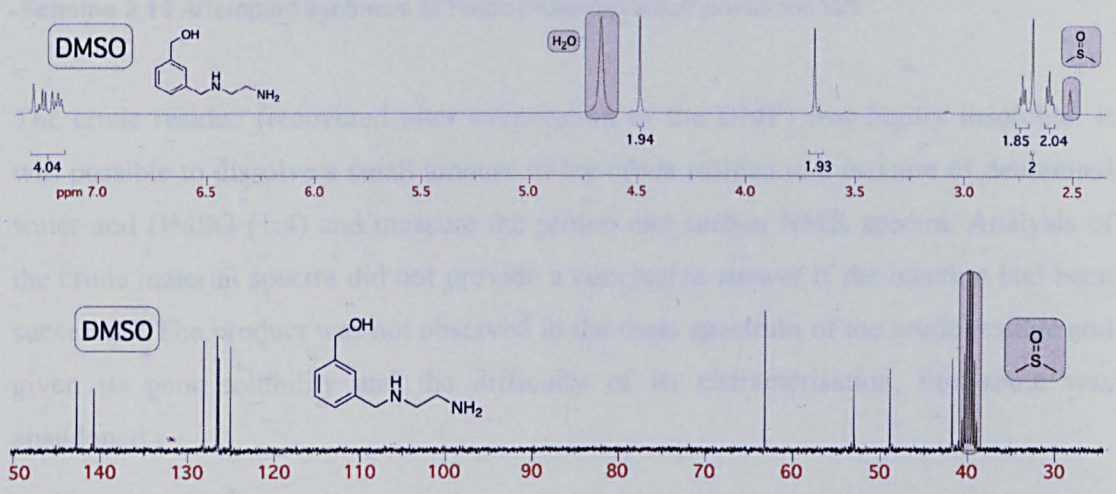
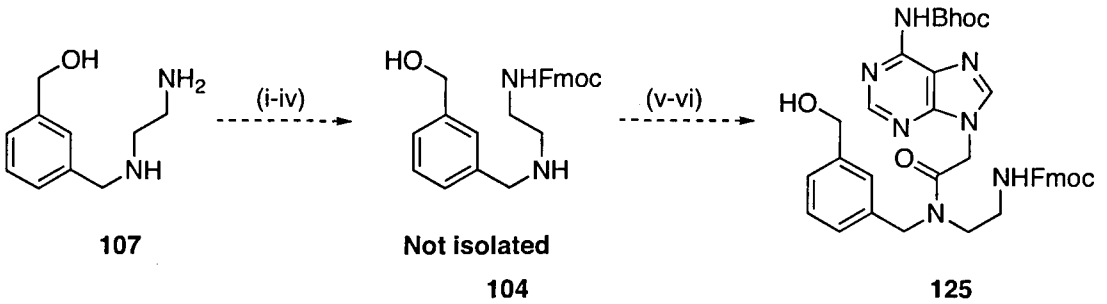


Figure 3.19 NMR data for intermediate **122**

To make use of **107** it must be possible to discriminate between the primary and secondary amine. Synthesis of linker **97** was envisioned to be completed by the selective protection of the primary amine of **107** with an Fmoc group followed by acylation with adenyl acetic acid **75** (Scheme 3.15). Initially it was hoped that the chemistry used to synthesise backbone **89** could be applied to the synthesis of **104**. There was concern however that **104** had the potential to deprotect itself, that the Fmoc carbamate could be destroyed by attack of the free secondary amine. If auto-degradation was occurring had to be determined empirically. The reaction was carried out at pH 7.0 (unbuffered) and dibenzofulvene, a distinctive byproduct of deprotection, was not

detected (by mass spectrographic analysis of the crude reaction mixture) during the reaction.



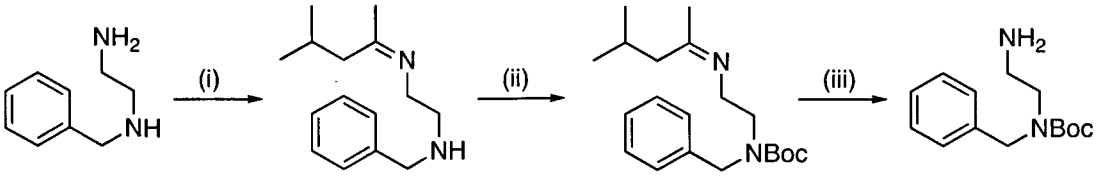
107 **Not isolated** **104** **125**

(i) Bistrimethylsilylacetamide, DMF, 1 h, RT
(ii) Fmocsuccinimide, DMF, 1 h, RT
(iii) MeOH then H_2O, 10min, ice bath
(iv) Evaporate to dryness, triturate with CH_2Cl_2
(v) Bhoc-Adenyl acetic acid, N-methylmorpholine, MeCN, RT, 5 min
(vi) Pivaloyl chloride, 0°C, 20 min then **104**, 1 h, RT

Scheme 3.15 Attempted synthesis of Fmoc protected linker precursor **125**

The crude residue (recovered after evaporation of the DMF) was highly insoluble. It was possible to dissolve a small amount of the crude residue in a mixture of deuterated water and DMSO (1:4) and measure the proton and carbon NMR spectra. Analysis of the crude material spectra did not provide a conclusive answer if the reaction had been successful. The product was not observed in the mass spectrum of the crude residue and given its poor solubility and the difficulty of its characterisation, this route was abandoned.

Laduron and coworkers reported that a primary amine could be selectively condensed with methyl isobutylketone (MIBK) to give the corresponding imine in the presence of a primary amine (Scheme 3.16).[10]



126

(i) MIBK, reflux, 2 h
(ii) 0°C, Boc anhydride, 30 min
(iii) Evaporate to dryness then water/isopropanol (1:1), 50°C, 1 h

Scheme 3.16 Laduron selective acylation of secondary amine in presence of primary amine

107 and benzyl EDA **126** are very similar in structure and it was thought that the conditions given in Scheme 3.16 could be used to convert **107** into **127** and then into either **128 or 129.** Either of these compounds could then be converted to **125** (Figure 3.20).

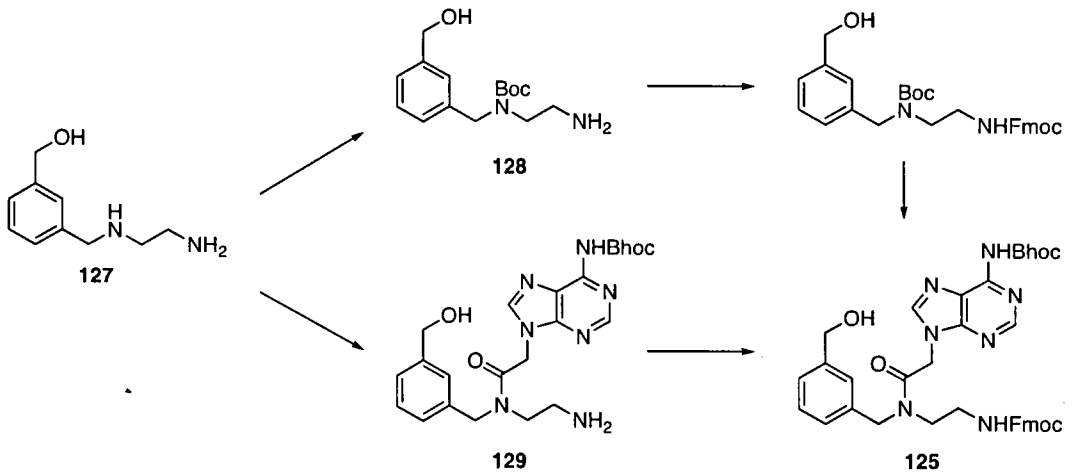


Figure 3.20 Approaches to **125** using the Laduron methodology

The imine **130** was successfully formed under the conditions described by Laduron (Figure 3.21).

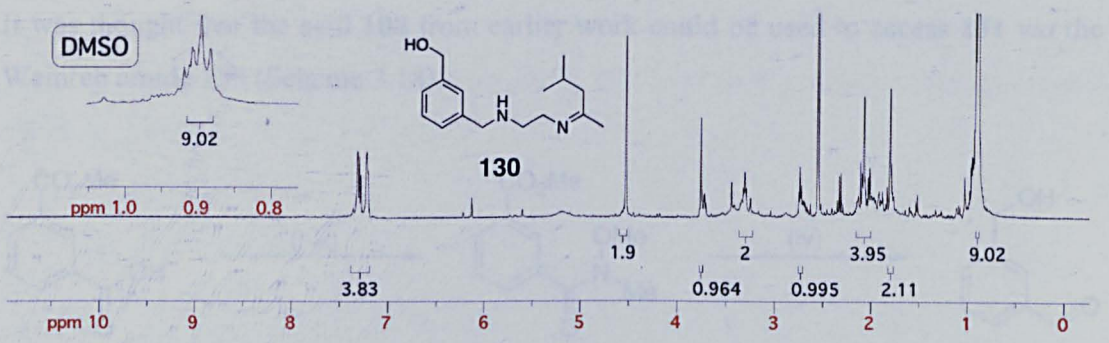
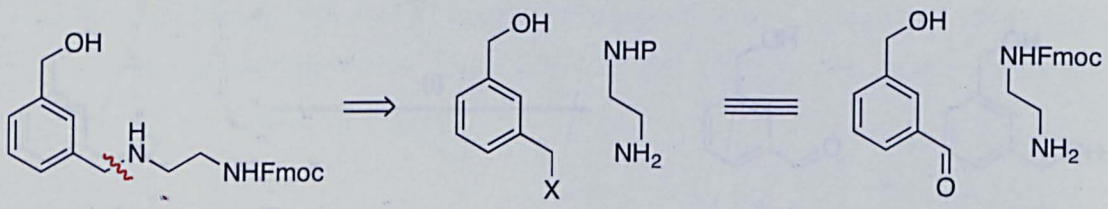


Figure 3.21 MIBK imine of **107**

Addition of the boc anhydride caused the reaction to become a dark brown colour, and after hydrolysis of the imine there was no product (**127**) present. NMR analysis of crude isolated material showed that the starting material had been transformed into something else that could not be identified. The reaction was repeated and identical results were observed. It is not clear why this reaction failed in light of the reported results depicted in Scheme 3.16 .

Another approach to the intermediate **104** was required. It was decided that **104** could be made from reductive amination between **131** and **132**.

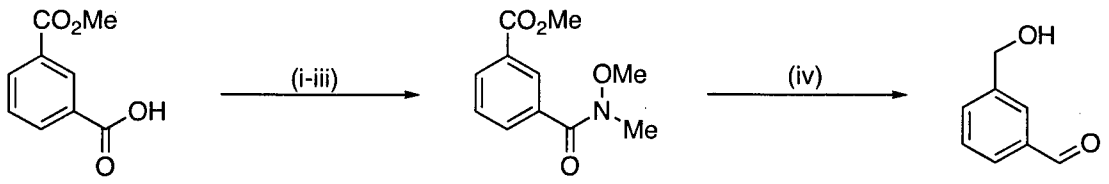


104 **131** **132**

Scheme 3.17 Revised retrosynthesis of linker precursor **104**

This approach had been previously considered but was not pursued for a number of good reasons. There is a known difficulty with reducing aryl imines – Schiff base complexes (aryl imines) are used as ligands for reduction catalysts, highlighting that such imines are stable to reducing conditions and as such, harsh conditions are required to reduce them. Such conditions could be incompatible with the Fmoc group. Previous experience with purifying **107** from metal reductant and waste metal salts proved to be troublesome and the potential for inseparable side products (forming *via* reductive amination of the product) could also occur and render the reaction inefficient.

It was thought that the acid **108** from earlier work could be used to access **131** *via* the Weinreb amide **133** (Scheme 3.18).

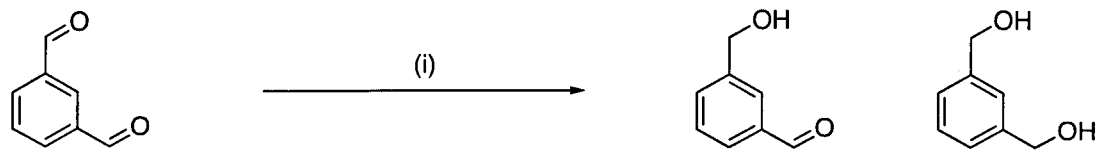


108 **88%** **133** **94%** **131**

(i) NMM, THF, RT, 5 min
(ii) Isobutylchloroformate, -20°C, 10 min
(iii) NEt_3 then *N,O*-dimethylhydroxylamine, DMF, -20°C, 30 min then RT, 30 min
(iv) Lithium aluminium hydride, THF, 0°C, 1 h

Scheme 3.18 Synthesis of aldehyde **131**

A more direct route of synthesis was also developed. The partial reduction of isophthalaldehyde **119** was attempted to access **131** in one step. The procedure of Takemoto *et al.* was followed and **131** was isolated in slightly greater yield than that expected for a statistical reduction. [11]

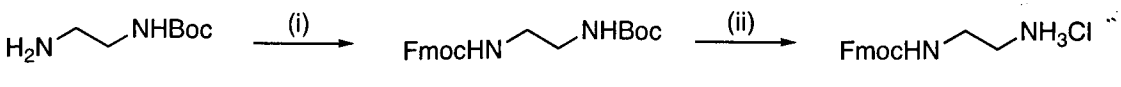


119 **58%** **131** **31%** **120**

(i) 0.34 equiv. $NaBH_4$, EtOH, 0°C, 1 h

Scheme 3.19 Alternative synthesis of aldehyde **131**

Amine **132** was prepared as according to the procedure of Boeijen (Scheme 3.20).[12]

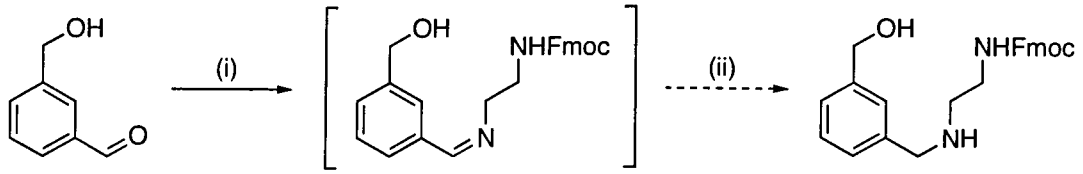


111 **89%** **134** **~99%** **132**

(i) H_2O, MeCN, Fmocsuccinimide, pH 9.0, 10 min
(ii) MeCN, 6M HCl, 18 h

Scheme 3.20 Synthesis of FmocEDA hydrochloride

With both the aldehyde and amine in hand, the reductive amination was attempted (Scheme 3.21).



131 **135** **Not observed 105**

(i) **132** 0.5 equiv., DIPEA, MeOH, RT, 1 h
(ii) $NaBH_4$ 1 equiv., RT, 1 h

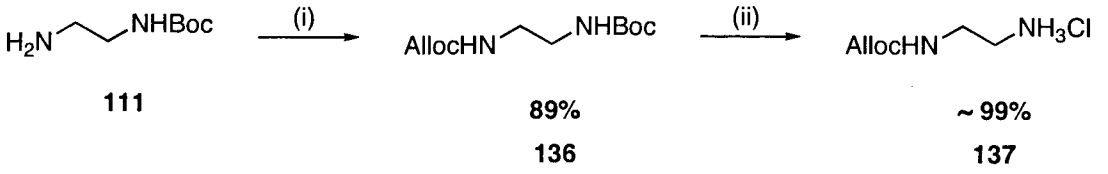
Scheme 3.21 Attempted synthesis of intermediate **105** via reductive amination

Salt **132** was suspended in stirred methanol and diisoproylamine was added until the reaction pH reached 7.0, wherein **132** was assumed to be the free base. An excess of aldehyde **131** was added to the reaction and allowed to stir for 1 h to complete imine formation. The reaction was placed in an ice bath and sodium borohydride was added to the suspension and stirred for 30 min before the reaction was quenched and worked up. The reaction returned only diol **120** and a baseline mixture of amines and borate salts. The results suggest that either imine **135** did not form or was not reduced and subsequently hydrolysed on workup.

The difficulty associated with introducing the Fmoc group into **97** was now thought to not be worth the convenience of its use as a familiar protecting group. The allyloxycarbonyl (Alloc) protected equivalent of **97** was thought to be a suitable analogue. The Alloc group is established in solid phase synthesis as a more chemically robust alternative to the Fmoc group for amine protection. It is removed under neutral conditions ($Pd(PPh_3)_4$, cation scavenger (eg $PhSiH_3$)) that are orthogonal and complementary to the acidic cleavage conditions used for the cleavage of peptide and PNA sidechain protecting groups. It was anticipated that **137** would be a superior choice of partner for **131** in imine formation for more than one reason. It should withstand harsher, more forcing reducing conditions that could be required to successfully reducing the aryl imine. It should not possess the solubility issues that accompany, and are peculiar to, the Fmoc group with this class of molecule. Furthermore, it will be

possible to accurately gauge imine formation by ^{1}H NMR without the Fmoc aromatic signals obscuring changes in chemical shift of the aryl protons of **131** or the presence of the imine proton.

137 was prepared in an analogous manner to FmocEDA **132** (Scheme 3.22):



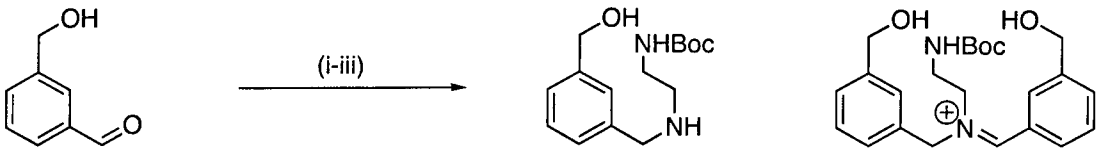
111 **89%** **~ 99%**

136 **137**

(i) NEt_3, CH_2Cl_2, RT, 5 min
(ii) Allylchloroformate (dropwise), CH_2Cl_2, 0°C, 30 min then RT, 24 h
(iii) Acetyl chloride 3 equiv,, MeOH, 0°C, 30 min then **136**, RT, 3 h

Scheme 3.22 Synthesis of Alloc EDA **137**

Concurrent to the synthesis of Alloc EDA **137**, it was decided to test reductive amination conditions using sodium triacetoxyborohydride and aldehyde **131** with the readily available Boc EDA **111** (Scheme 3.23).[13]



131 **Minor (<5%)** **Major (>95%)**

122 **122b**

(i) Boc EDA 1.0 equiv., RT, 20 min
(ii) AcOH (2% total volume)
(i) Sodium triacetoxyborohydride 2.0 equiv., RT, 3 h

Scheme 3.23 Reductive amination of **131** using Nagalingam conditions

The product (**122**) was found to be very reactive and could not be isolated (Figure 3.22).

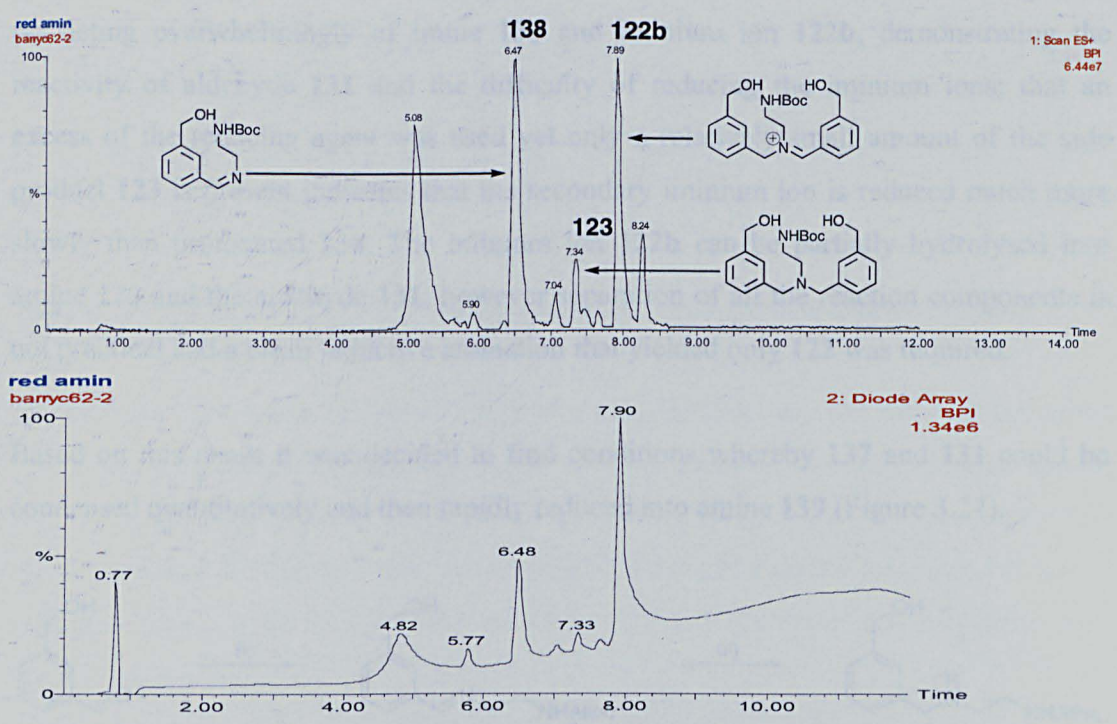


Figure 3.22 LC-MS traces of crude reductive amination mixture

Under the mildly acidic conditions, formation of imine **138** must be readily reversible, and the untethered product **122** outcompetes Boc EDA in the formation of an iminium ion with free aldehyde **131** (Figure 3.20).

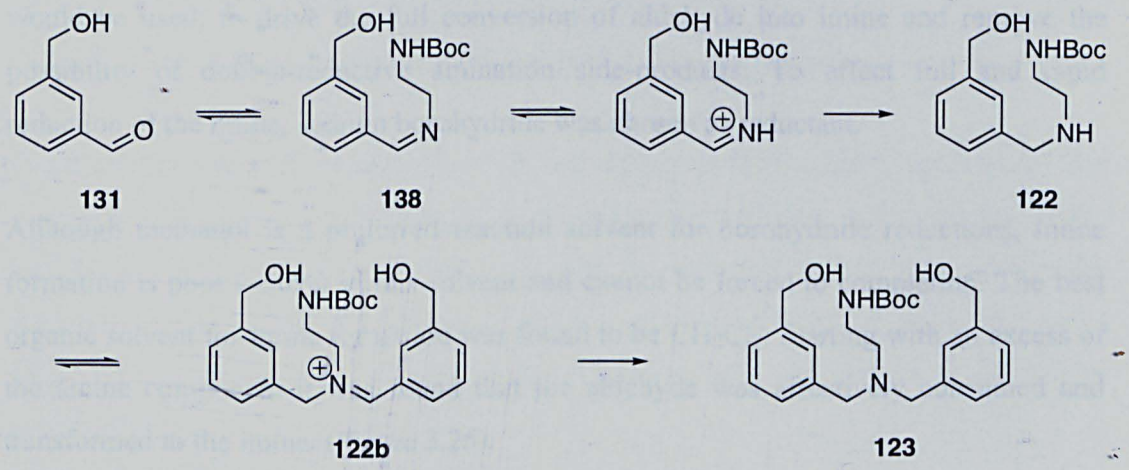
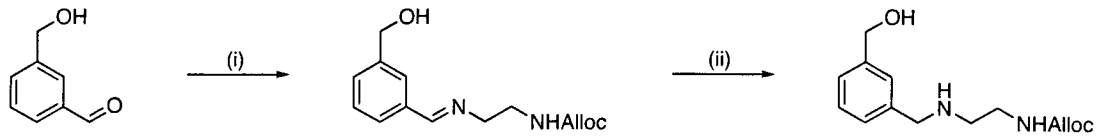


Figure 3.23 Formation of single and double reductive amination products

Product **122** is a much more competent nucleophile than Boc EDA, due to the +I effect from the aryl ring that raises the energy of the nitrogen lone pair of electrons, and readily sequesters aldehyde **131.** The overall observed reaction is therefore a mixture

consisting overwhelmingly of imine **138** and iminium ion **122b**, demonstrating the reactivity of aldehyde **131** and the difficulty of reducing the iminium ions; that an excess of the reducing agent was used yet only a relatively small amount of the side product **123** is present indicates that the secondary iminium ion is reduced much more slowly than protonated **138**. The iminium ion **122b** can be partially hydrolysed into amine **122** and the aldehyde **131**, however separation of all the reaction components is not practical and a clean reductive amination that yielded only **122** was required.

Based on this result it was decided to find conditions whereby **137** and **131** could be condensed quantitatively and then rapidly reduced into amine **139** (Figure 3.24).



131 **138** **139**

(i) Quantitative condensation
(ii) Rapid reduction to prevent side-product formation under reductive conditions

Figure 3.24 Approach to key amine **139**

To establish conditions for the reduction of **138**, imine formation and reduction was followed with ^{1}H NMR and LC-MS. It was decided that initially an excess of amine would be used, to drive the full conversion of aldehyde into imine and remove the possibility of double-reductive amination side-products. To effect full and rapid reduction of the imine, sodium borohydride was chosen as reductant.

Although methanol is a preferred reaction solvent for borohydride reductions, imine formation is poor (~50%) in this solvent and cannot be forced to completion. The best organic solvent for imine formation was found to be CH_2Cl_2. Starting with an excess of the amine component it was found that the aldehyde was effectively consumed and transformed to the imine (Figure 3.25).

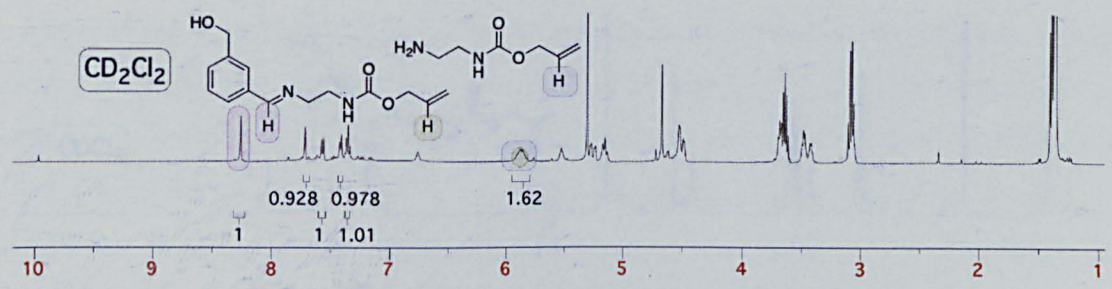


Figure 3.25 ^{1}H NMR spectrum of imine formation with an excess of the amine component

To estimate the excess of free amine present in the reaction, the ratio of the imine proton (8.30 ppm) and allyl protons in the ^{1}H NMR spectrum was calculated. It should be noted that the percentage of imine formation is identical in both deuterated chloroform and dichloromethane. An amount of aldehyde equal to the excess of amine was then added to the reaction and the crude ^{1}H NMR spectrum was measured once more (Figure 3.26)

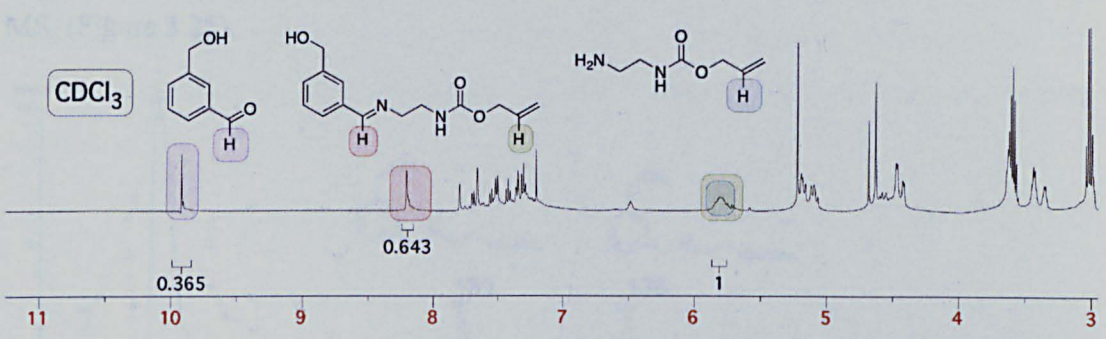


Figure 3.26 ^{1}H NMR spectrum of incomplete imine formation from equimolar reactants

Prolonged stirring and the addition of drying agents such as magnesium sulphate and 3 Å molecular sieves did not change the amount of imine present in the reaction. Imine formation was known to be sensitive to pH and an amount of DIPEA was added to the reaction to raise the pH to approximately 9.0 in the hope of driving imine formation. This was successful and imine formation was essentially quantitative at pH 9.0 or greater (Figure 3.27).

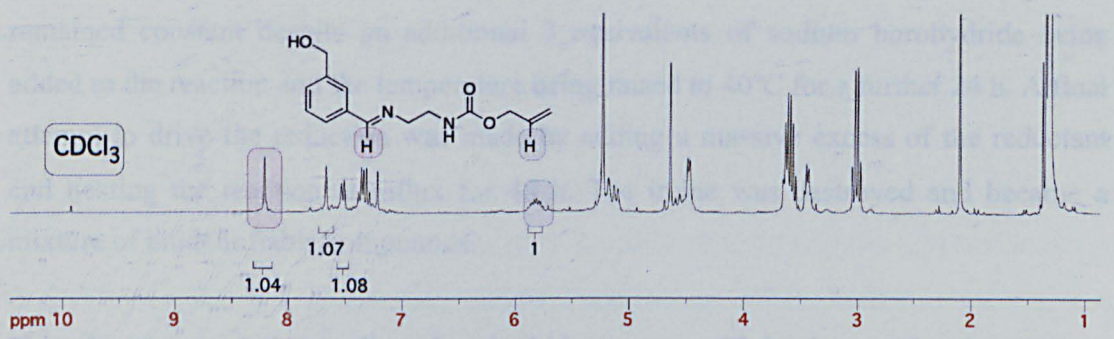


Figure 3.27 ^{1}H NMR spectrum of completed imine formation with equimolar reactants at high pH

An aliquot of the reaction was treated with solid sodium borohydride and no reduction was observed after 1 h of stirring. This was not an unexpected result given the insolubility of sodium borohydride in CH_2Cl_2. To effect reduction, the reaction was evaporated to dryness and a solution of sodium borohydride (2 equivalents)dissolved in THF was added. After 3 h of stirring, an aliquot from the reaction was analysed by LC-MS. (Figure 3.25).

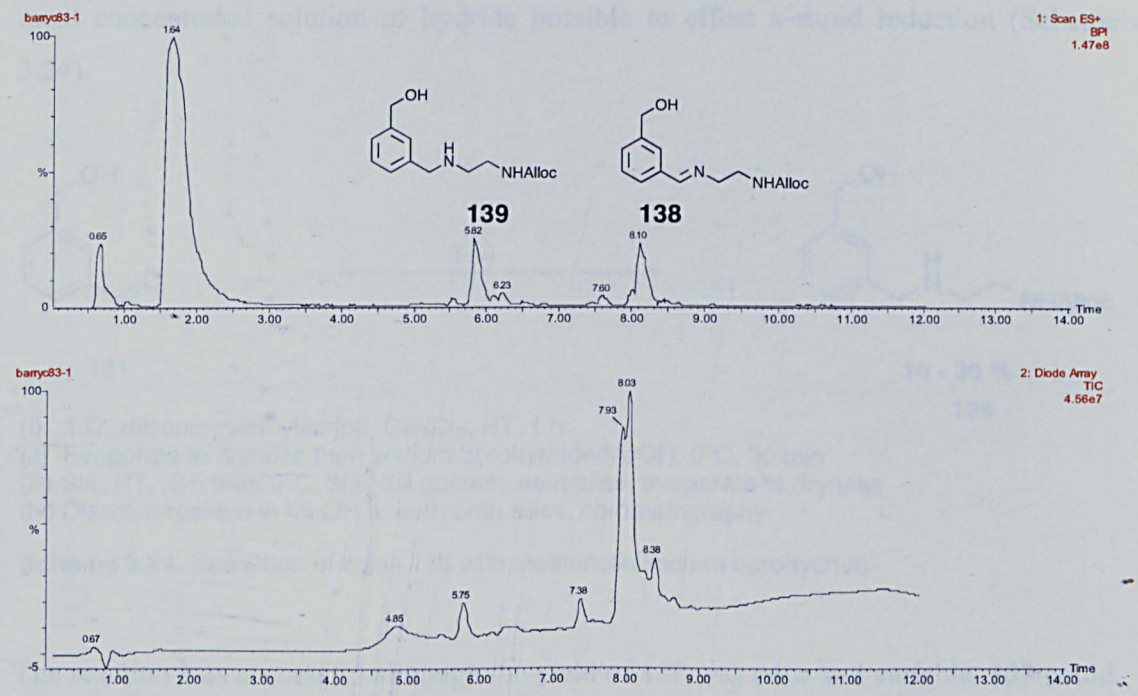


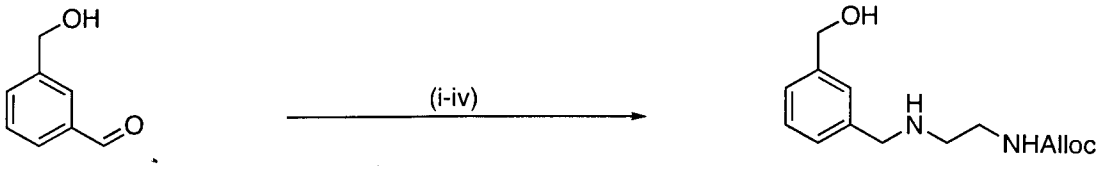
Figure 3.25. LC-MS traces of incomplete imine reduction with sodium borohydride

While a small amount of **139** was present in the mass spectrum trace, the UV trace revealed that the product was almost exclusively the imine **138**. The reaction was allowed to stir at RT for 12 h but no change was observed. The reaction LC-MS trace

remained constant despite an additional 3 equivalents of sodium borohydride being added to the reaction and the temperature being raised to 40°C for a further 24 h. A final attempt to drive the reduction was made by adding a massive excess of the reductant and heating the reaction at reflux for 48 h. The imine was destroyed and became a mixture of unidentifiable compounds.

This demonstrated that sodium borohydride is not sufficiently reactive in THF or CH_2Cl_2 to be able to reduce the imine and that the imine is unstable under harsh reductive conditions.

Reduction of imine **138** (dissolved in CH_2Cl_2) was revisited with the idea of using a solution of sodium borohydride dissolved in a minimum amount of methanol. It was hoped that the miscibility of methanol with CH_2Cl_2 would allow reactive hydride to be delivered to the imine and thus facilitate reduction. There was concern about the methanol disrupting the imine/amine equilibrium. It was hoped that using the least volume of methanol possible would minimise any such disruption, while providing the most concentrated solution of hydride possible to effect a rapid reduction (Scheme 3.24).



131 **10 - 30 %** **139**

(i) **137**, diisopropylethylamine, CH_2Cl_2, RT, 1 h
(ii) Evaporate to dryness then sodium borohydride/MeOH, 0°C, 30 min
(iii) Stir, RT, 3 h then 0°C, 3M HCl quench, neutralise, evaporate to dryness
(iv) Dissolve residue in MeOH adsorb onto silica, chromatography

Scheme 3.24. Reduction of imine **138** with methanolic sodium borohydride

The reaction was successful although the yield of **139** was poor and variable. **139** could not be extracted from aqueous solution and required multiple bouts of chromatography (methanol/ethyl acetate, 3:7) to separate it from the borate/diisopropylethylamine salts and traces of Alloc amine **137**. The yield was maximised by allowing the sodium borohydride to stir for 10 – 15 min once fully dissolved in the methanol. This presumably resulted in the imine being reduced by a methoxyborohydride deriviative,

such as $BH(OMe)_3$, that was almost certainly more miscible with the imine than sodium borohydride.

The poor yield was caused largely with the difficulty in separating **139** from the diisopropylethylamine used in the reaction. It was desired to remove nitrogenous base from the reaction to improve the reaction yield and expediate access to **139**. It was thought that the imine formation could be achieved in aqueous conditions with an inorganic base instead.

The aldehyde and amine were stirred in saturated sodium hydrogen carbonate solution water for 3 h to give an orange oil floating on a colourless solution. A small portion of the oil was dissolved in aqueous acetonitrile and analysed by LC-MS (Figure 3.28)

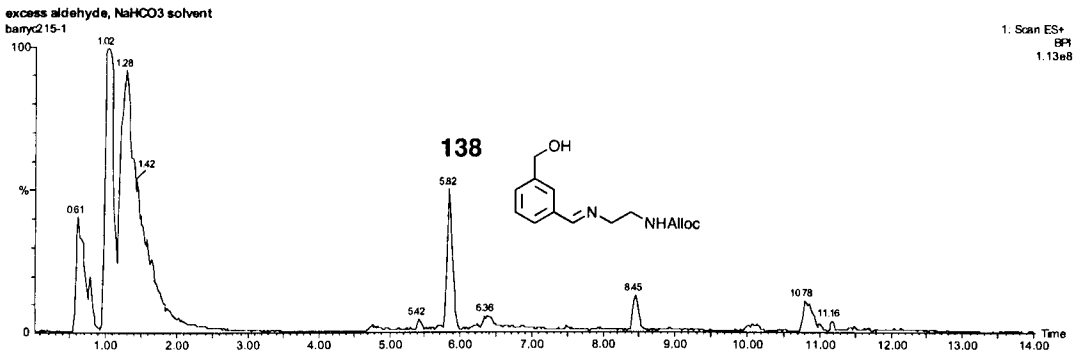


Figure 3.28 Imine formed in saturated sodium hydrogen carbonate solution

The reaction was concentrated and an excess of sodium borohydride (dissolved in water) was added and the imine was rapidly reduced to give **139** as the major component of the crude reaction solution (Figure 3.29).

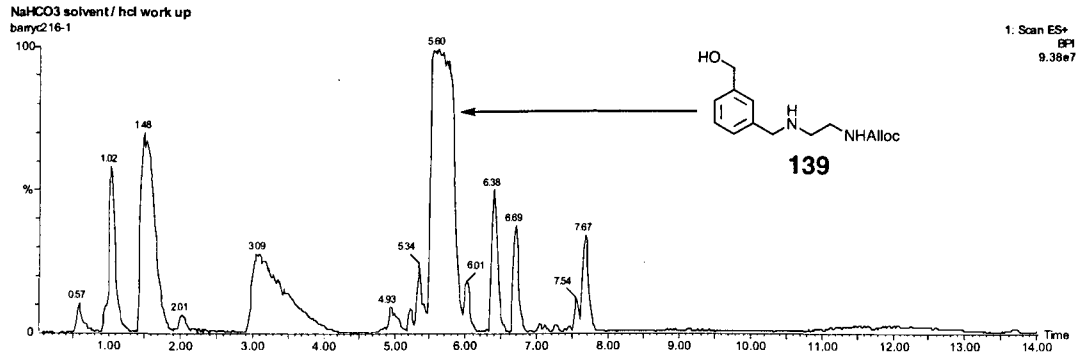


Figure 3.29 LC-MS trace of crude reduction mixture

Initially chromatography was used to recover the product to give an average yield of 55 – 60%. However a workup procedure was later developed that recovered the crude product much more cleanly (Figure 3.30) and in up to 90% yield.

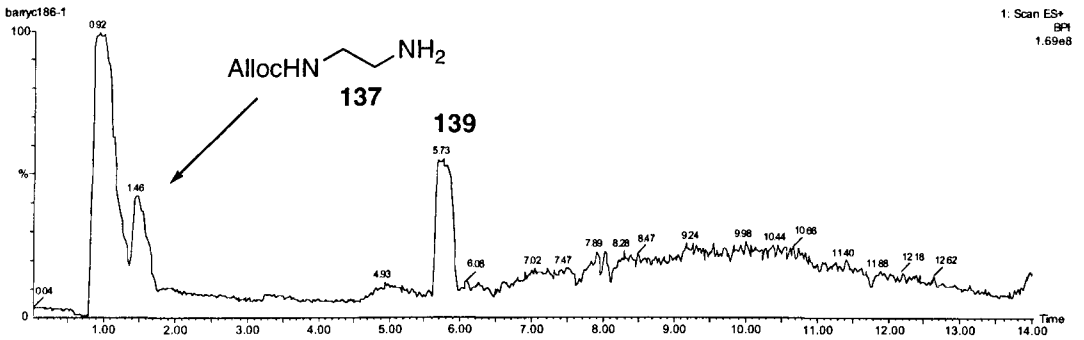


Figure 3.30 LC-MS trace of amine **139** recovered by new workup procedure

The crude mixture of amines were separated using ion exchange on a prepacked SCX column eluted first with methanol and then ammonia saturated methanol to give **139** as a single compound in 90% overall yield (Figure 3.31).

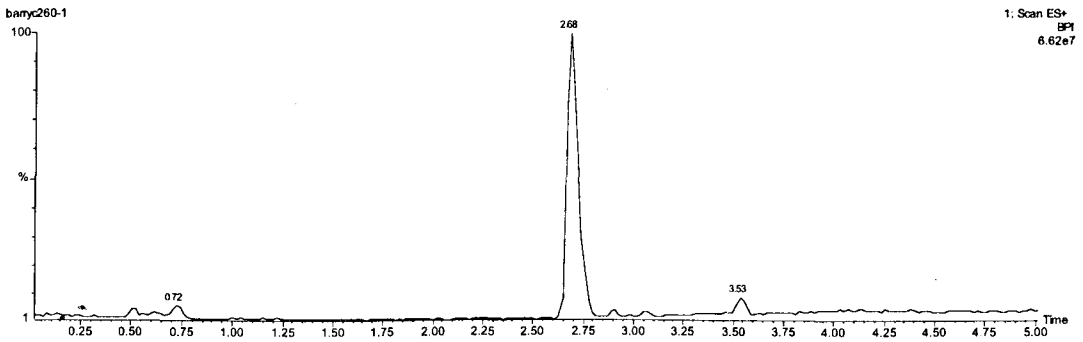
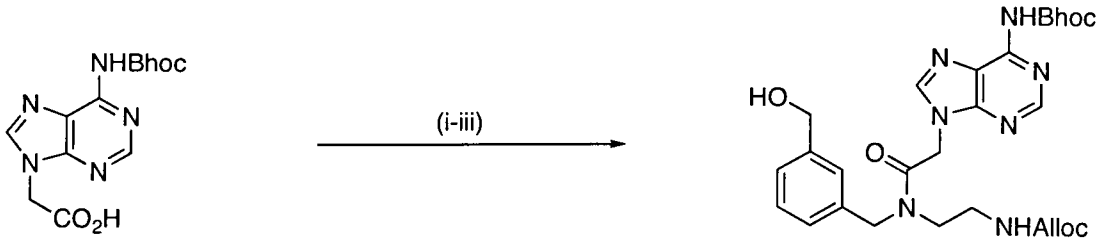


Figure 3.31 LC-MS trace of amine **139** purified *via* prepacked SCX column

Intermediate **140**, the penultimate compound in the synthesis of the linker, was seen to be accessible *via* a mixed anhydride coupling as per PNA monomer synthesis (Scheme 3.26).

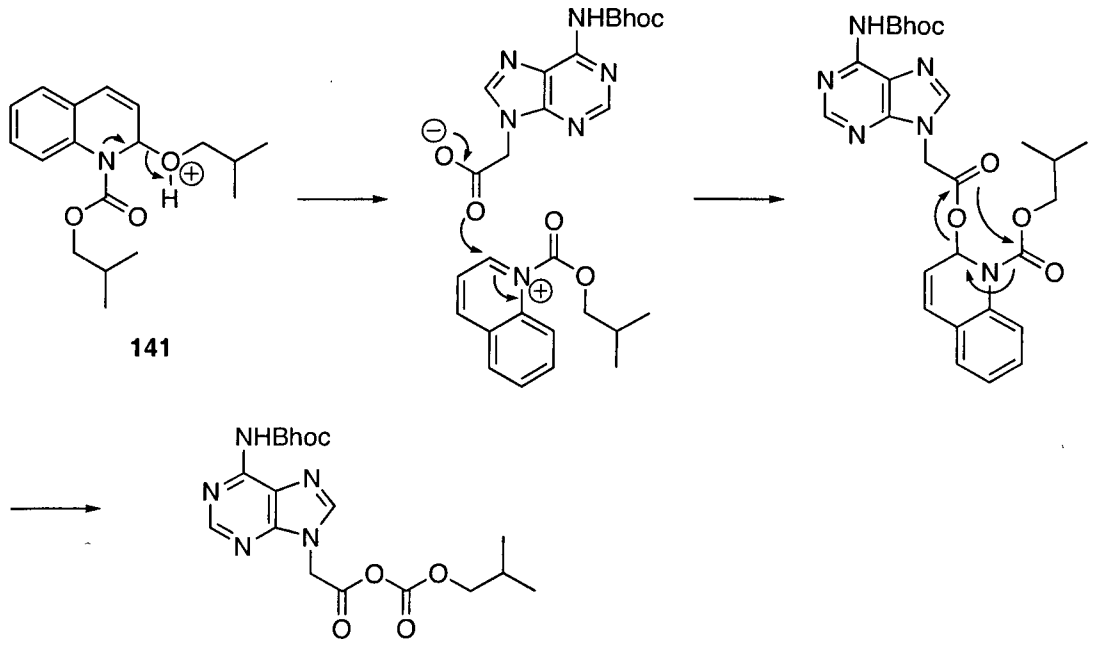


75 **20 %** **140**

(i) N-methyl morpholine 2 equiv., MeCN, RT, 5 min
(ii) Pivaloyl chloride, 0°C, 20 min
(iii) **139**, MeCN/H_2O (1:1), 30 min, RT

Scheme 3.26 Synthesis of **140** using PNA monomer synthesis conditions

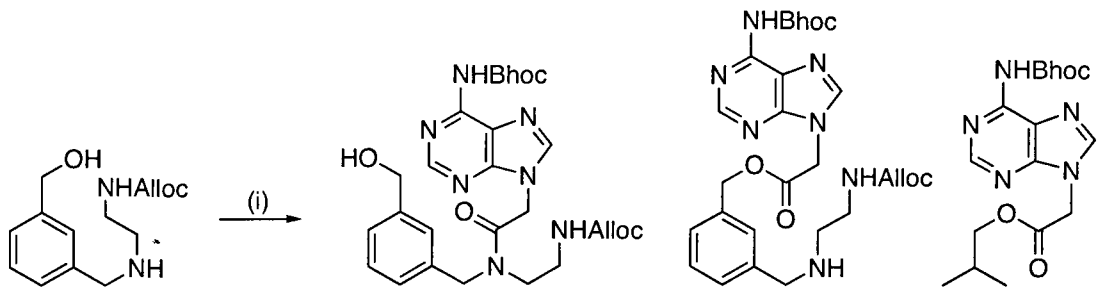
The poor yield was unexpected and reproducible. It was known from the synthesis of PNA monomers that activation of the nucleobase acetic acids using standard reagents other than pivaloyl chloride resulted in their destruction. However a mild carboxylate activating reagent (that cannot be used for the synthesis of the PNA monomers) that could be used to couple **75** and **139** was made investigated. *N*-Isobutyloxy-carbonyl-2-isobutyloxy-1,2-dihydroquinoline (IIDQ) **141** is an unusual peptide coupling reagent because it activates carboxylic acids without base and can do so in the presence of amines (Figure 3.32).[14, 15] In practical terms an amide can be synthesized by simply mixing equimolar amounts of an acid and amine with a slight molar excess of the IIDQ coupling reagent.[16, 17]



141

Figure 3.32 IIDQ activation of Bhoc adenyl acetic acid

An attempt was made to couple **75** and **139** (Scheme 3.27).



139 **140** **142** **143**

(i) **75,** IIDQ, NEt_3 2.0 equiv., CH_2Cl_2, RT, 24h

Scheme 3.27 IIDQ mediated coupling of **75** and **139**

In this instance the IIDQ activated intermediate did not discriminate (at RT) between the amine or alcohols present in the system and gave an inseparable mixture of compounds (Figure 3.33).

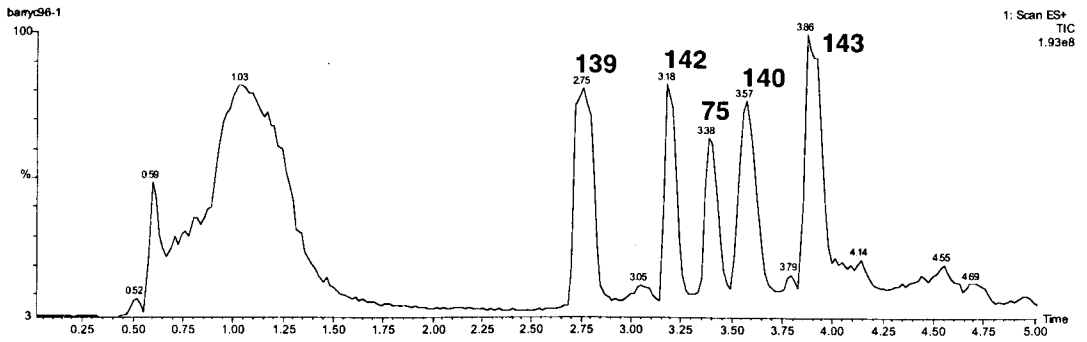
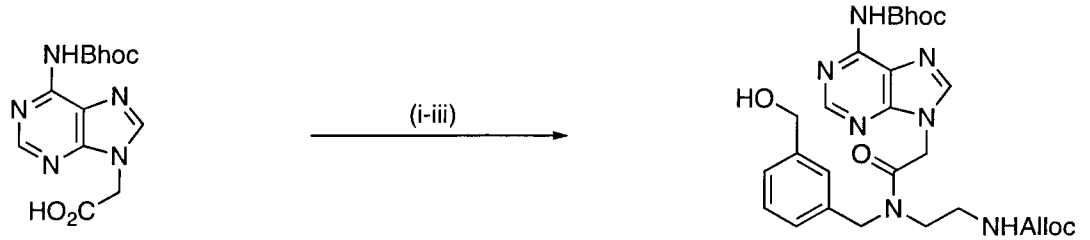


Figure 3.33 LC-MS trace of crude reaction mixture of IIDQ mediated formation of **140**

It was decided that it would be easier to optimise the original mixed anhydride method than the (slow) IIDQ reaction and this approach was abandoned.

A procedure optimised for the conversion of **139** into **140** was developed (Scheme 3.28).



75

96% relative to 139

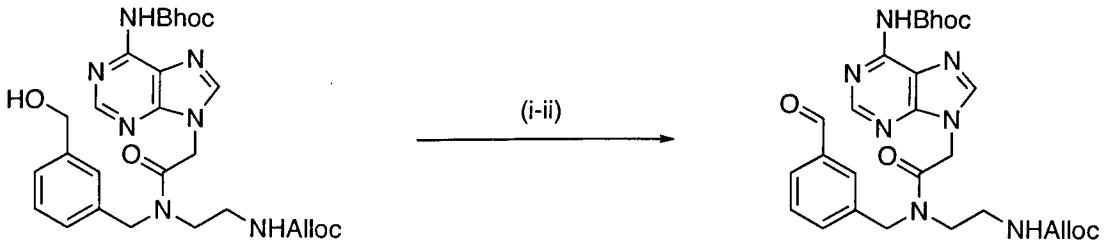
140

(i) N-methyl morpholine (3.0 equiv.), MeCN, RT, 5 min
(ii) -15°C then pivaloyl chloride (1.0 equiv.), -15°C, 30 min
(iii) **139** (0.66 equiv.), MeCN/H_2O (1:1), Et_3N (1.0 equiv.), RT, 36 h

Scheme 3.28 Optimised conditions for the acylation of **139** via mixed anhydride methodology

Amine **139** was considered to be the most precious component of the reaction and its conversion to **140** was maximized by increasing the amount of activated **75** in the reaction by 50%

The final step in the synthesis of the linker was the oxidation of the benzylic alcohol to the aldehyde. A mild oxidation procedure was required that would not disrupt the fragile Bhoc protecting group and ideally give complete conversion into **144** (ie no over-oxidation). TPAP mediated oxidation was initially selected (Scheme 3.29).



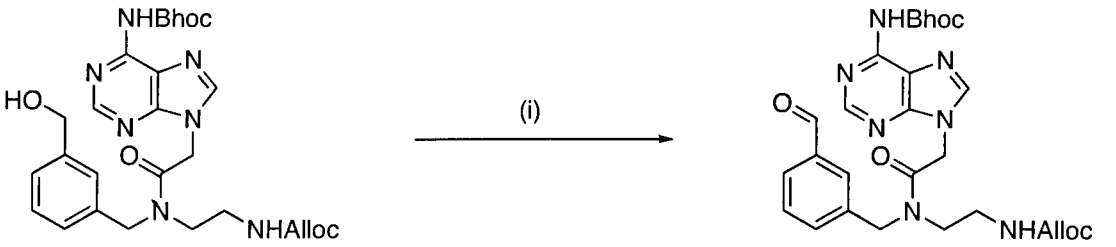
140 **49%** **144**

(i) NMO (1.5 equiv.), 4Å mol sieves, CH_2Cl_2, RT, 10 min
(ii) TPAP (20 mol %), RT, 3 h

Scheme 3.29 TPAP mediated oxidation of **140**

The reaction could not be driven to completion and separation of **144** from the alcohol precursor was very difficult as the two compounds co-elute. Removal of the spent ruthenium metal catalyst was also non trivial and contributed toward the low isolated yield of **144**.

Other methods of oxidation of **140** were investigated to maximize the yield of **144** and minimize waste of precious alcohol **140**. Swern oxidation failed to improve the yield, however Dess-Martin oxidation gave a markedly better yield (Scheme 3.30).



140 **88%** **144**

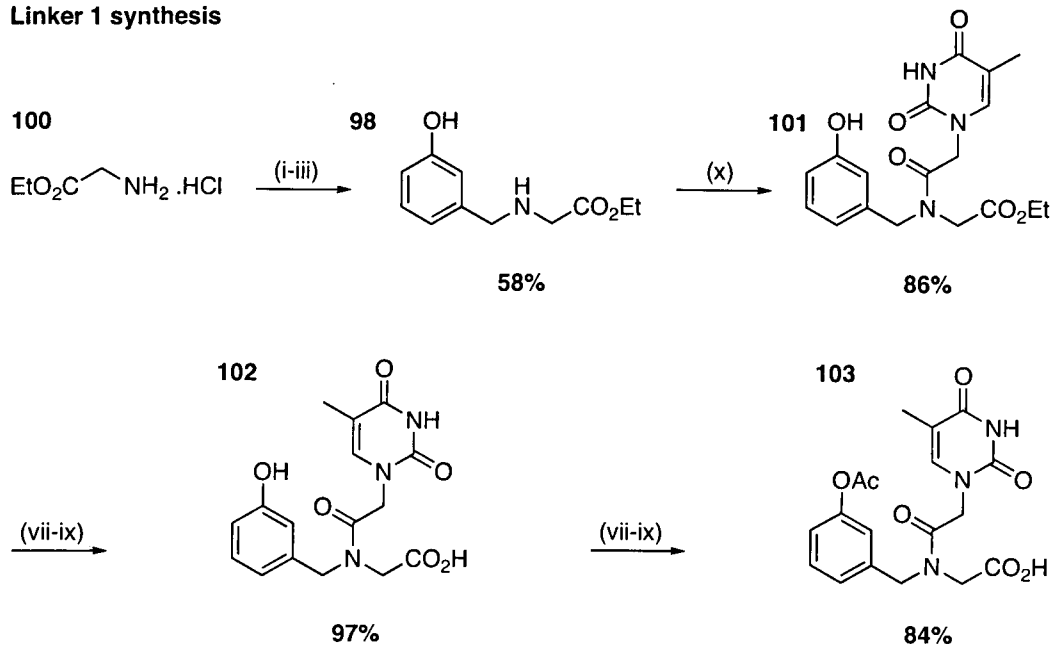
(i) Dess-Martin periodinane (2.0 equiv.), CH_2Cl_2, RT, 3 h

Scheme 3.30 Dess-Martin periodinane mediated oxidation of **140**

The reaction delivered **144** in much improved yield relative to the other methods tested. These conditions were deemed satisfactory for the conversion of **140** into **144** and completed the overall synthesis of the peptide to PNA linker.

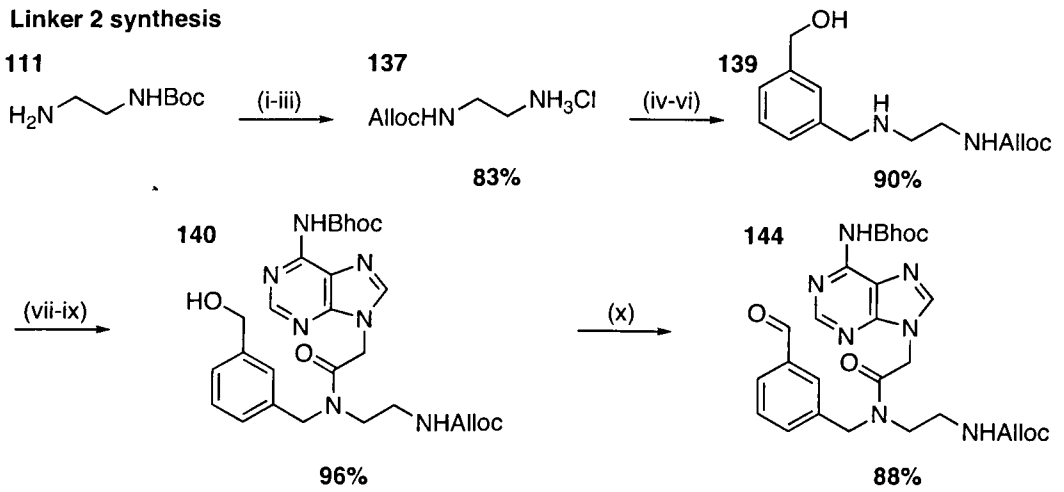
The objective of devising synthetic routes that could deliver both linkers on a multigram scale had been achieved (Figure 3.34).

Linker 1 synthesis



(i) EtOH, potassium hydroxide, 80°C, 10 min
(ii) 3-hydroxybenzaldehyde, RT, 90 min
(iii) Sodium borohydride, RT, 30 min
(iv) MeCN, NMM, RT, 5 min
(v) add to pivaloyl chloride, -15°C, 25 min
(vi) **98,** triethylamine, MeCN/H_2O (1:1), 16 h, RT
(vii) H_2O, lithium hydroxide, 0°C, 5-10 min
(viii) Acetic anhydride, 10% NaOH solution, ice, 1 min

Linker 2 synthesis



(i) NEt_3, CH_2Cl_2, RT, 5 min
(ii) Allylchloroformate (dropwise), CH_2Cl_2, 0°C, 30 min then RT, 24 h
(iii) Acetyl chloride 3 equiv,, MeOH, 0°C, 30 min then **136**, RT, 3 h
(iv) **137**, $NaHCO_3$ solution, RT, 1 h
(v) Evaporate to dryness then $NaBH_4/H_2O$, 0°C, 30 min
(vi) 0°C, 3M HCl
(vii) N-methyl morpholine (3.0 equiv.), MeCN, RT, 5 min
(viii) -15°C then pivaloyl chloride (1.0 equiv.), -15°C, 30 min
(ix) **139** (0.66 equiv.), MeCN/H_2O (1:1), Et_3N (1.0 equiv.), RT, 36 h
(x) Dess-Martin periodinane (2.0 equiv.), CH_2Cl_2, RT, 3 h

Figure 3.34 Summary of Linker synthesis routes

With the PNA monomers and two linker units in hand, the synthesis of the PNA-peptide chimeras was undertaken.

3.3 References

[1] V.J. Gillet, A.P. Johnson, P. Mata, S. Sike, P. Williams, *J. Chem. Inf. Comput. Sci.* **1993**, *7*, 127

[2] A. Gebert, A. Linden, G. Mloston, H. Heimgartner, *Heterocycles* **2002**, *56*, 393.

[3] P. Phukan, *Tetrahedron Lett.* **2004**, *45*, 4785.

[4] F. D. Chattaway, *J. Chem. Soc.* **1931**, 2495.

[5] S. Kurosawa, M. Takenaka, E. Dunkelblum, Z. Mendel, K. Mori, *ChemBioChem* **2000**, *1*, 56.

[6] S. S. Kulkarni, P. Grundt, T. Kopajtic, J. L. Katz, A. H. Newman, *J. Med. Chem.* **2004**, *47*, 3388.

[7] R. Stockman, *SyntheticPages* **2001**, http://www.syntheticpages.org/pages/161.

[8] A. Giannis, K. Sandhoff, *Angew. Chem.* **1989**, *101*, 220.

[9] H. Dodziuk, O. M. Demchuk, W. Schilf, G. Dolgonos, *J. Mol. Struct.* **2004**, *693*, 145.

[10] F. Laduron, V. Tamborowski, L. Moens, A. Horvath, D. De Smaele, S. Leurs, *Organic Process Research & Development* **2005**, *9*, 102.

[11] Y. Kobayashi, A. Fukuda, T. Kimachi, M. Ju-Ichi, Y. Takemoto, *Tetrahedron* **2005**, *61*, 2607.

[12] A. Boeijen, J. van Ameijde, R. M. J. Liskamp, *J. Org. Chem.* **2001**, *66*, 8454.

[13] A. Nagalingam, *University of Leeds* **2007**.

[14] Y. Kiso, H. Yajima, *J. Chem. Soc., Chem. Commun.* **1972**, 942.

[15] D. Hudson, *J. Org. Chem.* **1988**, *53*, 617.

[16] D. J. Cremin, A. F. Hegarty, M. J. Begley, *J. Chem. Soc., Perkin Trans. 2,* **1980**, 412.

[17] B. Belleau, G. Malek, *J. Am. Chem. Soc.,* **1968**, *90*, 1651.

Chapter 4

With reliable synthetic routes to each of the six starting materials established, it was now possible to attempt the construction of the PNA-peptide chimeras. The proposed work was without precedent, and as such a significant amount of solid phase chemistry remained to be developed for the synthesis of the full size chimeras. In particular, conditions had to be developed to incorporate the linker constructs **103** and **144** into the chimeras that are also compatible with SPPS. In preparation for the synthesis of full size chimeras, small model systems were devised to develop the chemistry required for the ligation experiments. Ultimately our aim with this work was to establish reliable and generic routes for the synthesis of PNA to peptide and peptide to PNA chimeras that would scale up to preparative quantities.

4.1 Synthesis of a PNA to peptide chimera

We were aware that the synthesis of the large target PNA to peptide chimera (Figure 4.1) could be non-trivial at some stages of its assembly.

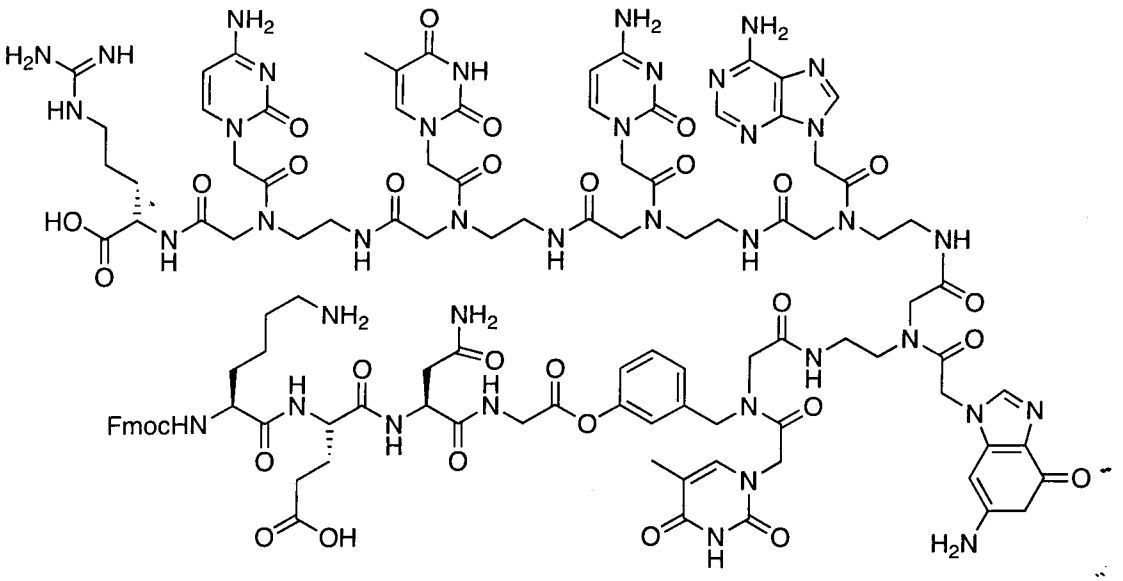
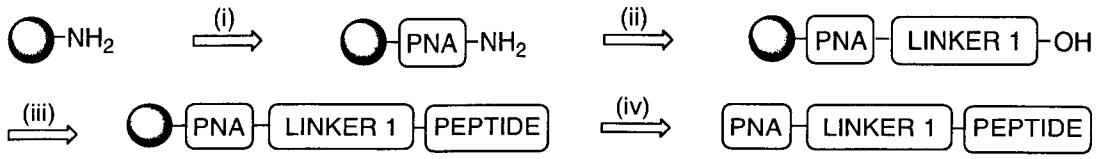


Figure 4.1 PNA to peptide chimera target molecule

Acylating the linker phenol with an amino acid had been anticipated to be potentially difficult, according to prior knowledge of phenolic ester formation within the group. Successfully growing a peptide from the amino acid anchored to linker **103** would also present its own set of problems if the peptidyl phenolic ester proved to be unstable to

standard SPPS conditions. It was decided to first make model chimeras to develop chemistry that would overcome these and any other problems that could arise and allow for a smooth synthesis of the full size chimera. The linear synthesis of the model chimera involved three stages before cleavage from the resin (Figure 4.2).



(i) PNA oligomer synthesis
(ii) Linker attachment and deprotection
(iii) Peptide oligomer synthesis
(iv) Resin cleavage

Figure 4.2 Assembly of a PNA to peptide chimera

To minimise the amount of PNA monomer used in these test studies it was decided to develop the solid phase chemistry in discrete out-of-sequence stages. The first stage was to establish conditions for loading the linker onto a solid-supported amine and deprotect it (Figure 4.3).

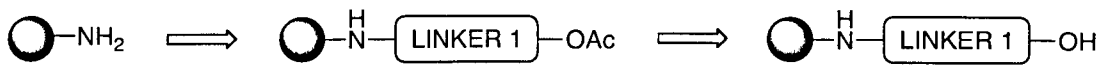


Figure 4.3 Stage one of the solid supported synthesis of the PNA to peptide chimera

The second stage built upon the results of the first stage by grafting amino acids onto the phenol. This represented two tasks. First and foremost was finding a method to quantitatively acylate the phenol with an amino acid. The second was to maintain the phenol ester linkage intact during the elaboration of the amino acid into a large polypeptide using standard SPPS conditions (Figure 4.4):

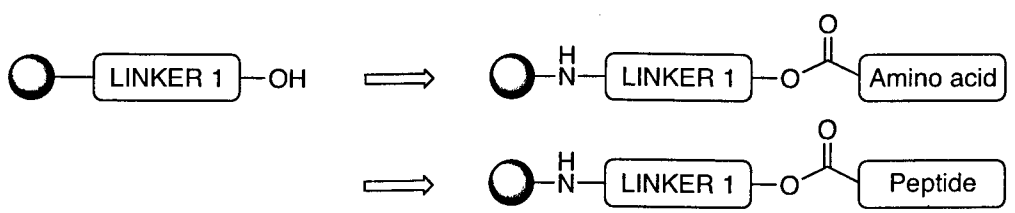


Fig 4.4 Stage two of development of the solid supported PNA to peptide chimera

The third stage was to establish conditions for the clean synthesis of a short PNA oligomer (Figure 4.5):

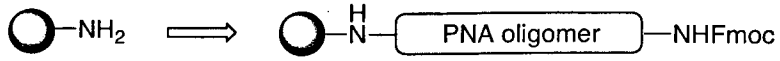


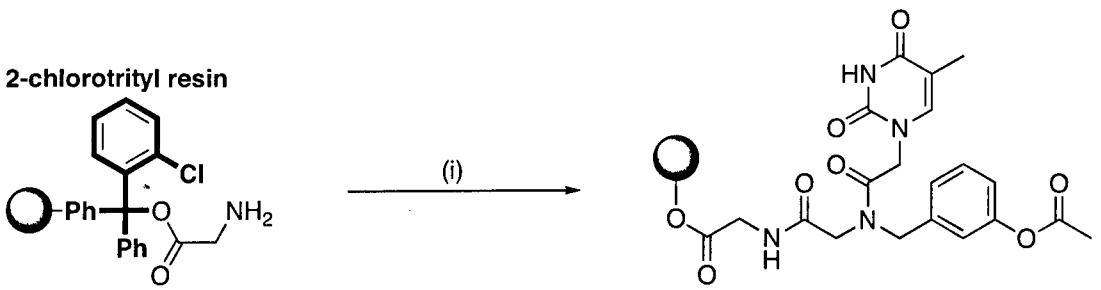
Fig 4.5 Stage three of development of the solid supported PNA to peptide chimera

When all three stages were completed the conditions established therein would be used in the correct order (3,1,2) to synthesise a PNA to peptide chimera. By developing stages one and two independently of stage 3 it was thought that potentially a substantial amount of PNA monomer material would be conserved.

4.1.1 Stage 1 – Linker 1 attachment and deprotection

It was decided to build the model chimeras upon 2-chlorotrityl resin as it offers mild cleavage conditions and can also be purchased in preloaded form.

2-Chlorotrityl resin (preloaded with glycine) **145** was coupled with linker **103** that had been preactivated with HCTU and DIPEA (Scheme 4.1).



145 **146**

(i) **103**, HCTU (4.9 equiv.), diisopropylethylamine (10.0 equiv), DMF, 1 h

Scheme 4.1 Attachment of **103** to glycine preloaded on 2-chlorotrityl resin

Treatment of an analytical sample of resin with a standard cleavage cocktail of TFA/triethylsilane/water (95.0:2.5:2.5) for 30 min was followed by LC-MS analysis (Figure 4.6).

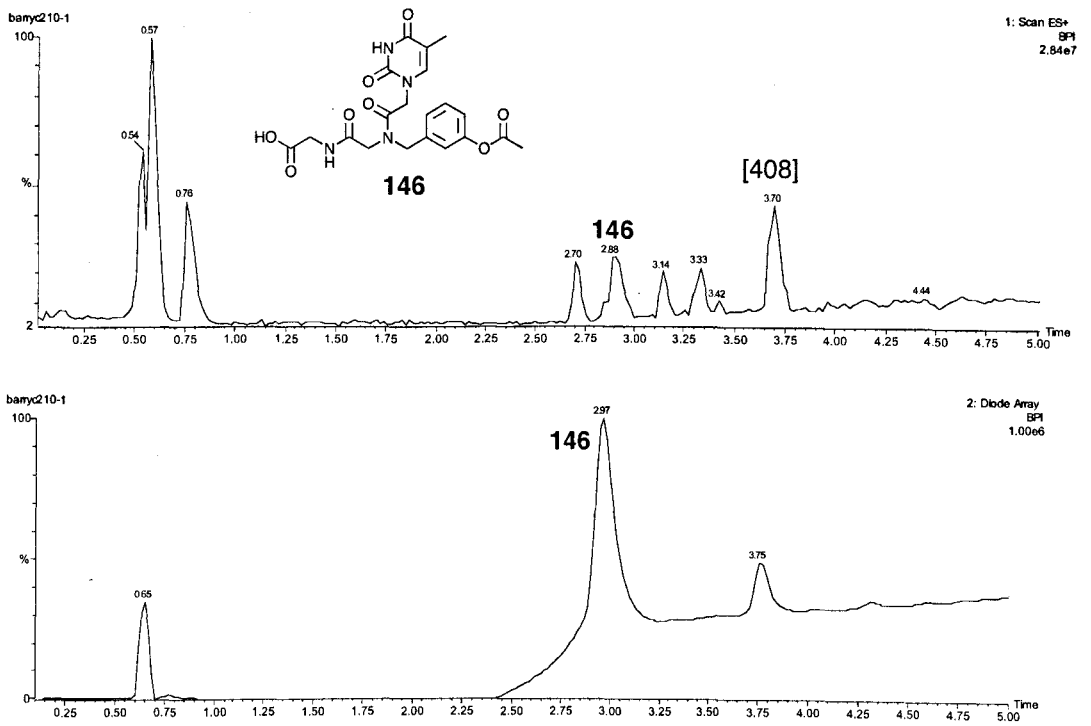
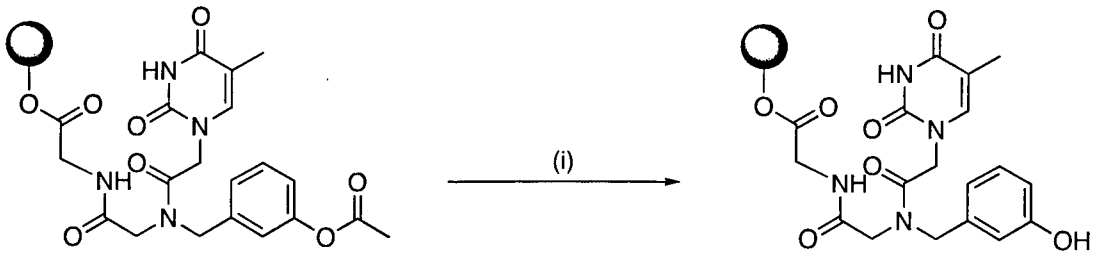


Figure 4.6 LC-MS (BPI, upper trace) and LC-UV (lower trace) data of cleaved glycine-linker hybrid

Despite considerable noise within the ion trace (a characteristic problem with weak analytical samples) the product was both present and was also the major UV-active component. The other minor UV-active compound could not be identified. With the linker in place, deprotection of the phenol was undertaken. The acetate protecting group is well known, especially in solid-supported carbohydrate chemistry. It was thought that conditions used from that class of chemistry would be equally adept in the deprotection of **146**. A procedure by Seeberger for the synthesis of β 1-4 linked polyglucosaccharides was investigated wherein an axial acetate was efficiently removed with sodium methoxide solution.[1] The same conditions were used to cleave the acetate of **146** (Scheme 4.2):

2-chlorotrityl resin



146 **147**

(i) MeONa (6.0 equiv.), MeOH/CH_2Cl_2(1:9), RT, 2 h

Scheme 4.2 Methanolysis of acetate protecting group

Treatment of the resin with sodium methoxide gave complete acetate cleavage. Standard acidic cleavage of the resin liberated **147b** and was detected by LC-MS analysis (Figure 4.7).

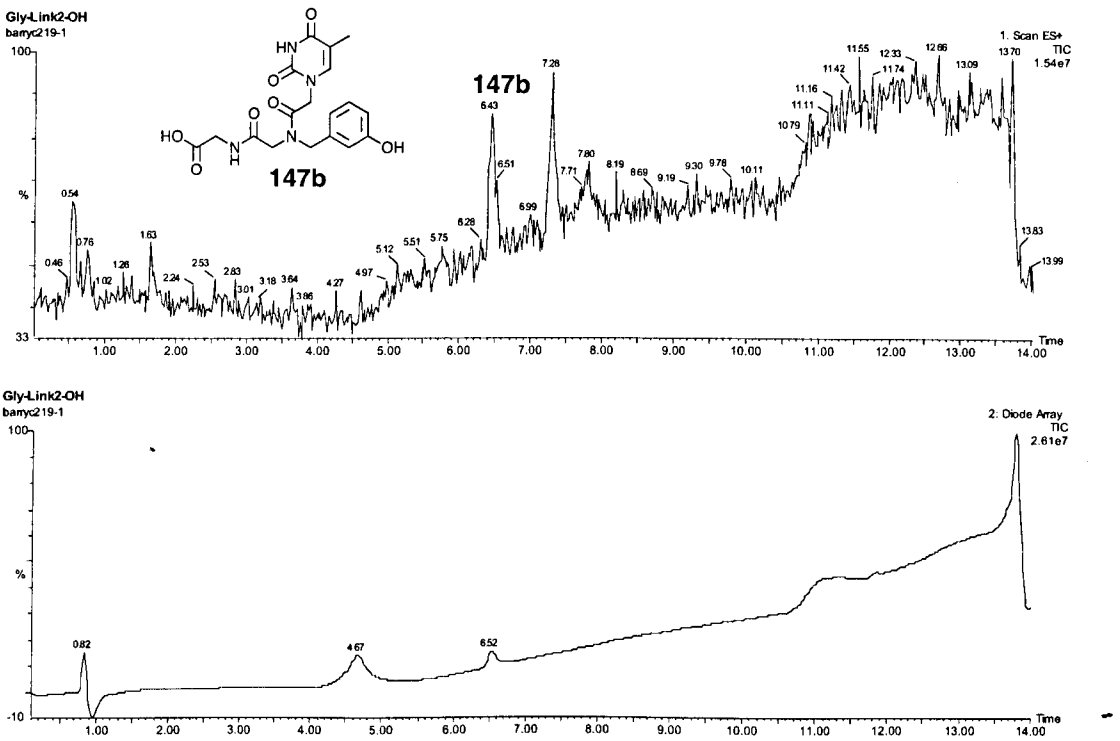


Figure 4.7 LC-MS data of the acetate deprotection

The scale of these test reactions was such that the small amounts of material used to make **147b** were insufficient for NMR analysis. Given the limited time remaining, it was decided to press ahead with the synthesis of the chimeras.

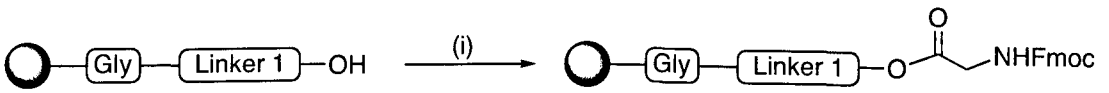
It was felt that the conditions developed during this work were suitable for attaching and deprotecting the linker on a full size chimera; the second stage of chimera synthesis was then undertaken.

4.1.2 Stage 2 – Growth of a peptide from the unmasked phenol

Ultimately, in the full-size chimera ligation experiments, it is desired that the peptide attached to the phenol will ideally be any sequence of choice. However for the purposes of the test ligation, and to promote the best chance of success in the initial ligation experiments, the first amino acid will always be a glycine residue. This was decided upon so to minimise any steric repulsion between the ligating junctions, caused by the clashing of amino acid side-chains; it was thought that such interactions could retard the ligation process.

The next stage in the synthesis to be developed was the stepwise assembly of a polypeptide from the unmasked linker phenol. It was known that forming and keeping a phenolic ester intact can be difficult, however we were confident that reaction conditions could be found to suit our needs. Furthermore a suitable degree of reactivity of the ester was welcomed as this would serve to promote the directed ligation reaction.

Acylation of phenol **147** with Fmoc-Glycine was attempted using conditions that were proven for phenolic ester synthesis (Scheme 4.3).[2]



(i) Fmocglycine (5.0 equiv.), EDC (5.0 equiv.), DMAP (5.0 Equiv.), DMF, RT, 16 h

Scheme 4.3 Attachment of fmocglycine to the linker phenol

A sample of the resin was cleaved under very mild acidic conditions (TFA/AcOH/CH_2Cl_2, 1:1:8) to minimise phenol ester hydrolysis that could occur as a result of low pH.

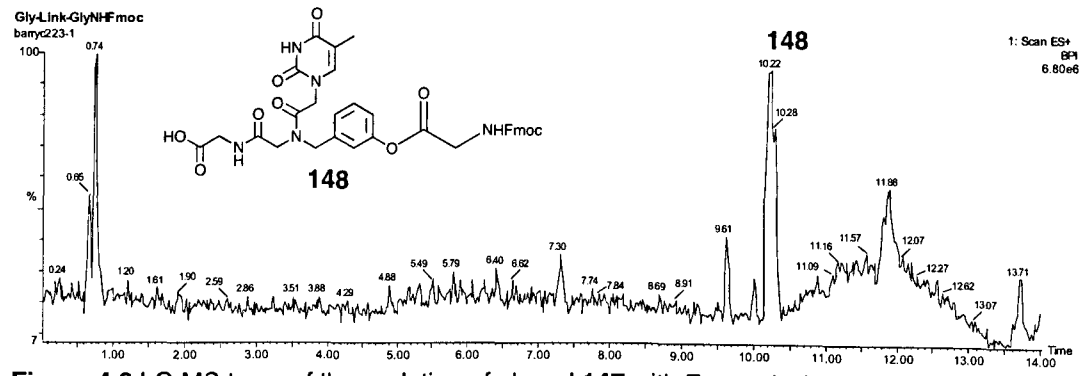
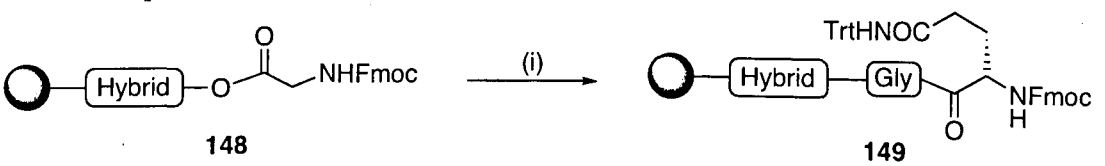


Figure 4.8 LC-MS trace of the acylation of phenol **147** with Fmoc glycine

148 was observed as the major product although a small amount of starting phenol **147** was also present. The resin was treated again with the peptide coupling conditions and a test cleavage revealed that the phenol was still present in the sample. This result suggested that double coupling does not improve the formation of ester **148**. There were a few possible reasons to explain why the reaction appears to be incomplete, however it was most likely that the phenolic ester was slightly unstable under the basic reaction conditions, resulting in partial hydrolysis of the ester.

With the Fmoc glycine residue appended onto the phenol, the remainder of the test peptide could then be assembled using standard SPPS conditions. The sequence of this test peptide was decided to be Gly-Gln-Lys-Val. After two residues had been attached (Gly-Gln), a portion of the resin was cleaved under standard conditions and analysed (Scheme 4.4).



(i) 5 x DMF/piperidine (4:1) wash , RT, 2 min then 5 x DMF, RT, 2 min
(ii) Fmoc Glutamine (Trt) (5.0 equiv.), HCTU (4.9 equiv.), DIPEA (10 Equiv.), DMF, RT, 1 h
(iii) 5 x DMF wash, RT, 2 min

Scheme 4.4 Extension of linker bound peptide

Mild acid cleavage of the resin repeatedly failed to give a good LC-MS trace and so it was decided to proceed with strong acid cleavage conditions instead. The strong acid cleavage conditions would however now bring the stability of the phenolic ester linkage under scrutiny once more. A portion of the resin was dried and cleaved using a cleavage cocktail of TFA/TIS (95:5) and analysed by LC-MS (Fig 4.9).

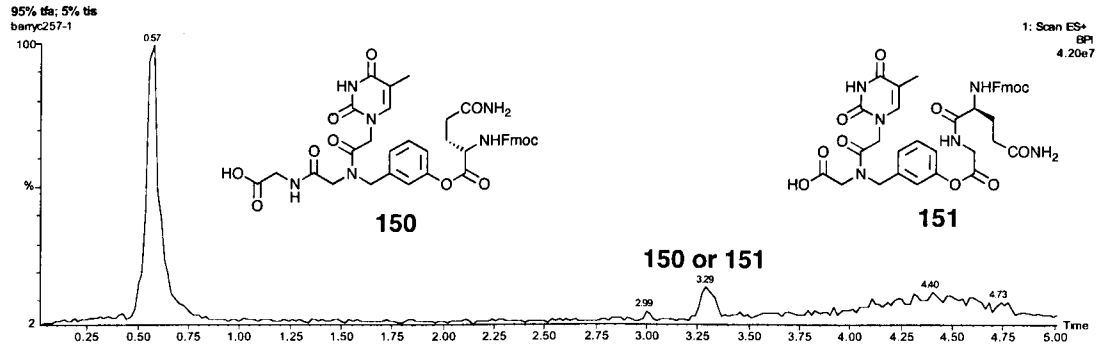
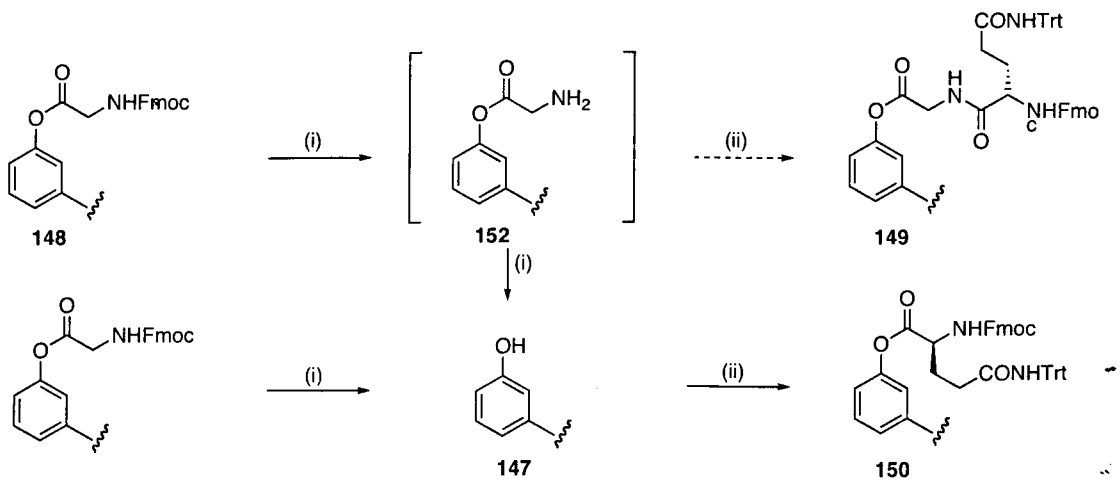


Figure 4.9 LC-MS trace of strong acid cleavage of resin bearing derivitised phenol

LC-MS analysis revealed that there was only one major product but that it not the desired compound. The mass ion observed was equal to the mass of product **148** but less the mass of glycine. There are two molecules of glycine in hybrid **148** but it could not be established if glycine had been lost from the product (giving **149** or **150**) or if the peak corresponded to a non-related compound. The most logical explanation was that **148** was unstable to Fmoc deprotection conditions (20% piperidine in DMF) and that the ester **148** was therefore cleaved, regenerating phenol **147** (Scheme 4.4). Treatment of **147** with the HCTU-activated glutamine would then give the observed mass ion. However, if Fmoc deprotection occurred before the proposed cleavage (**151**) and had activated the cleavage process (**151** to **147**) was not known.



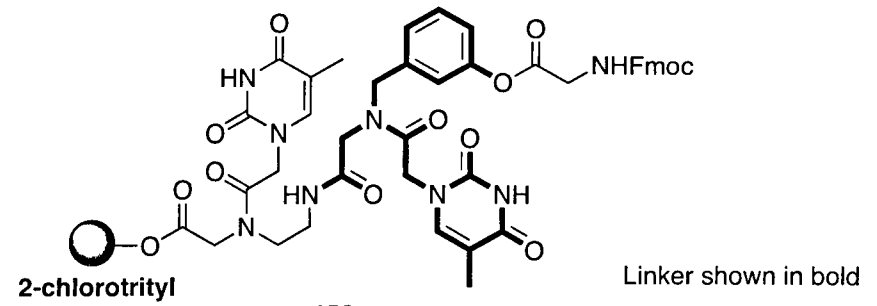
(i) Piperidine/DMF (1:4)
(ii) HCTU, diisopropylethyamine, Fmocglutamine(Trt)

Scheme 4.4. Possible fates of **148** undergoing second cycle of pepide synthesis based on observed MS data

To determine (unambiguously) if glycine was being lost during the strong acid resin-cleavage of the hybrid, an analogue hybrid needed to be built upon 2-chlorotrityl resin that was not preloaded with glycine. This analogue would therefore only contain one

instance of glycine in its structure (the phenolic glycine ester) and using this compound for the SPPS experiments would determine with absolute certainty if indeed the phenolic glycine ester was base-labile.

The hybrid analogue was designed with the 2-chlorotrityl-linked glycine being replaced by a T-PNA monomer, the most accessible and abundant of the four monomers (Figure 4.10).



153

Figure 4.10 Alternative test hybrid

2-Chlorotrityl chloride resin was loaded according to an established procedure with T-PNA monomer **37**.[3] Freshly loaded 2-chlorotrityl chloride resin is unstable in the presence of an Fmoc group[3] and PNA monomers are unstable when unprotected,[4] thus immediate deprotection and coupling with linker **103** was required (Figure 4.11). Resin loading was adjudged to be very poor as approximately half of the resin was required for cleavage to obtain a signal in the LC-MS trace. The desired intermediate hybrid **154** was the only product observed.

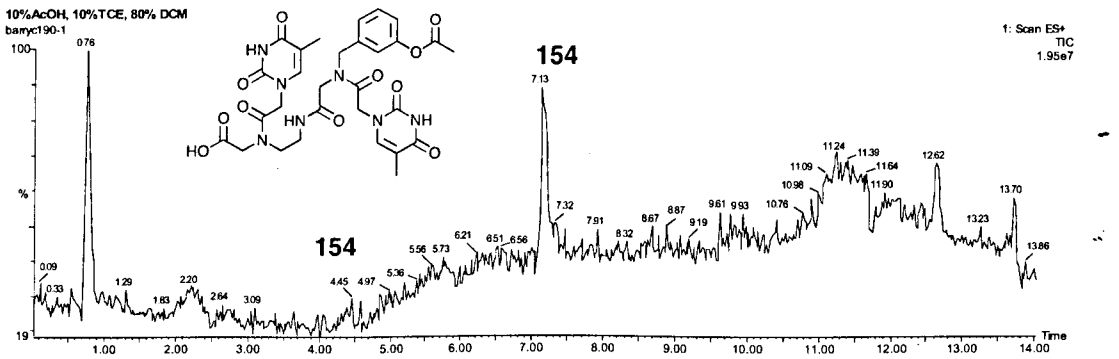


Figure 4.11 LC-MS of T-PNA/linker hybrid cleaved from manually loaded 2-chlorotrity resin

The resin was treated with the acetate cleavage conditions of Seeberger and half of the remaining resin (half was necessary due to poor loading) was cleaved for LC-MS analysis (Figure 4.12).

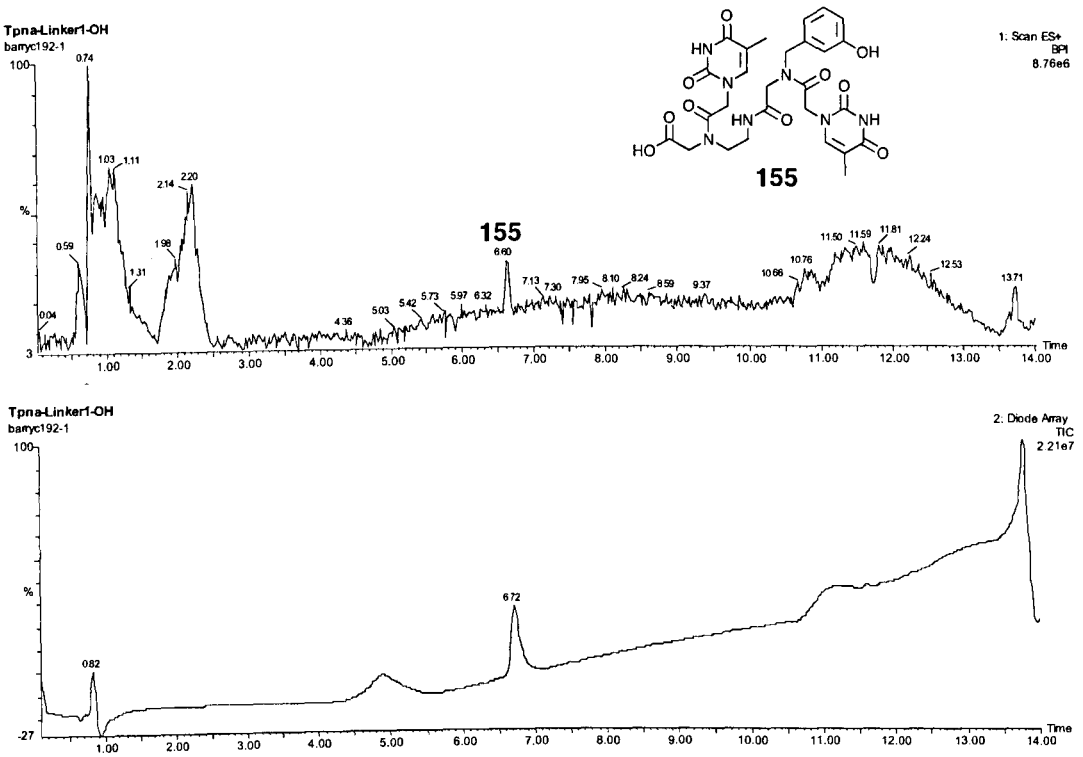


Figure 4.12. LC-MS data of deprotected T-PNA/linker hybrid

Despite the weakness of the LC-MS traces, it was observed that the phenol analogue had been synthesised. There remained three synthetic steps to reach the desired dipeptide ester; coupling of Fmocglycine, deprotection of the Fmocglycine ester and coupling of the next amino acid (Fmocglutamine). However, at this point there was only enough resin sufficient for a single analytical cleavage experiment, therefore the resin could not be analysed until the end of the three-step synthetic sequence.

The three step synthesis of the dipeptide was performed and all of the remaining resin was cleaved under anhydrous strong acid conditions (TFA/TES, 95:5). Unfortunately LC-MS analysis was inclusive as there was insufficient material to generate an observable signal; poor loading of the resin had prevented there being enough material for analysis of the final compound. It was thought that the resin would have to be preloaded to overcome this problem.

The model ester **148** was resynthesised on 2-chlorotrityl resin preloaded with glycine. This compound was used to study the stability of the phenol ester and it was now assumed that only the resin-linked glycine component was not labile.

A study to determine if (and how) the ester was being destroyed was then undertaken. There was conceivably more than one way in which the ester could be converted into the phenol (Figure 4.13) and it was necessary to determine the pathway of degradation.

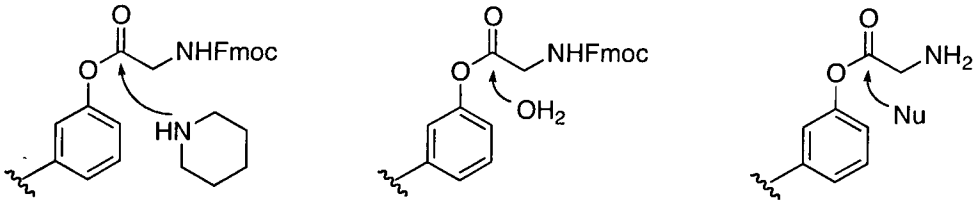
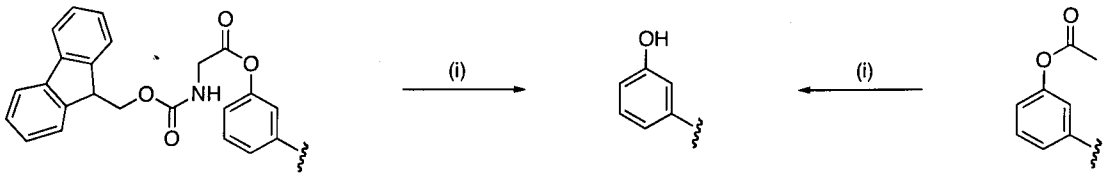


Figure 4.14 Illustration of the hypothesis that nucleophilic cleavage conditions affect **148** and **146** equally

To examine the effect of piperidine on the phenolic ester, **148** was treated with standard Fmoc deprotection conditions and an analytical portion of the resin was cleaved (TFA/TIS, 95:5) and analysed immediately (Figure 4.14). The sample showed quantitative conversion to the phenol, demonstrating that the Fmocglycine phenolic ester was not stable to piperidine and that SPPS in the presence of the ester was not possible using standard conditions.



148 **147** **146**

(i) Piperidine/DMF (1:4)

Figure 4.13 Possible pathways of phenolic ester degradation under SPPS conditions

To determine if the Fmocglycine ester was especially sensititive to the piperidine solution, the experiment was repeated with acetate-protected precursor **146** (Figure 4.14). Treating acetate **146** with 20% piperidine/DMF, as per standard fmoc deprotection conditions, did not fully remove the acetate protecting group (Figure 4.15).

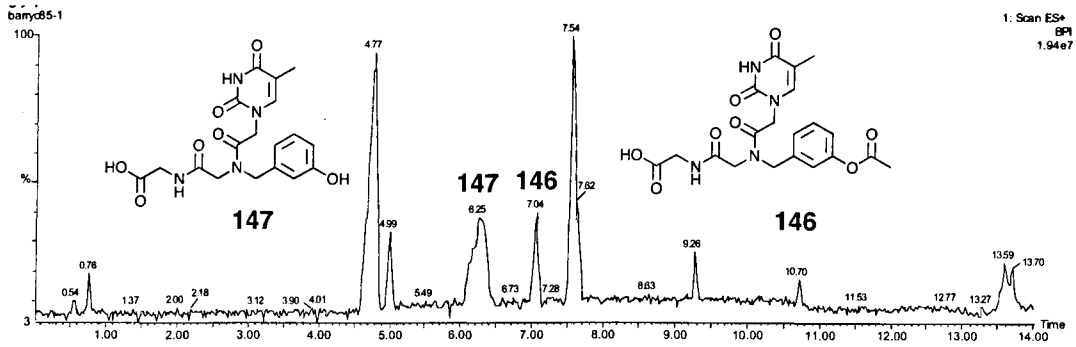


Figure 4.15 LC-MS trace of incomplete aminolysis of phenol acetate **146** with piperidine

Repeating the deprotection sequence completed aminolysis of acetate **146** (Figure 4.16).

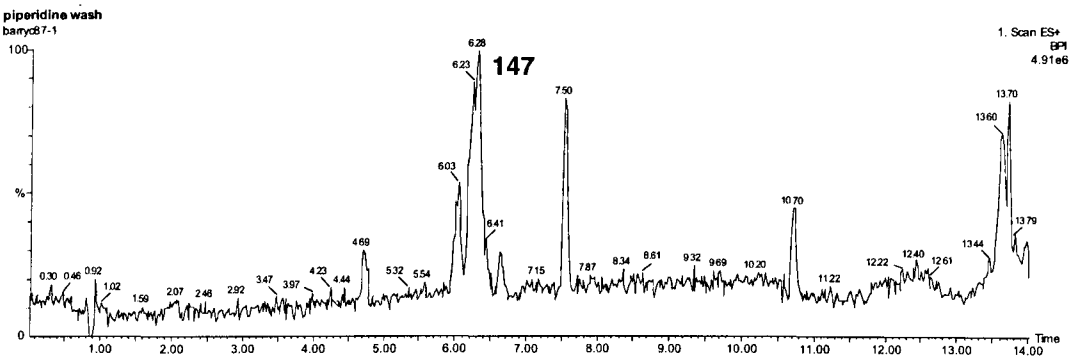


Figure 4.16 LC-MS trace of the total aminolysis of phenol acetate **146**

It was concluded that because the aminolysis of the acetate was sluggish compared to **148** that the Fmocglycine ester was not undergoing simple aminolysis. Interestingly, the aminolysis of **146** had demonstrated a new set of deprotection conditions for the acetate.

It was thought that the product obtained from the initial studies (Figure 4.9) was **150**. It was believed that piperidine mediated aminolysis of the Fmocglycine ester regenerated the phenol which in turn was then acylated with HCTU-activated Fmocglutamine (Figure 4.17).

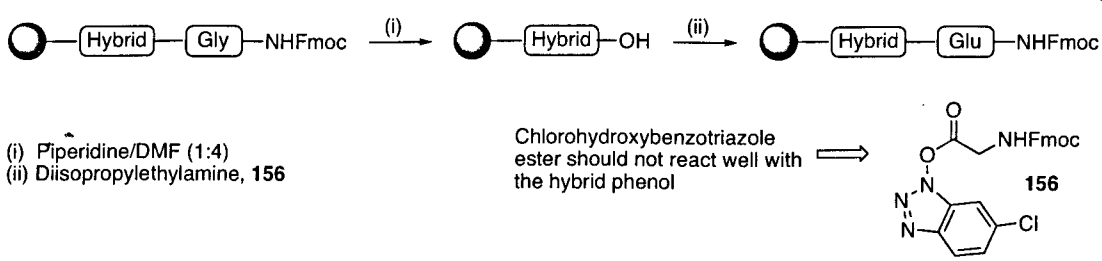
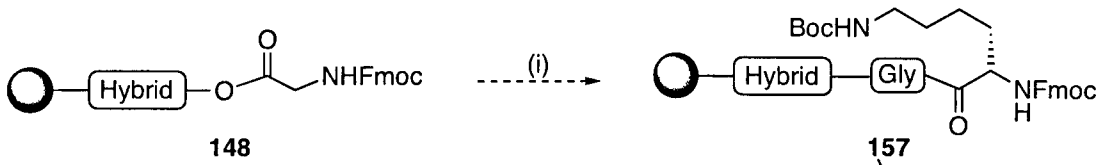


Figure 4.17 Schematic representation of possible HCTU-activated amino acid coupling with phenol

It was thought however that an amino acid activated with HCTU would not efficiently acylate the phenol. To determine if the phenol could be acylated with an HCTU-activated amino acid, phenol **147** was treated with Fmocglycine under standard coupling conditions. It was found that the phenol was easily converted to **148** using these conditions. This result confirmed that standard SPPS conditions caused ester aminolysis and concomitant HCTU-mediated acylation of the phenol with the next amino acid of the peptide sequence.

In order to overcome this problem, alternative Fmoc deblocking conditions were required that would not cause aminolysis of the ester.

Non-nucleophilic Fmoc deprotection conditions reported by Bu *et al.* detailed the synthesis of a decapeptide bound to tentagel resin *via* a reactive thioester linkage.[5] These conditions were thought to be suitable for the Fmoc deprotection (in the presence of an electrophilic ester) and were used during an attempt to couple lysine onto the glycine ester (Scheme 4.5).



148 **157**

(i) 3 x DBU/HOBt (0.80 mMol:0.74 mMol) DMF, RT, 1 min then 5 x DMF, RT, 1 min
(ii) Fmoc Lysine(Boc) (5.0 equiv.), HCTU (4.9 equiv.), DIPEA (10 Equiv.), DMF, RT, 1 h
(iii) 5 x DMF wash, RT, 2 min

Scheme 4.5 Non-nucleophilic peptide fmoc deblocking conditions

What was observed was the conversion of **148** into a 2:1 mixture of **158** and **157** (Figure 4.18).

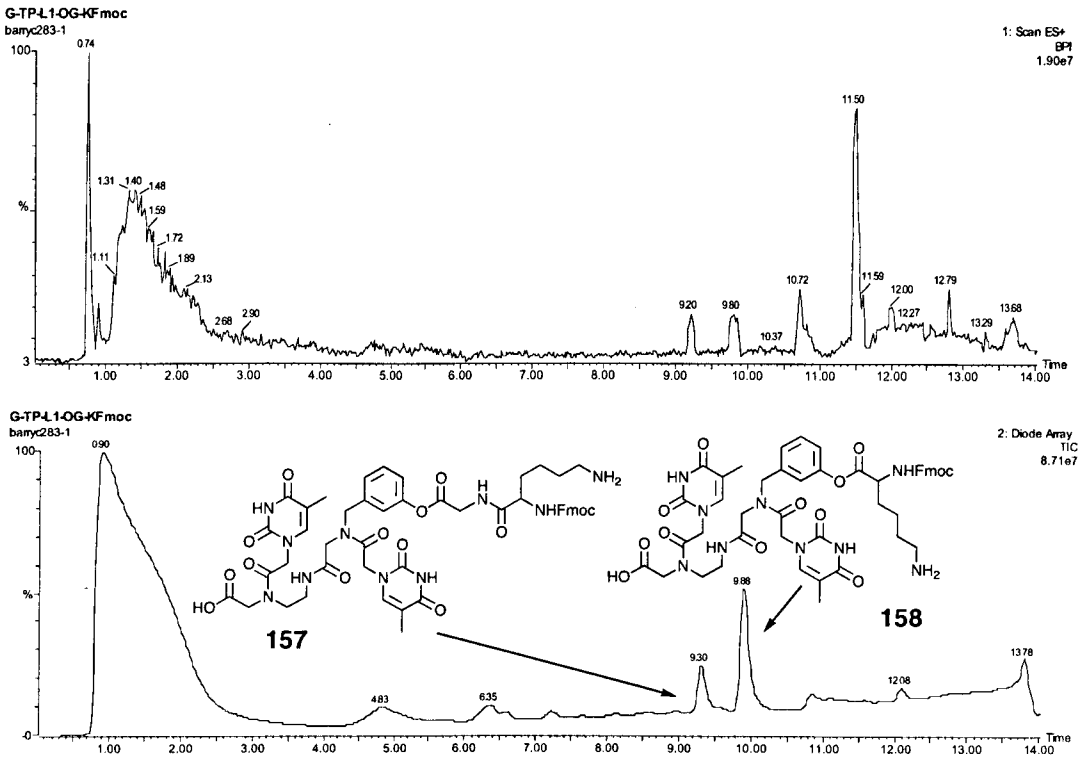


Figure 4.18 LC-MS trace of peptide elongation experiment using non-nucleophilic base

The Fmoc glycine was only exposed to DBU for a relatively brief period, however the (relative amounts of) products observed show that the majority of **152** had dissociated into the phenol before the coupling reaction could occur.

This result demonstrated that the Fmoc deprotected glycine ester is not stable at basic pH and that aminolysis (by the Fmoc-deblocking reagent) is not required for the ester to degrade.

To test if the Fmoc-protected ester was stable to hydrolysis at basic pH, an LC-MS sample of **148** was treated with aqueous sodium hydrogen carbonate until the pH of the sample had been raised to pH 9, and was then agitated for 20 min. Under these conditions the Fmoc group was stable and would remain intact. The sample was resubmitted for analysis by LC-MS, to search for evidence of ester hydrolysis at basic pH with the Fmoc group intact; the hydrolysed ester was not observed, strongly suggesting that the sample had remained unchanged. The results obtained thus far had shown that the Fmoc-protected ester was stable to hydrolysis up to at least pH 9.0 and that ester degradation occurred rapidly once deprotection had occurred. This degradation had been shown to be due to independent of the nucleophilicity of the amine used to effect Fmoc deprotection. These results therefore suggested that auto-aminolysis of the deprotected ester was occurring; inter-chain aminolysis was possible

but improbable as the resin loading was too low (<1 mmol per gram) for this to be a feasible consideration.

To test if the ester was undergoing auto-aminolysis, **147** was synthesised on 2-chlorotrityl resin (preloaded with glycine) and coupled with Boc-glycine. The idea was to trigger (acid-catalysed) global resin cleavage and ester deprotection to give the protonated glycine amine; in this form, auto aminolysis could not occur and therefore the deprotected ester would be observable by LC-MS. The resin was treated with a standard cleavage of TFA/TIS (95:5) and analysed by LC-MS (Figure 4.19).

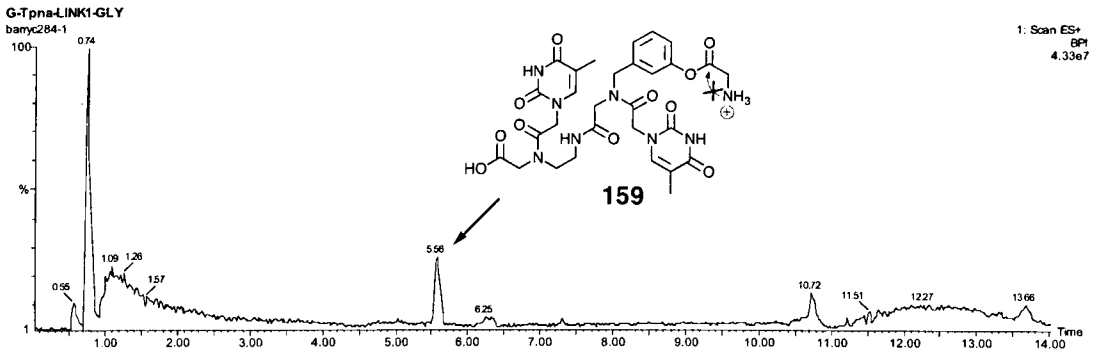
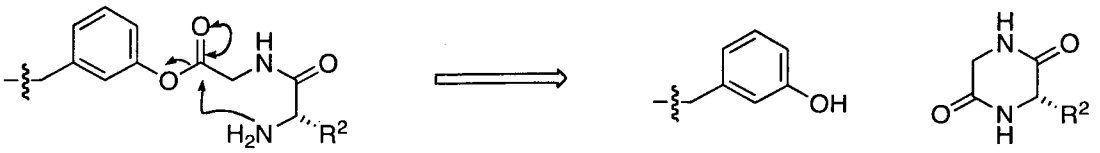


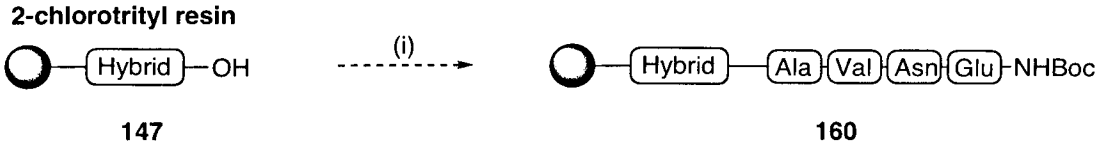
Figure 4.19 LC-MS trace of Boc deprotected variant of the model hybrid

The deprotected glycyl phenolic ester **148** was observed as the only product in the LC-MS trace. To show that the free amine was responsible in causing the ester degradation, the LC-MS sample of **159** was neutralised with non-nucleophilic diisopropylethylamine and resubmitted for LC-MS analysis. Phenol **147** was observed as the only product. The experiments had shown that the free amine was unstable at neutral or basic pH. For auto-aminolysis to occur however, a highly strained 3-membered transition state has to be passed though, and this seems unlikely to occur. How the ester is degrading is therefore still uncertain, although in the absence of other explanations it could be that inter-chain aminolysis was occurring on-resin. The main conclusion from the investigation is that it is not possible to grow a peptide from **147** using standard Fmoc solid phase chemistry. It was thought that the only option available to overcome this problem was to attach a fully assembled polypeptide onto the phenol. Although this solution increased the labour involved for the chimera synthesis, it was also welcomed as it increased synthetic convergency and circumvented a potential problem that could have occurred when growing the polypeptide from two to three residues (diketopiperazine formation, Figure 4.20).



Scheme 4.20 Potential diketopiperazine (DKP) formation after Fmoc deprotection of second amino acid

To complete the synthesis of the chimera, the coupling of a small tetramer peptide with **147** was investigated. Phenol **147** was synthesised again on 2-chlorotrityl resin and treated under standard coupling conditions with the protected tetramer peptide AVN(Trt)E(Boc)NHBoc that had been prepared previously in the group by Nagalingam (Scheme 4.7).[6]



147 **160**

(i) AVN(Trt)E(Boc)NHBoc (5.0 equiv.), HCTU (4.9 equiv.), DIPEA (10 Equiv.), DMF, RT, 1 h
(ii) 5 x DMF wash, RT, 2 min

Scheme 4.7 Coupling of tetramer peptide to **147**

The coupling reaction was unsuccessful. Analytical portions of the resin were cleaved under both mild and strong acid conditions and analysed by LC-MS; both sets of conditions gave the same trace (Figure 4.21).

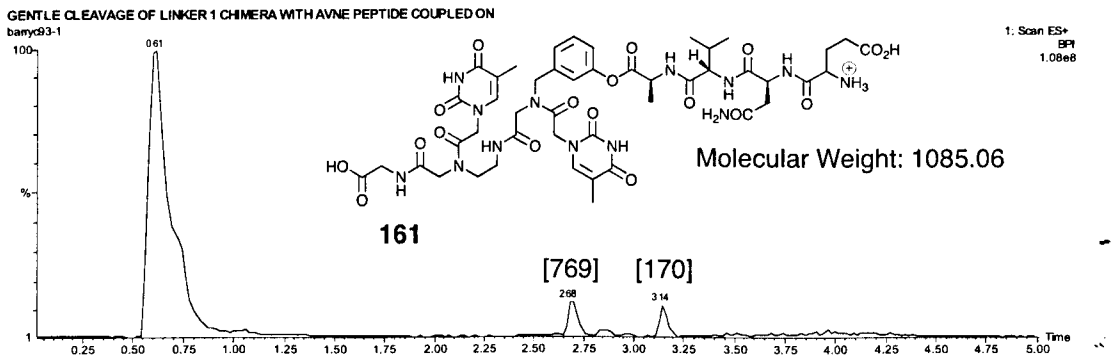


Figure 4.21 LC-MS trace of tetramer peptide coupling experiment

161 has a $[M+H]^+$ of 1086 *m/z* and sidechain protected **161** has a $[M+H]^+$ of 1383 *m/z*; neither of the two peaks observed in the LC-MS spectrum were related to the masses of

the product or the starting materials. The sum of the two observed masses (assuming they are $[M+H]^+$ ions) does not equal, nor is close to, the mass required for **161** (nor its protected form), implying that product **160** had not degraded under acid cleavage into two smaller fragments ie protolysis of benzylic amine bond.

The short peptide used as the coupling partner for **147** was analysed for contaminants by LC-MS and found to be a pure sample of the expected tetramer. The coupling experiment with this peptide was repeated with similarly unsuccessful results.

To establish if and how the peptide was involved in the synthesis of the unknown compound (of mass [768]) the phenol-peptide coupling experiment was repeated with a different tetramer.

In preparation of the new coupling experiment, peptide fully protected G-N(Trt)-Glu(^{t}Bu)-Lys(Boc)NHBoc **162** was synthesised on 2-chlorotrityl resin preloaded with glycine on a large (115.5 μMol) scale. For the purpose of the qualitative coupling experiment crude **162** was found to be sufficiently pure (Figure 4.22).

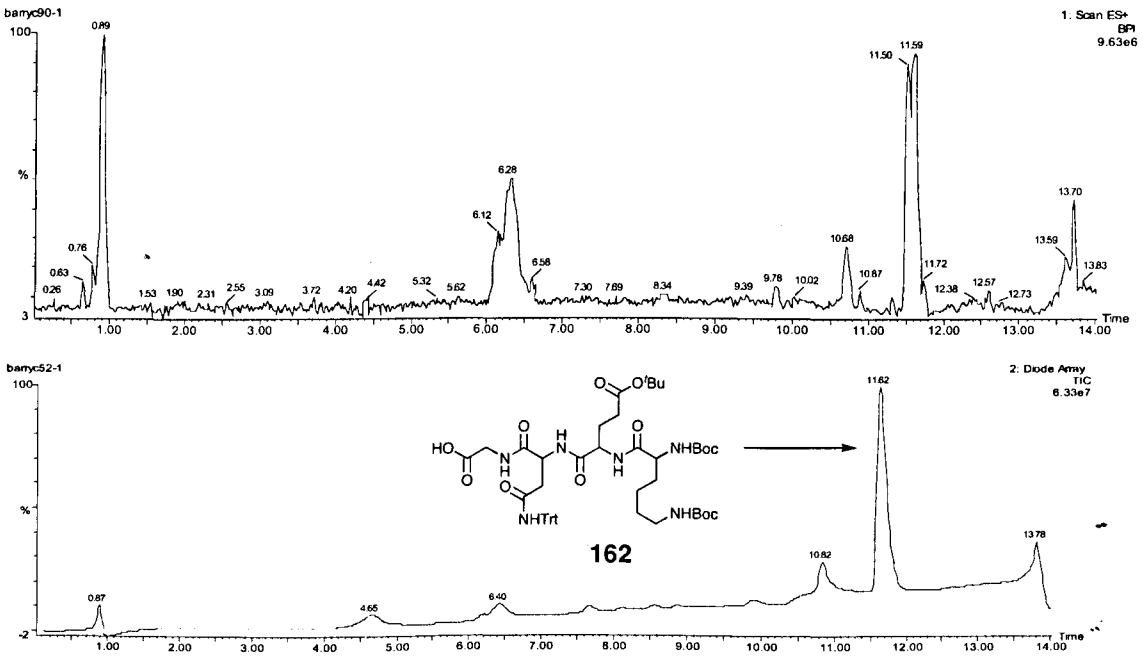


Figure 4.22 LC-MS trace of crude tetramer peptide 162

An attempt to couple peptide **162** with **147** was made using HATU and DIPEA (Scheme 4.8). HATU was used ahead of HCTU for its superior reactivity, so to maximise the success of peptide activation and coupling.

(i) GN(Trt)E(^{t}Bu)L(Boc)NHBoc (10.0 equiv.), HATU (5.0 equiv.), DIPEA (10 equiv.), DMF, RT, 1 h

(ii) 5 x DMF wash, RT, 2 min

Scheme 4.8 Coupling of alternative tetramer peptide to phenol **147**

The resin was treated with standard cleavage cocktail and analysed by LC-MS (Figure 4.23).

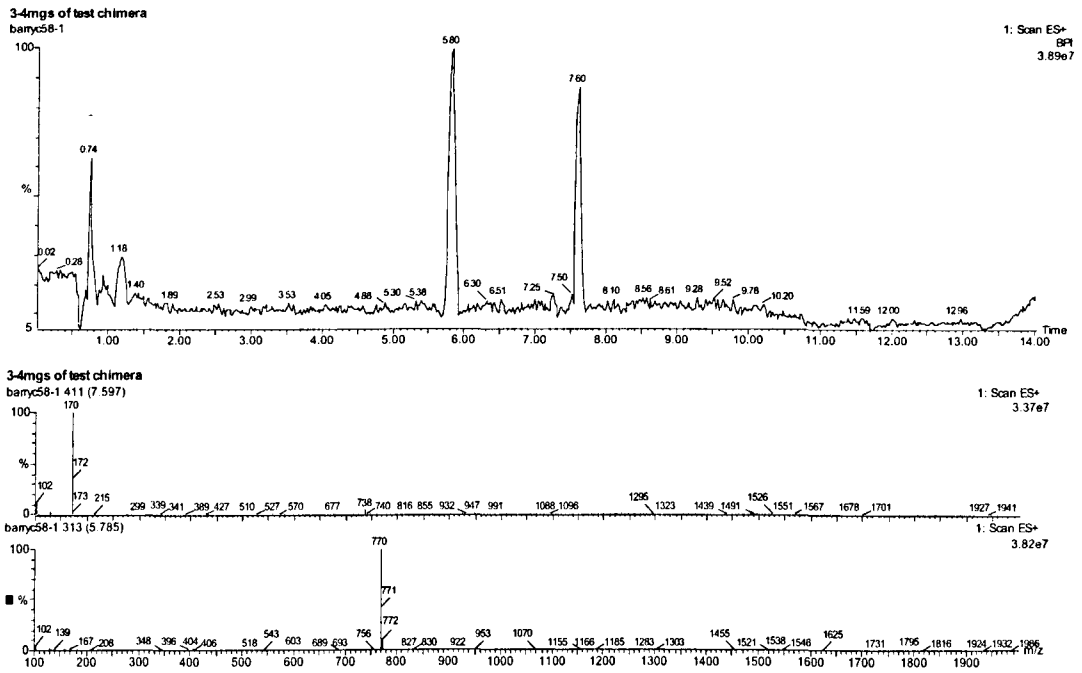
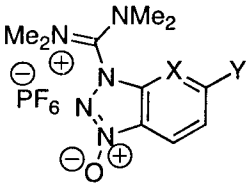
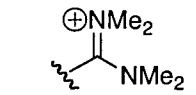


Figure 4.23 LC-MS traces of G-N-E-L peptide coupling experiment

The reaction had failed, but obtaining the same result as before with a different peptide was a telling result. This result proved that the unknown compound (of mass [769]) was not comprised in any part from the peptide component, and must therefore some other derivitive of the phenol. The difference in mass between the phenol and the mass of the observed peak is [100] and it was determined that this was the uronide cation component **163** of the coupling reagents HCTU/HATU (**164/165**, Figure 4.24). Uronide **163** was the only constant in the tetramer-peptide coupling experiments and it is thought that the product of the peptide coupling experiments is therefore cation **166**.

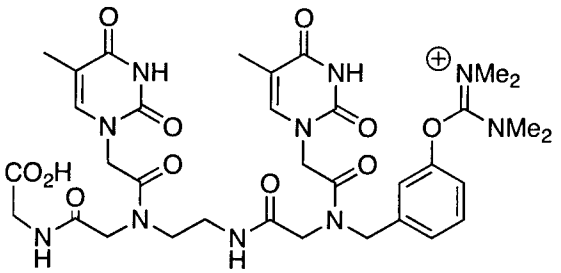


X = CH Y = Cl : HCTU **164**
X = N Y = H : HATU **165**



Mol. Weight: 100.16

163



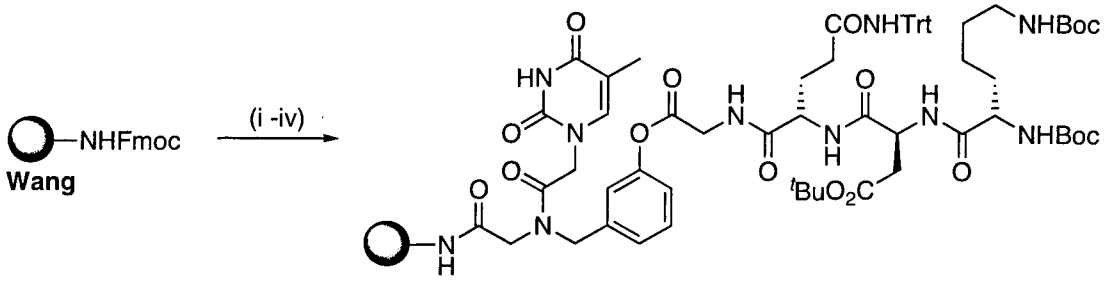
166 [Mol. Weight: 769.78]

Figure 4.24 Uronide capped phenol

This result was unexpected and did not follow prior observation wherein identical conditions had successfully coupled **147** with Fmoc glycine (without observation of uronide **166**). **166** could only be formed by reaction with either HCTU or HATU. In order for all of **147** to be capped with uronide **163**, two conditions must be true; There must have been a substantial amount of unreacted coupling reagent present in the reaction and **147** must react much faster with the coupling reagent than with the activated peptide.

Given that the coupling experiment with **147** and **162** involved a large excess of peptide, relative to the amount of coupling reagent, it was concluded that the tetramer peptides must react sluggishly (relative to single amino acids) with uronium reagents **164** and **165**. It was believed that increasing the pre-activation time of the peptides five-fold to 5 min would allow activation of the peptide to go to completion. This would remove the suspected reservoir of the uronide coupling reagent from the reaction, thus preventing formation of **166**. A quick model study was performed to test this hypothesis.

Wang resin (only available resin at that time) was loaded with linker **103** and the phenol was unmasked *via* piperidine mediated aminolysis to give **167**. Peptide **162** was preactivated with HCTU for 5 minutes and then agitated with the resin for 1 hour (Scheme 4.9).



(i) Piperidine/DMF (1:4)
(ii) **103**, HATU (4.9 equiv.), DIPEA (5.0 equiv.), DMF, RT, 1 h
(iii) Piperidine/DMF (1:4)
(iv) GN(Trt)E(^{t}Bu)L(Boc)NHBoc (5.0 equiv.), HCTU (4.9 equiv.), DIPEA (10 equiv.), DMF, RT, 1 h

Scheme 4.9 Working conditions for coupling short peptide to phenol **167**

A test cleavage of the resin with TFA/TES/water (95.0:2.5:2.5) gave a crude mixture that was analysed by LC-MS (Figure 4.25).

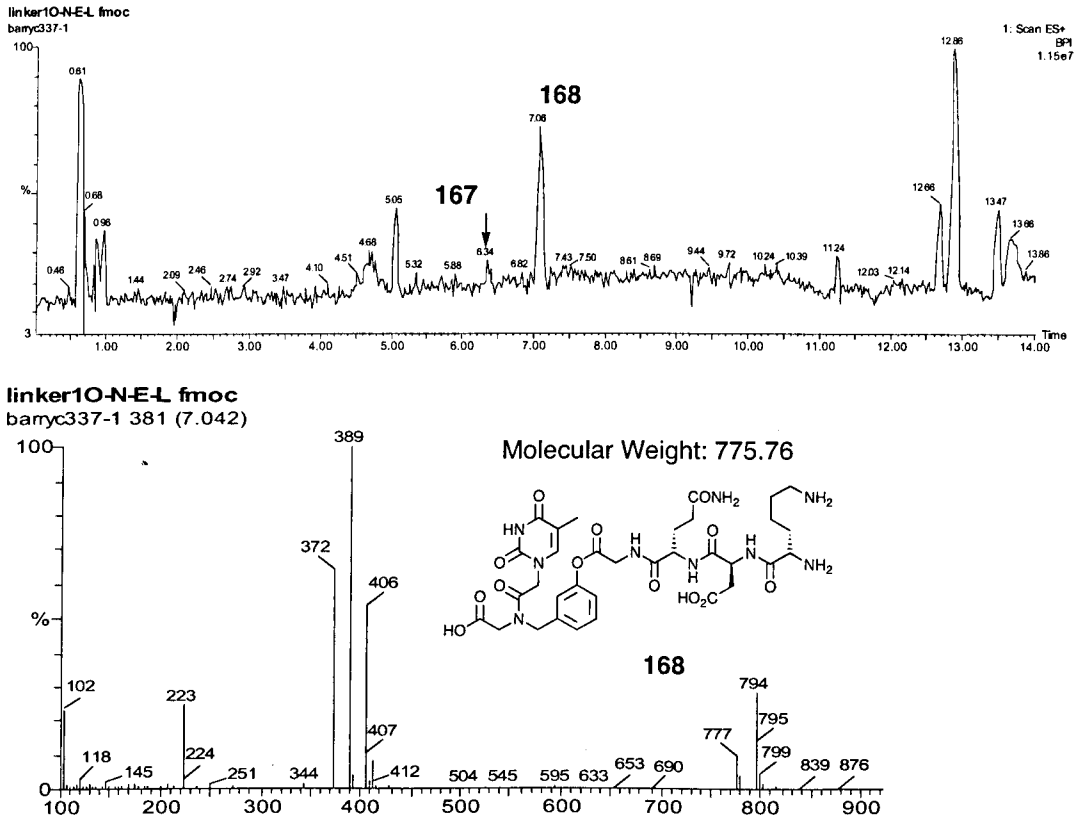


Figure 4.25 LC-MS traces of linker-pepide hybrid **168**

168 was the only product within the crude reaction mixture; the uronide capped phenol was not observed. A very small amount of the starting phenol was observed, however prolonged coupling time or double-coupling should drive the reaction to completion. This result suggests that the original hypothesis concerning activation time was correct

and that it is practical to attach polypeptides to the linker in this manner. With the phenol/peptide coupling procedures established, the final stage of the chimera synthesis was undertaken.

4.1.3 Stage 3 – PNA oligomerisation and full assembly of a PNA – peptide chimera

A simple PNA-peptide chimera that was suitable for use in test ligations was drafted for synthesis (Figure 4.26).

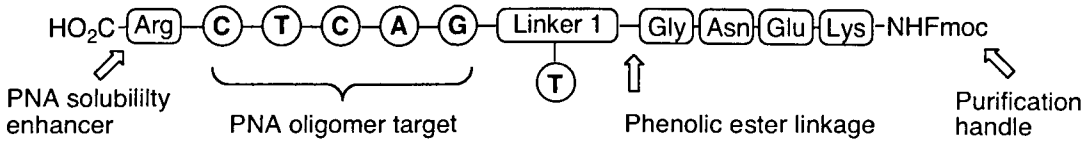


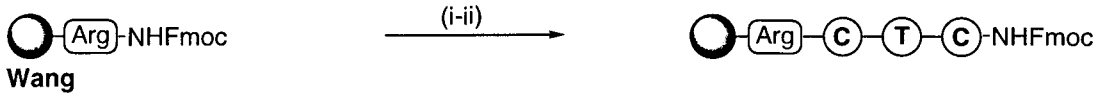
Figure 4.26 Target PNA-peptide chimera

The PNA sequence was designed with several criteria in mind. It was required that it contained all four PNA monomers / nucleobases, with the stronger G-C pairs being distributed evenly throughout the helix, to ensure that a tight helix would form with the complementary strand. A G-C pair was placed at the helix terminus to minimise the probability of helical 'unzipping' (three hydrogen bonds compared to two hydrogen bonds in an A-T pair) and there were no repeat sequences (eg T-T-T) to help prevent the formation of hairpin structures.

Ideally the linker unit would have been tagged with guanine or cytosine to give the most stable ligation site that was possible; however it was necessary to preserve the limited amount of G and C nucleobase acetic acid precursors available, and so T was chosen instead.

It was known that PNA oligomers are poorly soluble in most solvents but that simply tagging either the *N* or *C*-terminus with an amino acid possessing a hydrophilic sidechain greatly overcomes this problem.[7] It was decided to construct the PNA oligomer upon (readily available) Wang resin that was preloaded with arginine in order to accommodate this design feature. Wang resin is acid labile and readily cleaved using standard conditions of TFA/TES/H_2O (95.0:2.5:2.5).

A PNA trimer was assembled using standard SPPS conditions upon the arginine-preloaded Wang resin (Scheme 4.10).



ii) Piperidine/DMF (1:4)
(ii) PNA monomer (5.0 equiv.), HCTU (4.9 equiv.), DIPEA (10 equiv.), DMF, RT, 1 h; repeat sequence for next PNA monomer in sequence

Figure 4.10 Partial PNA oligomer synthesis

A portion of the resin was cleaved under standard conditions and analysed by LC-MS (Figure 4.28).

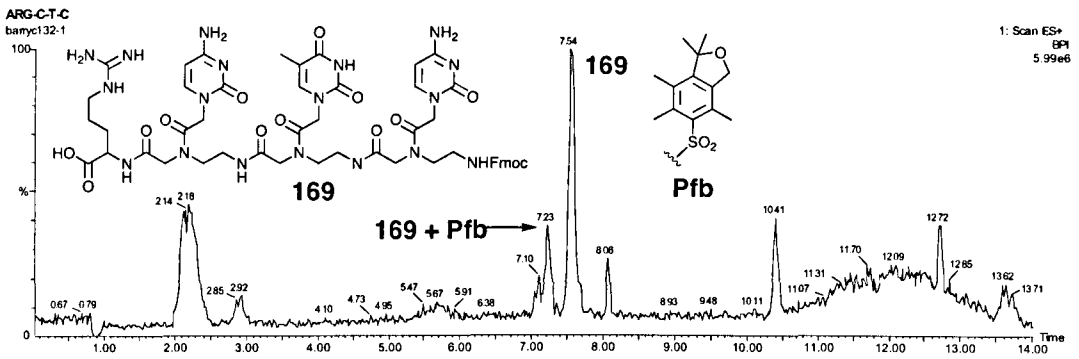
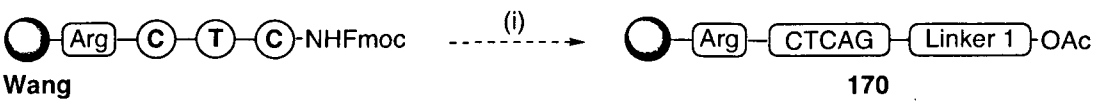


Figure 4.28 LC-MS trace of PNA trimer hybrid Arg-C-T-C

The main peak at 7.54 min contained a large number of mass ions, however fully deprotected PNA trimer was present in the mix of mass ions detected, with the $[M+H]^+$ and $[M+2H]^{2+}$ ions being observed. The smaller peak at 7.23 min is comprised of what appears to be the $[M+Pfb+H]^+$, $[M+Pfb+2H]^{2+}$ and $[M+Pfb+3H]^{3+}$ ions of **169**. Pfb is relatively robust to acidic conditions and requires prolonged cleavage times to affect full cleavage and so its presence was not surprising. It was decided that the result had demonstrated successful trimer synthesis. The sequence for the PNA oligomer synthesis was completed and the linker unit coupling was attempted (Scheme 4.11).



(i) 5 x Piperidine/DMF (1 : 4), RT, 2 min then 5 x DMF , RT, 2 min
(ii) PNA monomer or linker (5 equiv.), HATU (4.9 equiv.), PNA base solution (3 μL per mg of monomer), RT, 1 h
(iii) 5 x DMF , RT, 2 min

Scheme 4.11 Partial chimera synthesis

A portion of the resin **170** was treated with formal cleavage conditions (TFA/TES, 95:5), RT, 1 h) ensuring complete deprotection of the Arg/PNA side chains. The resin filtrate was analysed by LC-MS (Figure 4.29).

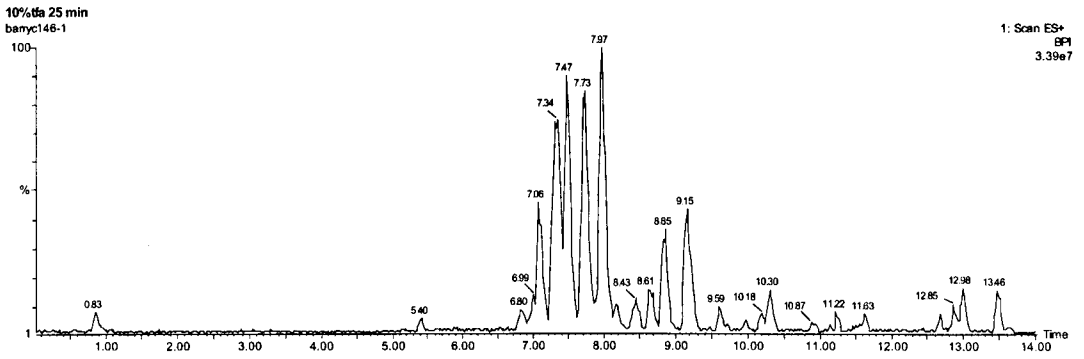


Figure 4.29 LC-MS trace of partially completed PNA to peptide chimera

The synthesis had failed, giving an array of effectively inseparable compounds. Each peak contained many strong mass ion signals and the desired chimera was not present amongst them.

To account for the many sideproducts, it was thought that the PNA monomers must contain contaminants. HP-LC analysis of the PNA monomers revealed that small amounts of PNA monomer starting materials were present in each of the respective A, C and G-monomers. T-PNA monomer was found to be pure. All impurities were removed from each PNA monomer (see Chapter 1).

There was another possible reason or contributing factor for the failure of the PNA synthesis. Christensen demonstrated that backbone-unprotected PNA can undergo *N*-terminal rearrangement at basic pH when using standard SPPS conditions (Figure 4.30).[8]

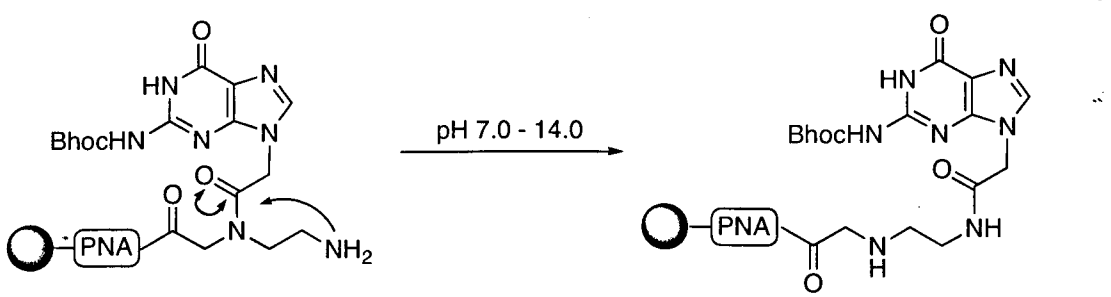
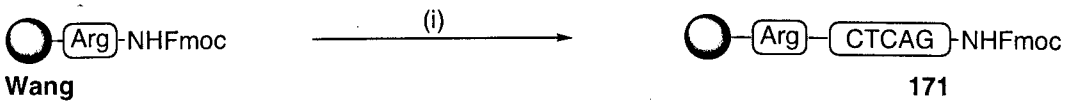


Figure 4.30 *N*-acyl transfer of deprotected PNA

Half-lives of the monomers (at pH 11) range from 1 day to several days, however the potential for generating small amounts of inseparable (structurally isomeric) failure sequence compounds was not tolerable. The optimised PNA synthesis protocol of Corey (for the manual oligomerisation of Bhoc/Fmoc PNA monomers) was chosen to replace the standard SPPS conditions that had been used in the previous failed experiment.[9] Using this method does not use a capping stage – for short PNA sequences, each coupling stage is instead driven to completion, as judged by the Kaiser test for free amines (no blue/black colour). This approach is valid as the likelihood of a failure sequence is negligible (for short PNA sequences).

PNA pentamer hybrid **171** was synthesised using the purified PNA monomers and the alternative PNA synthesis method (Scheme 4.12).



Wang **171**

(i) 1 x Piperidine/DMF (1 : 4), RT, 3 min then 2 x DMF , RT, 2 min
(ii) 1 x Piperidine/DMF (1 : 4), RT, 3 min then 6 x DMF , RT, 2 min
(ii) PNA monomer (5 equiv.), HATU (4.9 equiv.), PNA base solution (3 μL per mg of monomer), RT, 1 h
(iii) 6 x DMF , RT, 2 min

Scheme 4.12 Hybrid PNA pentamer synthesis

A portion of the resin **171** was cleaved under standard conditions and analysed by LC-MS (Figure 4.31).

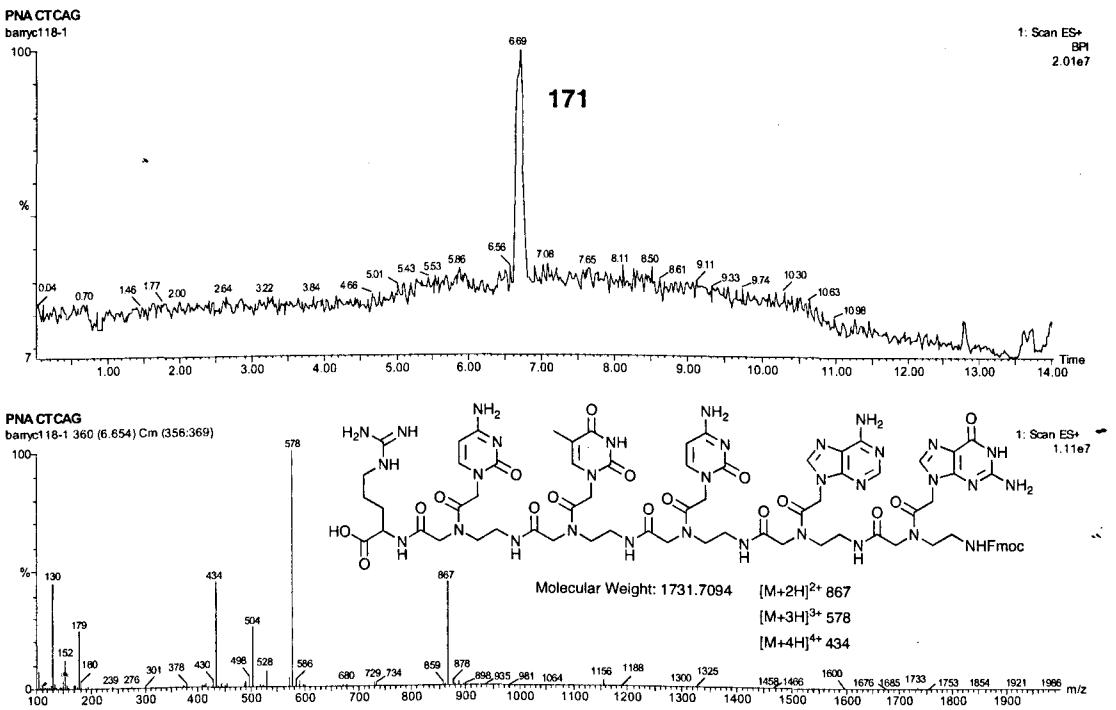


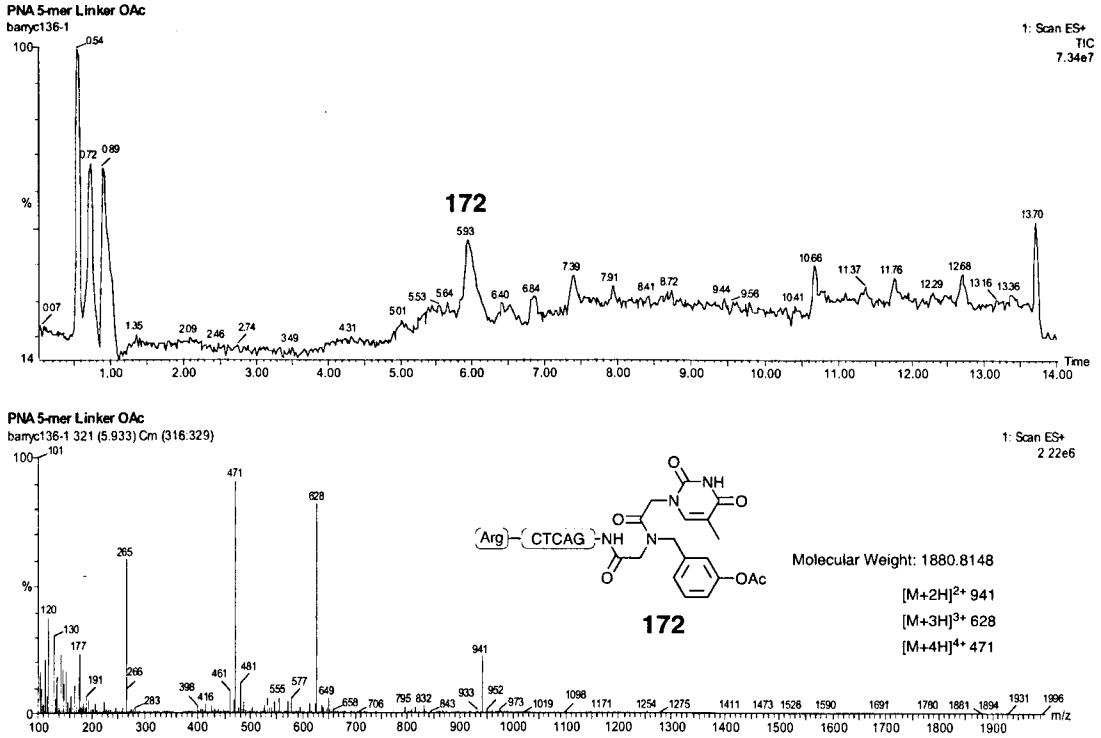
Figure 4.31 LC-MS traces of PNA oligomer CTGAG

The measures introduced to improve PNA synthesis had been effective.

As all of the steps required for the synthesis of a PNA-peptide chimera had been achieved, the chimera assembly could then be undertaken using the methods and conditions developed. PNA oligomer **171** was now used in a continuing synthesis toward a PNA-peptide chimera.

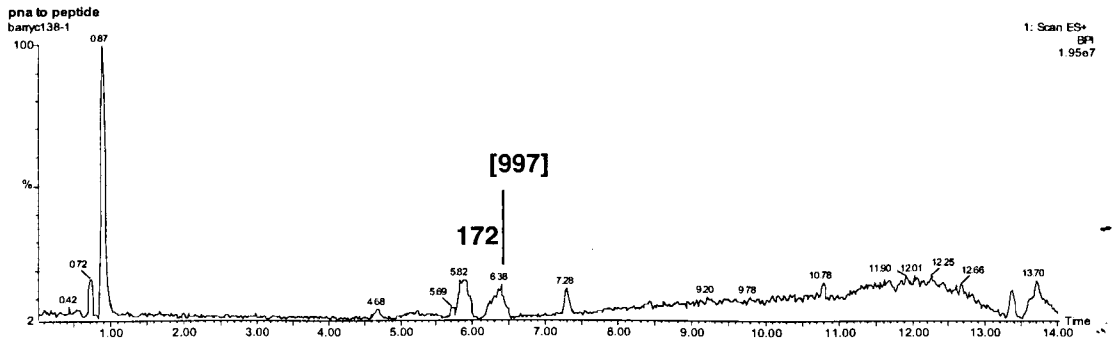
To expedite both the synthesis of the chimera and the first PNA-directed ligation reaction, it was decided that the peptide component of the chimera would be a single glycine residue. Transferring a single amino acid instead of a large polypeptide was a valid experiment for the purpose of a test ligation reaction. Using Fmocglycine as the transfer component also had strategic benefits for maximising success of the ligation reaction. Using the smallest amino acid available was thought to overcome any steric clashing at the ligation junction. Furthermore, a long polypeptide is less desirable (for the initial validation chemistry) because it could behave in a manner that would work against the ligation reaction e.g. obstructing the site of ligation or precipitating from solution. If the ligation was proven to not work with this simple chimera then it would be known that the linker architecture was most likely to be at fault and required redesign. From a practical viewpoint, transfer of Fmocglycine would be readily observed by the analytical techniques used to follow and characterise the reaction; It has a substantial mass [298] and a very UV active chromophore (Fmoc). It was decided that only when working conditions for PNA-directed ligation had been established that the scope of the work would expand toward larger chimera targets.

To continue toward the completion of the chimera synthesis, PNA oligomer **171** was deprotected and thymine-containing linker **103** was appended onto the chain (Figure 4.32).



Scheme 4.32 LC-MS of PNA-linker hybrid

Deprotection of the phenol was then carried out using the Seeberger conditions for the prescribed 2 h.[1] Piperidine was not employed for acetate deprotection as it was found to introduce impurities (as assessed by LC-MS analysis) not observed with methoxide mediated deprotection. The resin was washed thoroughly with DMF and a small portion was cleaved and analysed by LC-MS (Figure 4.33).



Scheme 4.33 LC-MS of partially deprotected PNA-linker hybrid

In contrast to the earlier test studies, methoxide-mediated cleavage of **172** had not been completed within 2 h, evident by the presence of a significant quantity of acetate **172**. The signal at 6.38 min could not be identified, although it was thought that it could be related to the chimera, given that its mass lies inbetween that of the phenol and the

target chimera. The resin was treated again with the phenol ester cleavage conditions for a longer period of 12 h. The result was not expected or favourable (Figure 4.34).

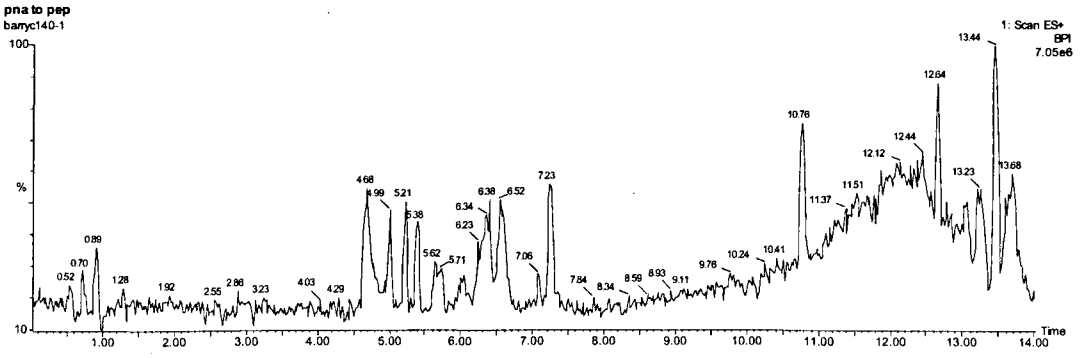


Figure 4.34 LC-MS of extended methoxide cleavage experiment

The expected phenol was not observed within the trace. It could not be determined if the proliferation of signals occurred because the hybrid had degraded under prolonged exposure to methoxide or if the sample had been contaminated. The PNA oligomer was resynthesised with freshly prepared reagents and the linker was appended to give the acetate **172** (Figure 4.35).

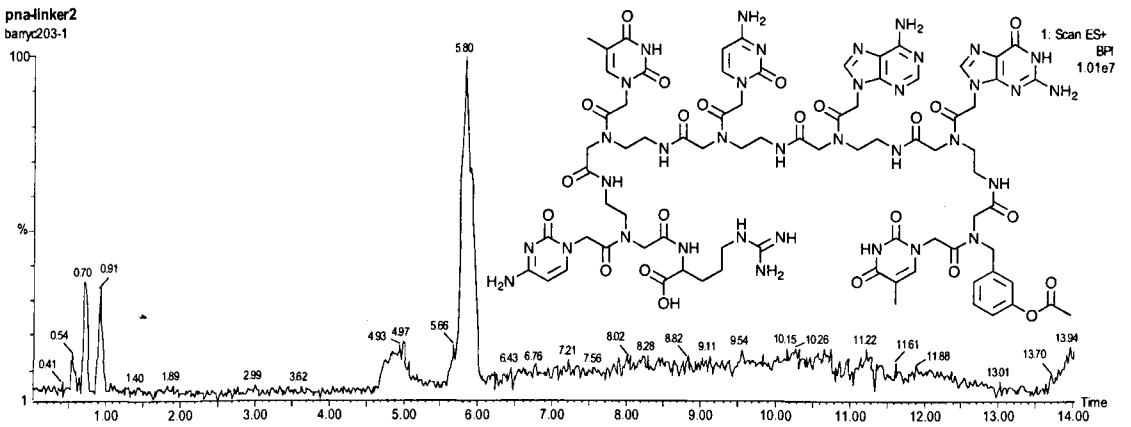


Figure 4.35 LC-MS of the PNA linker hybrid

Deprotection of the phenol was again attempted using the Seeberger conditions and an analytical portion of the resin was cleaved and then examined *via* LC-MS analysis (Figure 4.36).

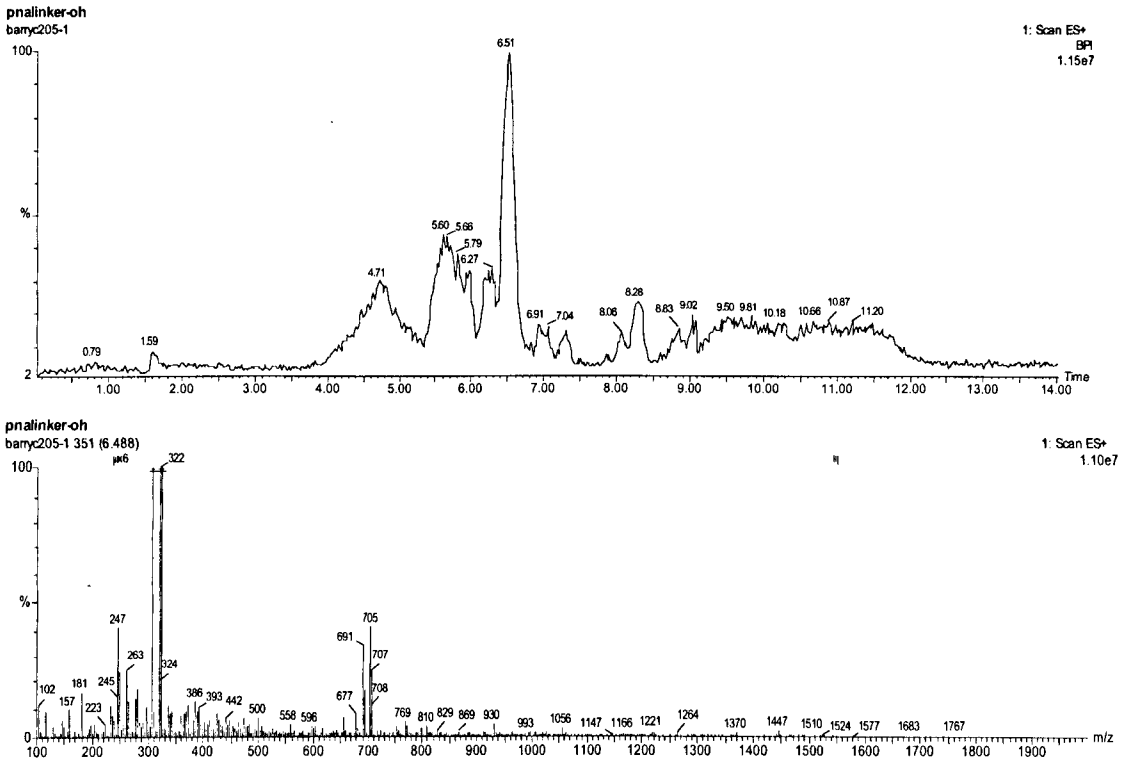
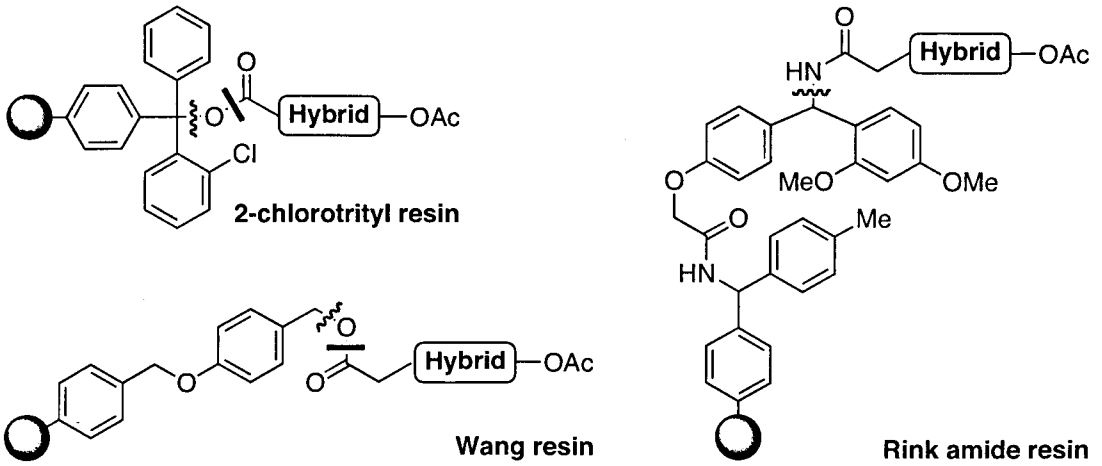


Figure 4.36 LC-MS trace of methoxide-mediated acetate cleavage experiment

The phenol was again not observed within the LC-MS trace. It was hypothesised that the sodium methoxide was cleaving the PNA hybrid from the solid support. Chimera 2 is attached to Wang resin *via* a benzylic ester linkage (Figure 4.37). Whereas good LC-MS data of the **171** and **172** was easily and reliably obtained, the phenol had proven very difficult to observe. Initially it was thought that the phenol was not easily ionised. However it was now thought that there had been insufficient material within the cleavage sample for adequate detection of the phenol by the LC-MS instrument.

Stability of the hybrids linkage to the resin had not been considered during the switch from 2-chlorotrityl to Wang resin. It had been assumed that the (generally) more chemically robust Wang resin would behave similarly to 2-chlorotrityl resin (on which the acetate deprotection reaction had been proven successful) under the conditions used. However, 2-chlorotrityl resin is known to suppress DKP formation, whereas Wang resin does not.[3] It is possible therefore that 2-chlorotrityl resin is more resistant to nucleophilic cleavage by methoxide than Wang resin. To test if methoxide was cleaving the product from the Wang resin, the target chimera needed to be synthesised and deprotected on Rink amide resin. Rink amide resin is not susceptible to nucleophilic

cleavage conditions and was predicted to retain the phenol during acetate cleavage with methoxide (Figure 4.37).



Acid cleavage sites shown as red wavy lines;
nucleophilic cleavage sites shown as bold red lines

Figure 4.37 Illustration of how the hybrids are linked to different solid supports

Rink amide resin was not available in preloaded form and it was felt that it was not necessary to append an amino acid onto the chimera *C*-terminus. PNA pentamer CTCAG was therefore synthesised upon Rink amide resin and linker **144** was appended in the usual manner. The acetate-protected hybrid was treated with the sodium methoxide deprotection conditions and a portion of the resin was cleaved and analysed (Figure 4.38).

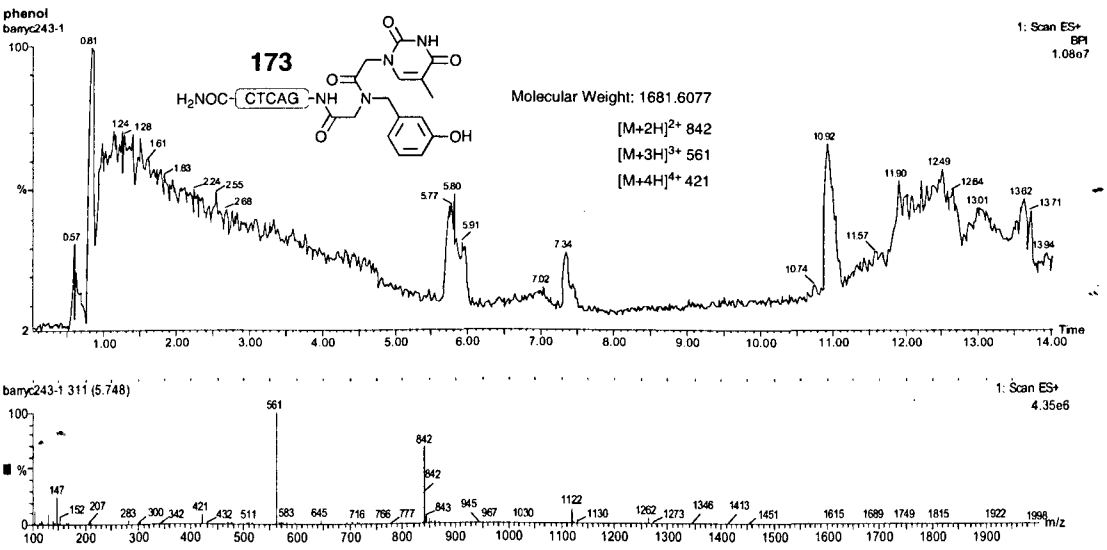


Figure 4.38 LC-MS data of hybrid phenol **173**

173 was detected as the only major compound within the crude cleavage filtrate, proving that the Wang resin had been cleaved under the methoxide cleavage conditions. **173** was then coupled with Fmoc glycine under standard conditions and the resin was shrunk, dried and cleaved with a standard cleavage cocktail (TFA/triethylsilane/water, 95.0:2.5:2.5). The resin filtrates were triturated from cold ether and purified by preparative HPLC to recover 15.4 mg (53%) of **174** (Figures 4.39, 4.40).

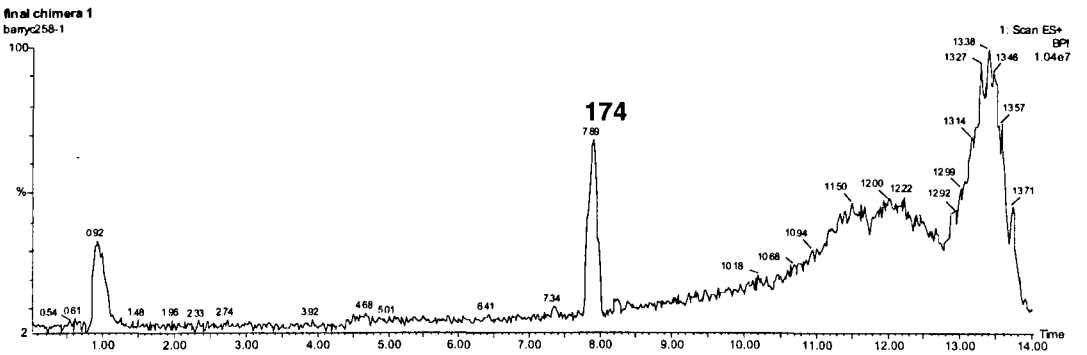


Figure 4.39 LC-MS trace of purified chimera **174**

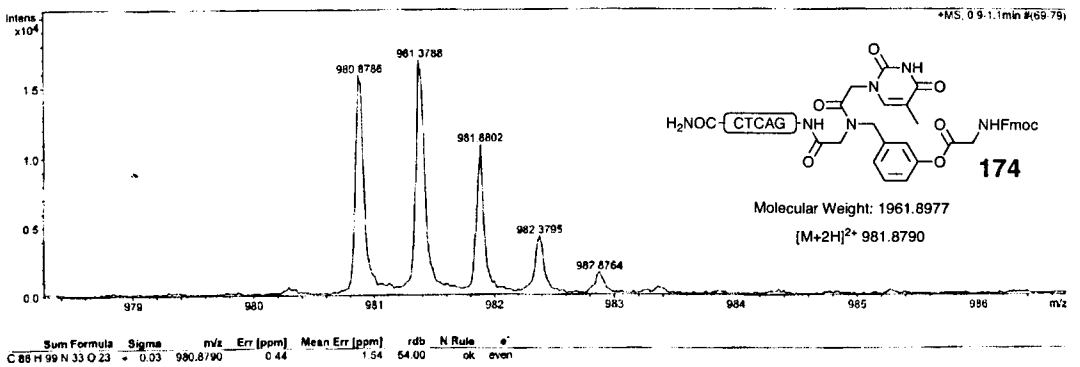


Figure 4.40 Hi-resolution mass spectrum of chimera **174**

There was insufficient time to prepare a full-sized chimera bearing a large polypeptide however the chemistry to achieve this has been established. Chimera **174** had been completed and could be used in PNA-directed ligation experiments.

4.2 Peptide to PNA chimera synthesis

Similar to the synthesis of the PNA to peptide (Figure 4.41) the main challenge in constructing this chimera was the development of conditions to incorporate the linker unit **144** into the chimera.

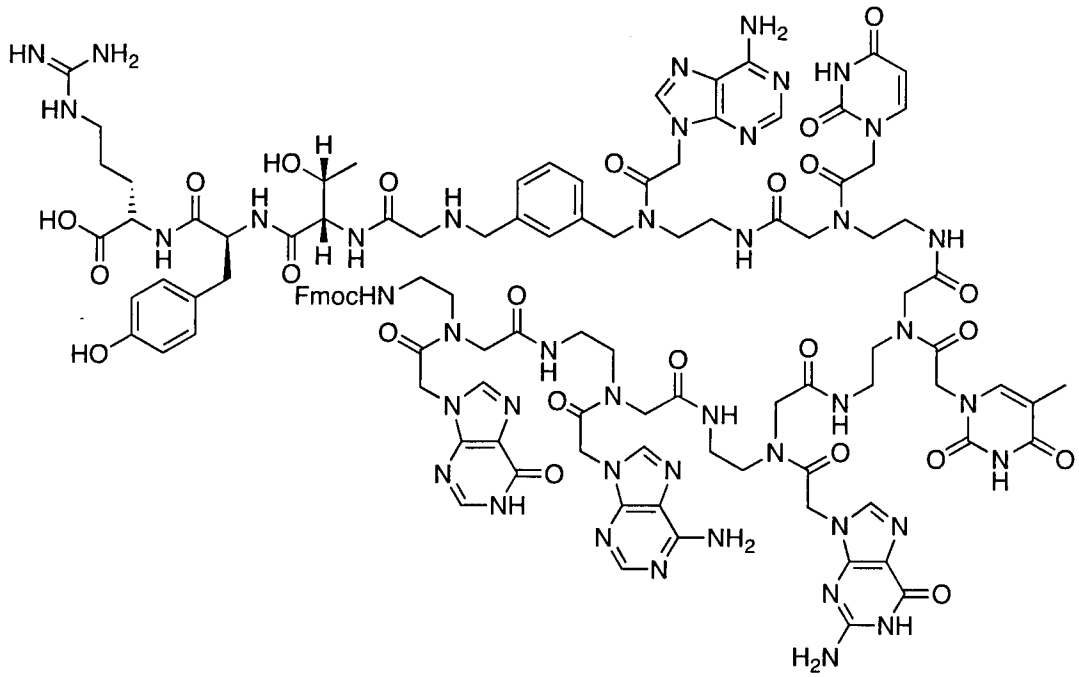


Figure 4.41 Peptide to PNA chimera target

There was no synthetic advantage to performing model studies for the attachment and deprotection of **144** because the preceding polypeptide to which the linker and PNA are appended was synthesised from (readily available) inexpensive materials and reagents. It was decided therefore that conditions for introducing the linker could be developed while trying to synthesise the full chimera target.

4.2.1 Chimera design

Design and synthesis of the chimera was constrained by three design features. The PNA sequence (by default) had to be complementary to its partner strand in **174** [other chimera]; Linker **144** had to be attached to the peptide *via* a glycine residue to minimise steric congestion about the ligation junction; Chemistry utilised in assembling the chimera had to be compatible with PNA, amino acids and the solid support (stable point of attachment).

Excepting the presence of the *N*-terminal glycine, design of the short peptide sequence was mostly arbitrary. It was convenient however to choose arginine as the first residue of the sequence given that a supply of Wang resin preloaded with arginine was readily available. The size of the peptide was purposefully kept short to preclude physical complications that could arise from using longer peptides e.g. aggregation (Figure 4.42)

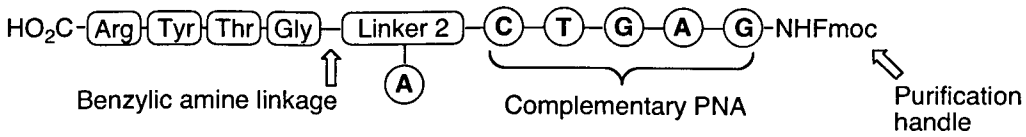
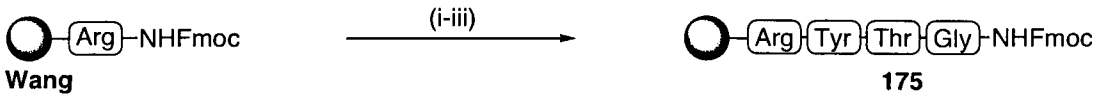


Figure 4.42 Schematic representaiton of target peptide-PNA chimera

Tetramer peptide **145** was synthesised upon Wang resin preloaded with arginine, aided by the qualitative Kaiser test (Scheme 4.13).



Wang **175**

(i) 5 x Piperidine/DMF (1:4), RT, 2 min then 5 x DMF, RT, 2 min
(ii) Amino acid (5.0 equiv.), HCTU (4.9 equiv.), DIPEA (10 Equiv.), DMF, RT, 1 h
(iii) 5 x DMF wash, RT, 2 min then repeat sequence

Scheme 4.13 Synthesis of peptide quadramer **175**

Double coupling of the threonine residue was required as indicated by a failed Kaiser test; the second coupling returned a negative Kaiser test. Cleavage and LC-Ms analysis of an analytical portion of resin showed that the peptide had been synthesised successfully without side-products. With the first component of the chimera in hand, the introduction of the linker was undertaken.

4.2.2 Attachment of linker 2

It was intended that linker **144** would be appended to the peptide using reductive amination conditions developed within the group by Nagalingam.[6] This chemistry is optimised for solid phase to overcome the problem of double-reductive amination of the starting amine that would otherwise create failure sequences and dramatically lower final product yield; double-reductive amination is the sequential condensation and reduction of two equivalents of aldehyde with one equivalent of a primary amine.

Nagalingam reported two different sets of conditions for the solid-supported reductive amination of peptides with aryl aldehydes, however both utilise an amount of AcOH. The stability of the PNA/linker sidechain protecting group (Bhoc) was known to be very sensitive to acid and it was unknown if it would survive these reductive amination conditions. To determine the stability of **144** under these conditions, a mixture of **140** and **144** (recovered from chromatographic purification of **144**) was agitated in 0.5M solution of AcOH dissolved in CH_2Cl_2 for 72 h and then analysed to determine the extent of any side-chain deprotection (Figure 4.43).

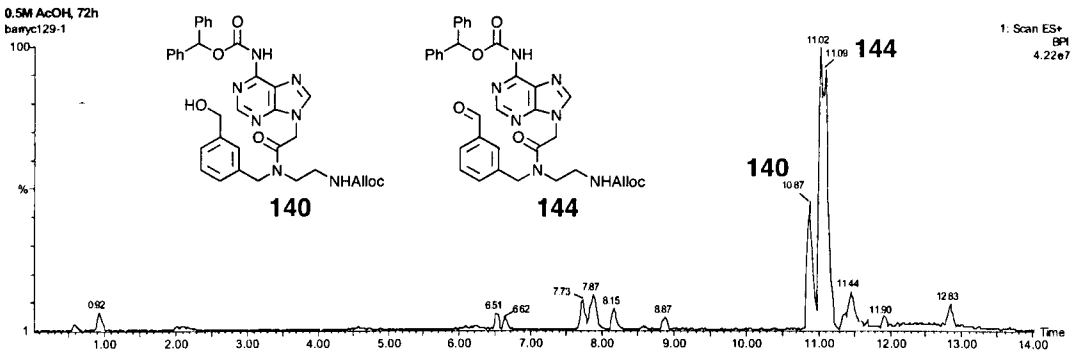
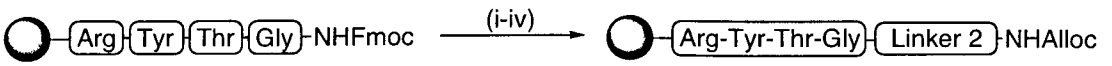


Figure 4.43 LC-MS trace of crude linker **144** after prolonged exposure to 0.5M acetic acid

0.5M AcOH solution was the greatest concentration of AcOH utilised by Nagalingam and did not deprotect **140** or **144** to any detectable degree. With the stability of **144** in the acidic reaction conditions assured, the reductive amination was attempted (Scheme 4.14).



(i) 5 x Piperidine/DMF (1:4), RT, 2 min then 5 x DMF, RT, 2 min
(ii) Linker **144** (1.1 equiv.), CH_2Cl_2, RT, 1 h
(iii) AcOH (5.0 equiv.), $NaBH(OAc)_3$ (2.0 equiv.), CH_2Cl_2, RT, 3 h
(iv) 5 x CH_2Cl_2 wash, RT, 2 min

Scheme 4.14 Reductive amination for attachment of linker **144** to a solid supported peptide

In this method the aldehyde was condensed with the resin bound amine to preform an aryl imine; addition of AcOH generated an iminium cation, the activated species that undergoes reduction with sodium triacetoxyborohydride.

Analysis of filtrate from an aliquot of resin cleaved after 3 h (neat TFA, 5 min) indicated that the reduction was sluggish. The reaction was 20% complete after 3 h and did not appear to be evolving further. It was decided to add a further two equivalents of the hydride and agitate with the resin for another 3 h. Reduction was then gauged to be approximately 50% complete after the additional 3 h. There was no evidence that the addition of the excess reducing agent or prolonging of the reaction time had generated side products – it was therefore decided to force the reaction to completion. An additional eight equivalents of reducing agent were added to the resin mixture and the resin was agitated for a further 24 h. Analysis of a crude cleavage filtrate revealed that the peptide had been fully derivitised into the hybrid **176** (Figure 4.44)

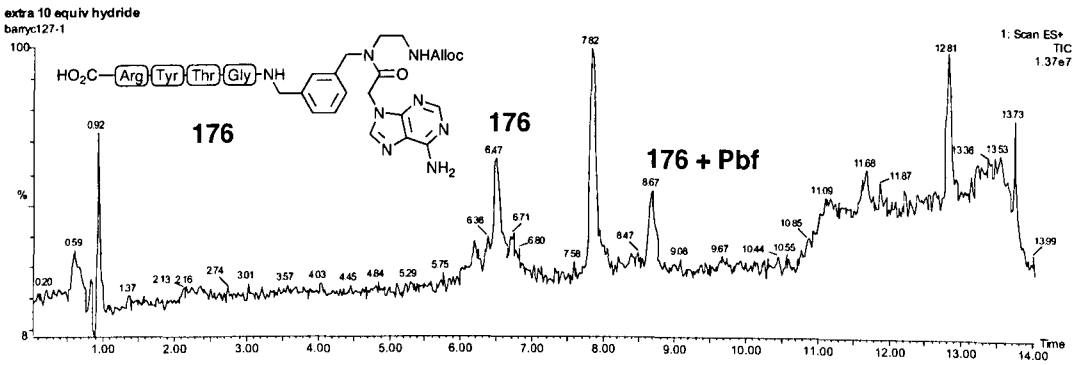
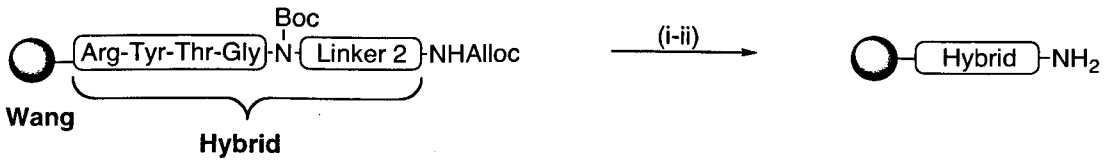


Figure 4.44 LC-MS trace of reductive amination after excess reducing agent added

176 was treated with an equi-molar solution of Boc anhydride and diisopropylethylamine in DMF to protect the glycine-linker nitrogen (protection required during PNA synthesis). It had to be assumed that the reaction was successful as it was not possible to cleave the chimera from the resin with the Boc group intact. If the amine had not been successfully protected, then it was thought that it would also not react with activated PNA monomers.

4.2.3 Hybrid linker deprotection

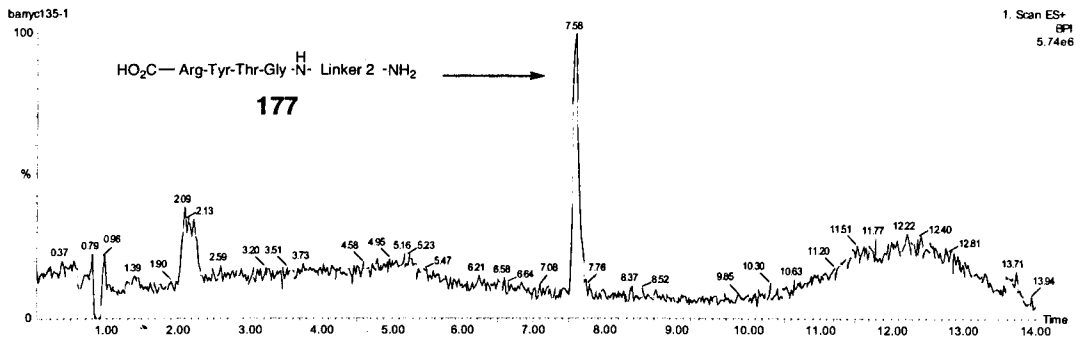
Alloc deprotection of **177** was performed using the method of Zorn.[10]



(i) 10 mol% $Pd(PPh_3)_4$, 10 mol % DABCO, CH_2Cl_2, 1h
(ii) 8 x CH_2Cl_2, 2 min, RT; 2 x TFA/CH_2Cl_2 (0.2:99.8), 1 min, RT;
5 x CH_2Cl_2, 2 min, RT; 3 x DIPEA/CH_2Cl_2 (5:95); 5 x CH_2Cl_2, 2 min, RT

Scheme 4.15 Deblocking the alloc protected hybrid

Removing the palladium catalyst from the polymer matrix was attempted using the procedure of Thieret .[11] Post reaction the resin took on a grey colour and substantially increased in density, indicating that the Palladium^0 catalyst had formed an adduct with the liberated amine. The workup had to be repeated at least twice to fully remove the palladium from the resin. It was later found that a general workup procedure used by Burt was much more effective at removing the palladium from the resin, and this procedure became the standard method.[12] Cleavage and analysis of the resin filtrate proved the deprotection to be successful (Figure 4.45).



Scheme 4.45 LC-MS trace of the hybrid after alloc deprotection

4.2.4 Attachment of PNA; completion of the chimera

The chimera was to be completed by appending with the PNA oligomer CTGAG onto **177**. The hybrid was passed through the first two cycles of the PNA oligomer synthesis, and elaborated with a C-TFmoc PNA dimer (Scheme 4.16).

(i) 1 x Piperidine/DMF (1 : 4), RT, 3 min then 2 x DMF , RT, 2 min
(ii) 1 x Piperidine/DMF (1 : 4), RT, 3 min then 6 x DMF , RT, 2 min
(iii) PNA monomer (5 equiv.), HATU (4.9 equiv.), PNA base solution (3 μL per mg of monomer), RT, 1 h
(iv) 6 x DMF , RT, 2 min; repeat (i-iv) to attach next PNA monomer

Scheme 4.16 Partial peptide to PNA chimera synthesis

Analysis of the test cleavage filtrate indicated that **178** was the only compound present.

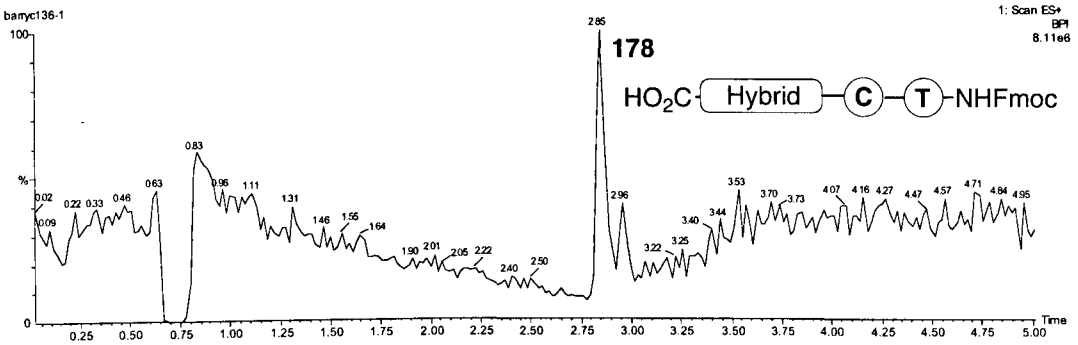


Figure 4.46 LC-MS trace of incomplete chimera **178**

The synthesis of chimera **178** was completed and the resin was dried and then cleaved with a standard cleavage cocktail (TFA/TIS/H_2O). Evaporation of the acid filtrate however did not yield a discernable amount of material and the product could not be detected by LC-MS analysis. It could not be determined why the synthesis had failed and so the chimera synthesis was undertaken again using the same methods.

Peptide **175** was synthesised as per the previous method. Linker **144** was attached to the peptide using modified conditions to overcome the sluggish reduction that had been encountered earlier; the quantity of reducing agent was increased from two equivalents to ten and the reaction time was increased from 2 to 24 hours. Analysis of the resin however showed that the desired product **176** had not been synthesised – an unknown compound (28 amu heavier than **176**, mass [944]) was the only signal observed in the LC-MS. This result could not be explained. It was decided that deviating from the original reductive amination conditions had caused a major side reaction.

Peptide **175** was synthesised again and treated with **144** and the **original** reductive amination conditions of Nagalingam. After the 3 h reaction period had elapsed, the resin

was analysed. It was found to contain a roughly equimolar mixture of the peptide and product – the side-product of mass [944] was not detected. The reaction was stirred for an additional 24 h (without further addition of hydride) and analysed again; the relative amounts of product and starting material had remained unchanged. This indicated that the active hydride in the reaction had been exhausted. An additional two equivalents of the reducing agent were therefore added and agitated with the resin for a further 12h. Analysis of the resin revealed that the mixture was now mostly product although a small amount of peptide still remained (Figure 4.47).

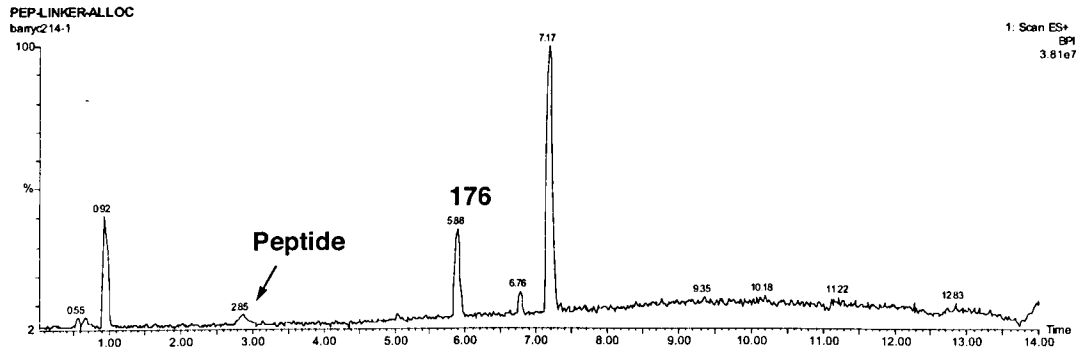
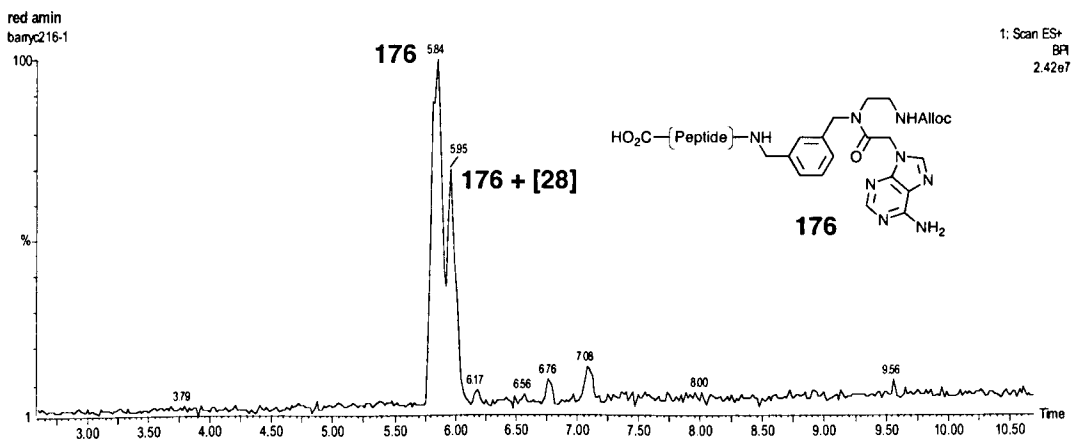


Figure 4.47 Incomplete reductive amination

With the reaction almost complete, it was decided to drive the reaction to completion with a final addition (two equivalents) of the hydride and agitate for 12 h. Cleavage and analysis of the resin proved to be illuminating (Figure 4.48).



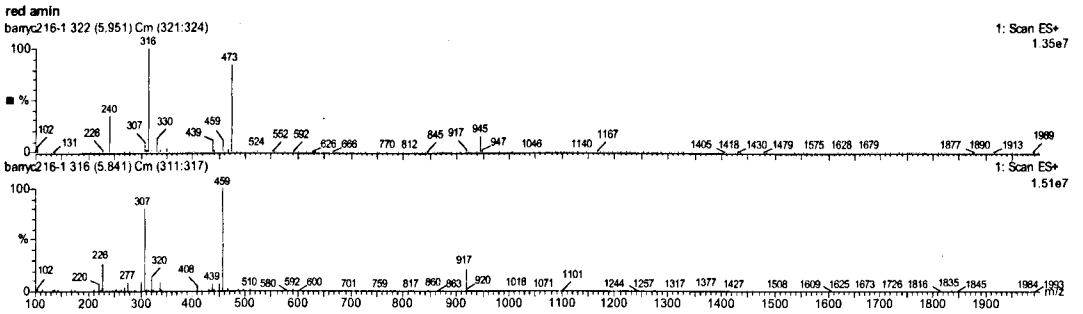


Figure 4.48 LC-MS traces of reductive amination experiment

Looking closely at the trace it was seen that the starting materials had been consumed and that there were two overlapping peaks. The major peak was the product and the smaller peak was the unknown compound of mass [944]. It appeared that following the complete consumption of the peptide, partial conversion of **176** into the mystery compound had occurred. It is thought that this process is occurring slowly relative to the reductive amination process and only in the presence of an excess of reducing agent.

176 and the unknown compound were treated with boc anhydride//DIPEA for 1 h and then the Alloc group was removed. Cleavage and analysis of a small portion of the resin was conducted (Figure 4.49).

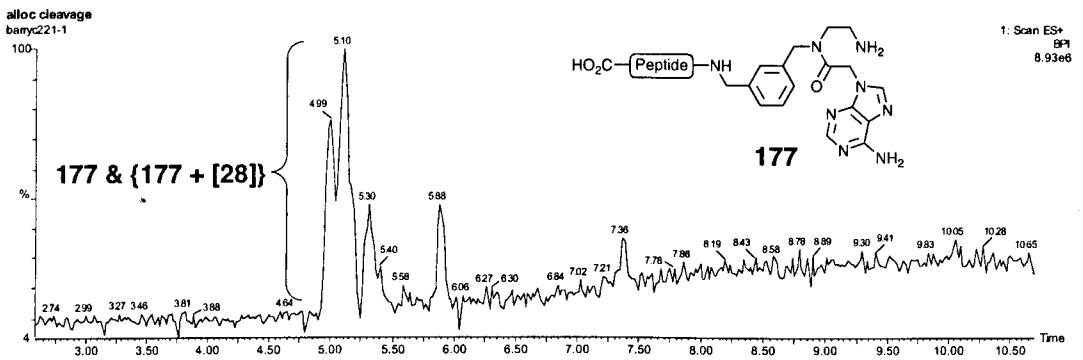


Figure 4.49 Alloc deprotection of the crude chimera mixture

Within the overlapping peaks found at 5.10 min it was observed that both the product and a compound of mass 28 amu greater than **177** were present. This data demonstrated that the Alloc group was not changed during the transformation of **176** into (**176** + **[28]**) as the side product had been proven to have had an Alloc group.

The mixture of **177** and the unidentified side product were then elaborated with the PNA oligomer **CTGAG**. Cleavage and LC-MS analysis were not conducted during PNA assembly so as to expedite the synthesis of the chimera target. The terminal Fmoc

group was removed from the hybrid prior to cleavage from the solid support to maximise its aqueous solubility, increase the potential for ionisation (LC-MS detection) and minimise any clashing that could inhibit PNA duplex formation. Analysis of the resin revealed a number of different components within the crude cleavage mixture, however **179** was observed as a strong signal. The resin was cleaved with a standard reagent cocktail of TFA/TIS/water (95.0:2.5:2.5) for 1 h at RT to give approximately 10 mg of a crude residue that was purified by semi-preparative LC-MS. Of the four fractions collected, only one fraction was found to contain chimera **179** (Figure 4.50). The remaining three fractions gave incoherent LC-MS traces.

The absolute mass of isolated **179** was less than 1 mg, an insufficient amount to properly execute a PNA directed ligation experiment.

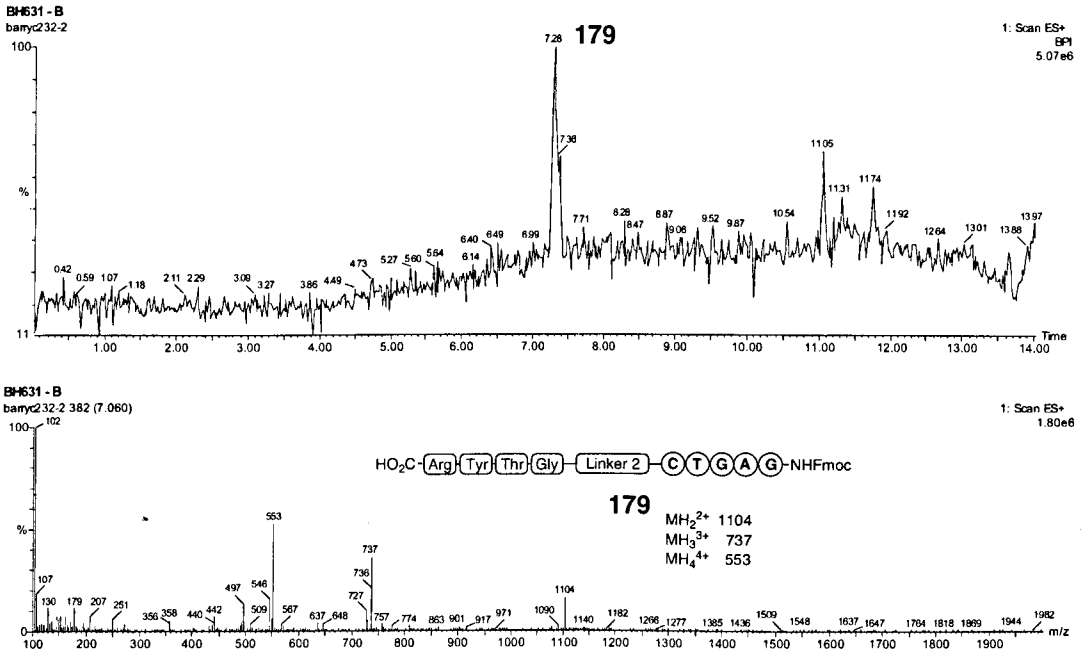


Figure 4.50 LC-MS data of chimera **179**

179 needed to be synthesised in multi-milligram quantities. The formation of the unknown side product formed during the reductive amination had to be prevented in order to achieve this goal. It had been noted that the side-product formed gradually under prolonged exposure to large amounts sodium triacetoxyborohydride added gradually; It was thought therefore that a short reaction time with a single addition of a large amount of reducing agent would promote rapid and full formation of **176**.

Peptide **175** was synthesised, deprotected and then agititated in CH_2Cl_2 in the presence of 1.1 equiv. of linker **144** for 1 h at RT. 5.0 equiv of AcOH and 10.0 equiv. of sodium triacetoxyborohydride were added to the resin and agitation was continued for 3 h. Analysis of the resin after 3 h showed the reaction to be complete without formation of the side-product **176** (Figure 5.51):

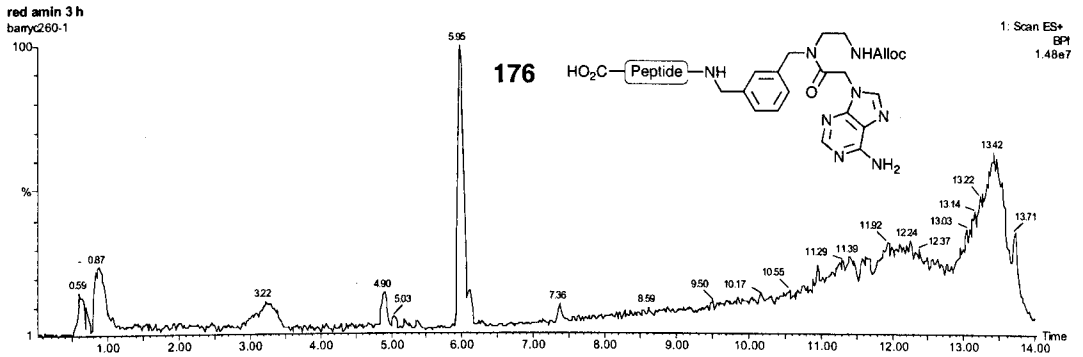


Figure 5.51 LC-MS trace of optimised reductive amination conditions

176 was treated with Boc anhydride/diisopropylethylamine in DMF to protect the naked amine as per the previous work. Deprotection of the hybrid *via* palladium-catalysed removal of the Alloc group proceeded smoothly (Figure 5.52).

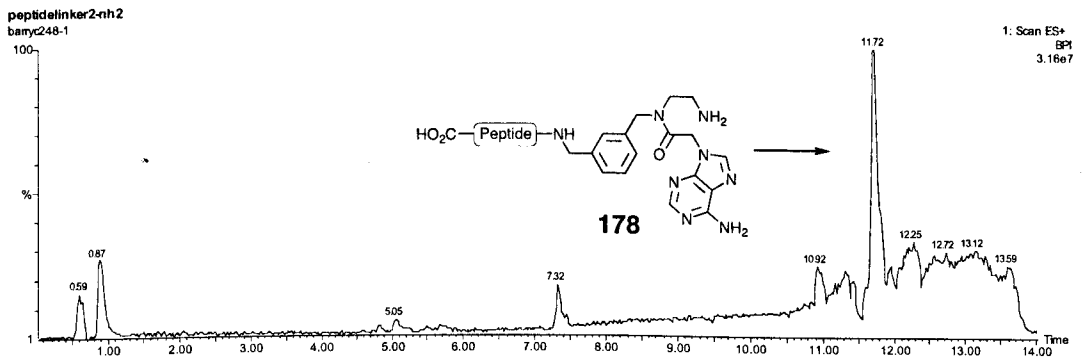


Figure 5.52 LC-MS trace of alloc-deprotected hybrid **178**

178 was coupled with a C-PNA monomer under standard conditions and the resin was analysed (Figure 5.53).

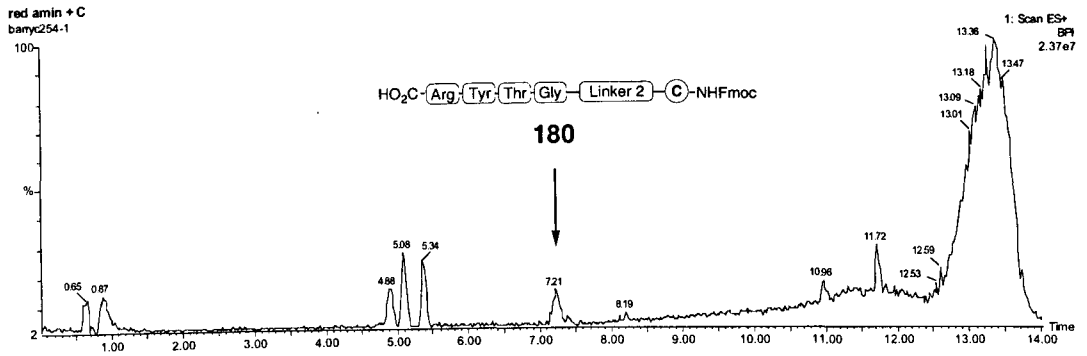
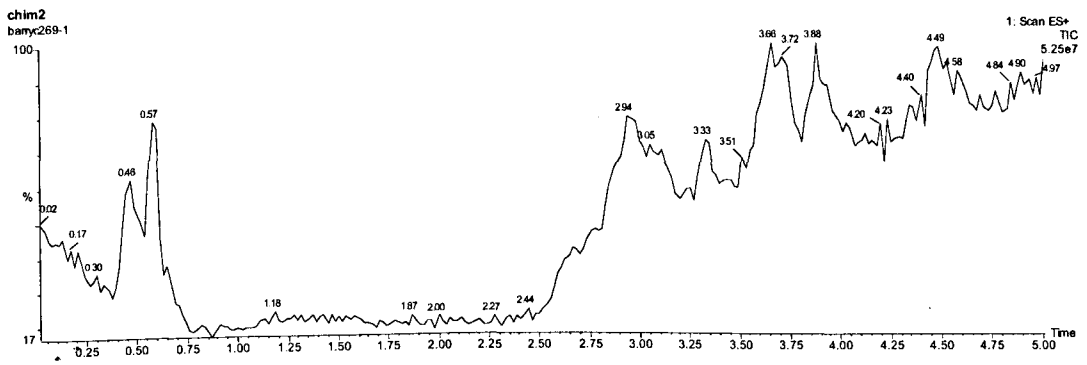


Figure 5.53 LC-MS data of partial chimera **180**

180 is observed at 7.21 min, however the triad of peaks at 4.88, 5.08 and 5.36 min could not be accounted for. It was assumed that the additional peaks were derived from general solid phase waste found in crude cleavage mixtures.

180 was carried through to the end of the synthetic sequence without further analysis and prepared for cleavage in the usual manner; it was decided to retain the Fmoc group prior to cleavage to aid purification and also maximise the mass of the product that could be recovered. The dry resin was treated with a standard cleavage cocktail of TFA/TIS/water (95.0:2.5:2.5) to afford 28 mg of crude material. In contrast to the other chimera, **181** was found to be only sparingly soluble in either water or organic solvents at RT. It was found that small amounts of the crude material would dissolve in DMF or DMSO at temperatures greater than 50°C, although slow precipitation occurred after heating was removed. Given the solubility issues, both purification and characterisation of the crude material was difficult. Analysis of a dilute sample (1 mg per mL, 0.41 μMol maximum amount) of crude **181** dissolved in DMF was carried out with LC-MS (Figure 5.54)



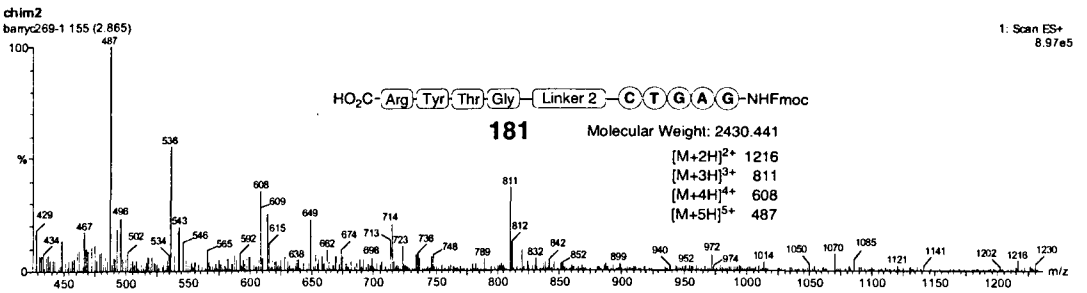


Figure 5.54 LC-MS data of chimera **181**

Although the LC-MS trace of the chimera was largely incoherent and noisy, **181** was observed within the trace. It was not known how well **181** can be detected by the LC-MS instrument. Considering chimera **181** was readily detectable, the sample should also have been detectable. To complement the LC-MS analysis, the crude sample was examined by HP-LC and Hi Resolution mass spectrometry (Figures 5.55 - 5.57).

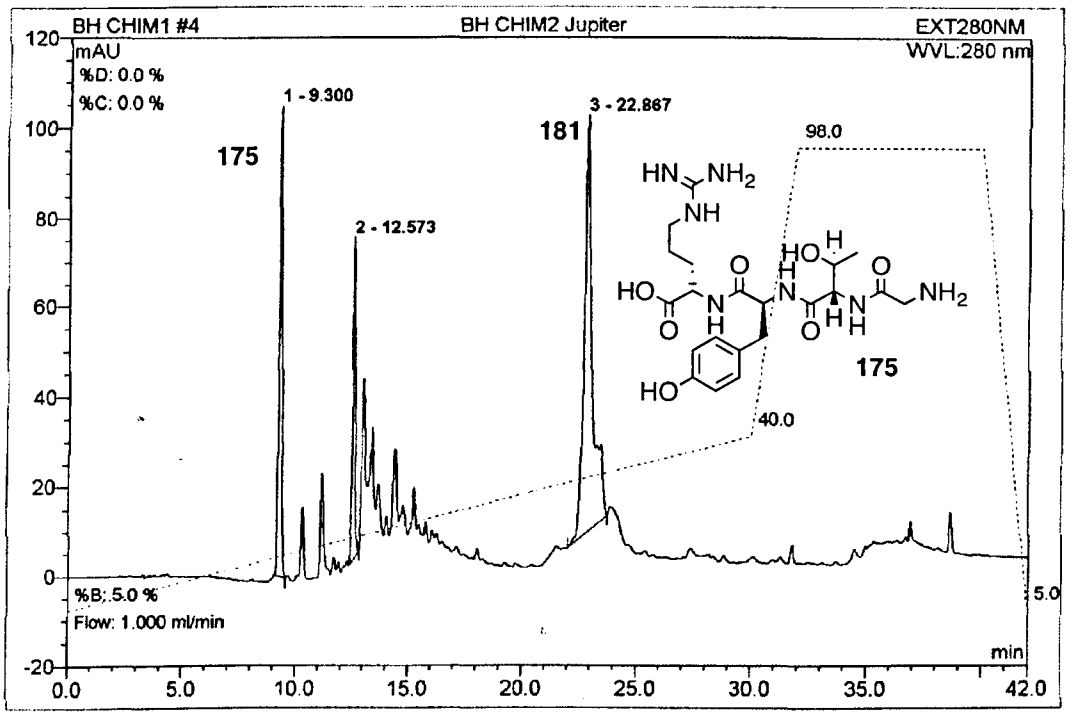


Figure 5.55 Analytical HPLC trace of crude chimera **181**.

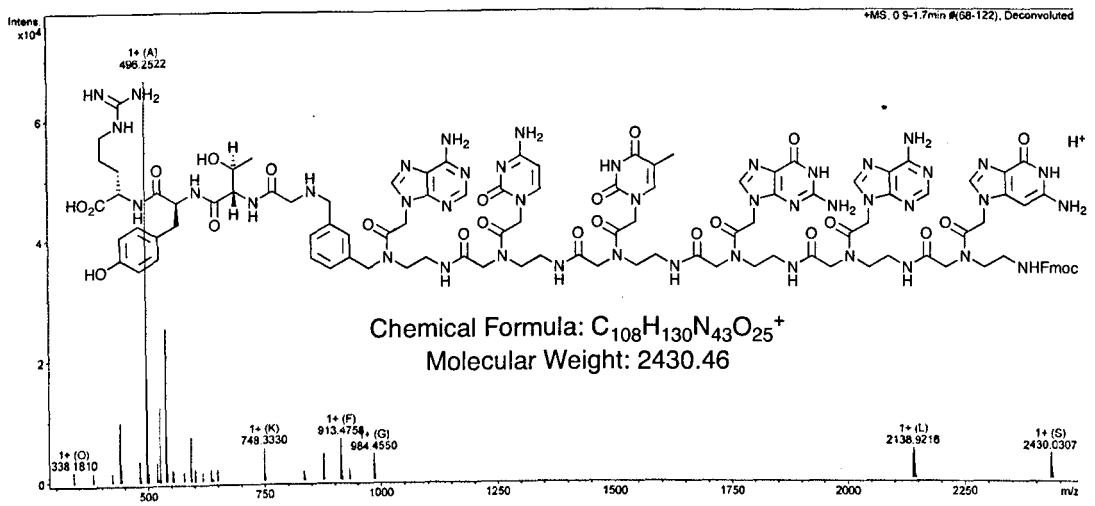


Figure 5.56 High resolution mass spectrum of crude chimera **181**

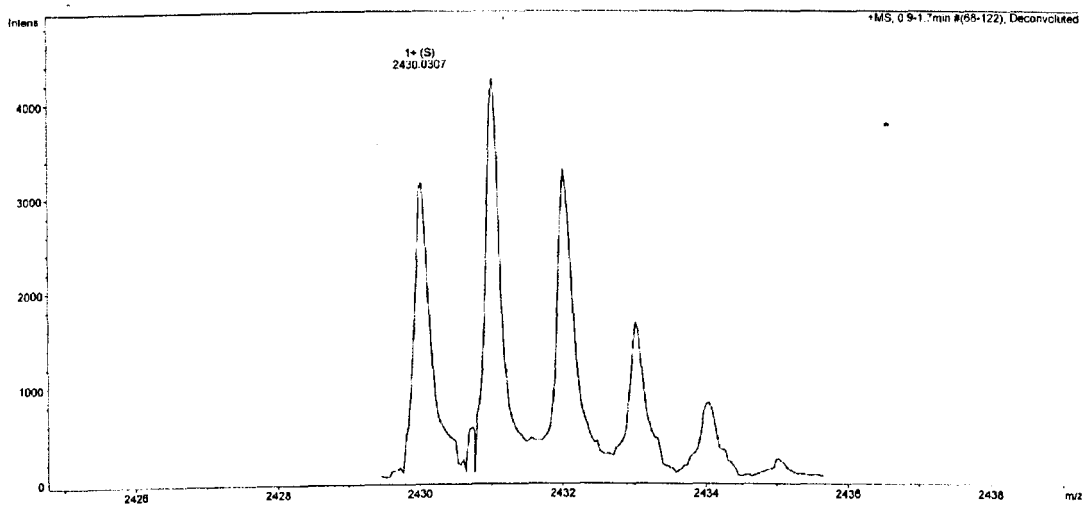


Figure 5.57 Close-up of high resolution mass-ion of crude chimera **181**.

HPLC analysis was run at a wavelength of [280 nm] which is more suitable for detecting nucleic acids and a strong signal was indeed observed at 22.86 min. The accurate mass spectrum was troubling. Although product **181** is clearly present, peptide **175** dominates the ion count by approximately two orders of magnitude relative to **181**. Peptide **175** can only have been generated by cleavage of the benzylic linker bond. It was of paramount importance to know if this cleavage occurred during ionisation of the sample or if **181** was destroyed during cleavage from the solid support. It is conceivable that Boc protection of the amine, prior to the removal of the Alloc group from the hybrid, activated the benzylic amine bond toward protolysis (Figure 5.58).

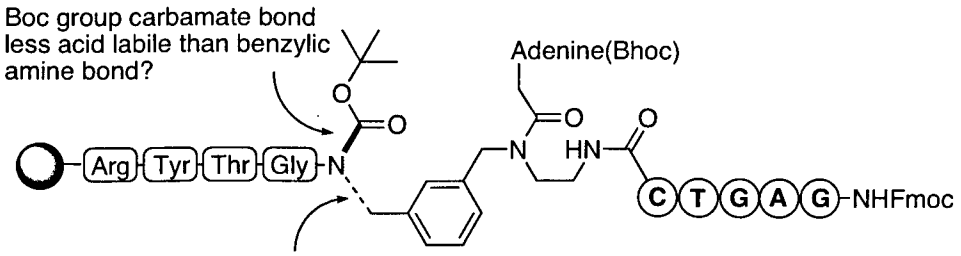


Figure 5.58 Schematic representation of chimera **181** illustrating potentially acid-labile benzylic bond

If this was true, then Boc protection of the amine is unsuitable for the chimera synthesis and a very acid labile nitrogen-protecting group is required instead. However, a silver lining to this development is that after a successful PNA-directed ligation, the post-ligation peptide-PNA hybrid would be sufficiently activated for simple protolytic cleavage of the unwanted directing group. It had been thought originally that initial ligation products would not be sufficiently labile toward cleavage of the linker/PNA hybrids and that a linker redesign would be required to accommodate this requirement of the methodology. It appeared that this would not be necessary and all that was required now was to overcome the chimera cleavage problem.

However it should be noted that the peptide was not observed during the test cleavage of intermediate **180**; it would have been expected to observe the peptide by LC-MS if indeed chimera **181** was sensitive to acid.

At this juncture, the time remaining for the project had effectively elapsed. There was neither sufficent time nor material available to resynthesise the chimera. There was now only the opportunity to perform a simple test to search for evidence that the two chimeras would interact and ligate. A solution of crude **181** (6 mg) dissolved in DMSO/DMF (1:1) 1 mL was purified by semi-preparative HPLC and fractions corresponding to the major UV-active elution peak (from the prior analytical HPLC trace) were collected. The combined fractions were evaporated to dryness to give 1.5 mg of (what was assumed to be) chimera **181**. This was estimated to be equal to approximately 0.61 μMol of **181** and was taken up in 500 μL of water/acetonitrile (1:1). This was added to a solution of 2mg (0.90 μMol) of the complementary chimera **174** dissolved in 500 μL of water/acetonitrile (1:1). The reaction mixture was neutralised with sodium hydrogen carbonate solution and stirred for 15 min. Analysis of the

mixture by LC-MS was difficult although UV activity correlated with two weak signals within the mass ion trace (Fig 5.59).

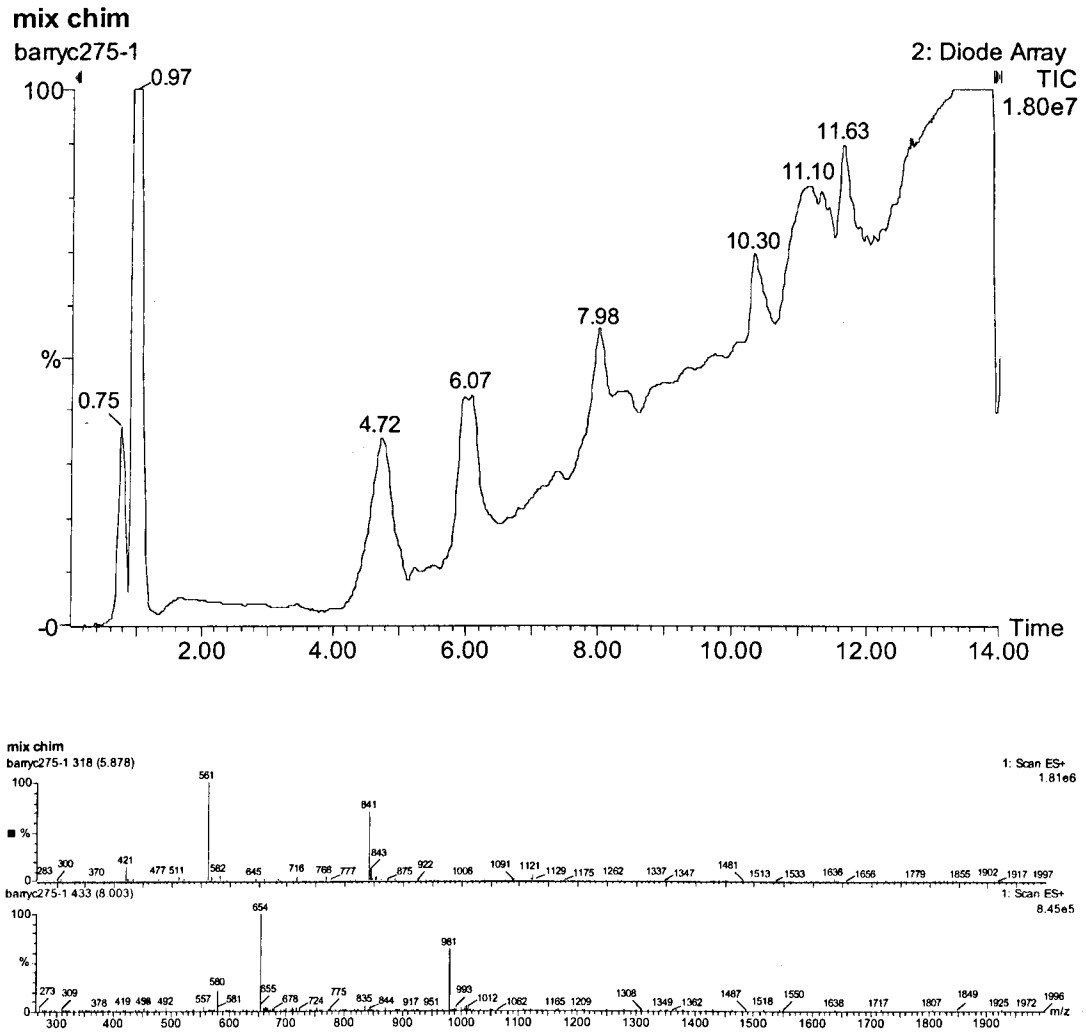


Figure 5.59 LC-MS data from a crude PNA directed ligation experiment

The signal at 7.98 min was **175** which (as the major component of the reaction) was expected to be observed. The second of the two major peaks observed in the UV-vis trace was identified as being PNA-linker phenol **173**. How **175** was hydrolysed is not clear, however it had been proven earlier that **175** was stable to saturated sodium hydrogen carbonate solution. The ligation product and **181** could not be detected; if **181** was present in the solution it would have been expected to have been annealed with **175** and this had not been observed. Overall however there was insufficient evidence to state (with absolute confidence) what precisely caused chimera **175** to dissociate. The presence of phenol **173** could be construed as tentative evidence for a PNA-directed ligation, however the LC-MS data was too weak to be able to draw any firm

conclusions. The time allotted for the project the work towards PNA directed ligation of peptides expired at this point in the project and all work was then brought to an end.

4.3 Conclusions

The existing literature for the synthesis of PNA monomers is not clear nor easily reproducible. The absence of any example in the academic literature of the synthesis of Bhoc/Fmoc PNA monomers is a strong indication of this. These procedures have been modified so that they now reliably allow access to multi-gram quantities of the four major PNA monomers.

The design and development of a set of novel linker molecules has been carried out. Difficulties associated with the synthesis of the PNA-derived linker units have been overcome and they are now accessible in multi-gram quantities.

The chemistry required for the synthesis of the very large and architecturally rich and dense chimeras has been developed. Access to multigram quantities of the PNA to peptide chimera has been achieved and analytical quantities of the peptide to PNA chimera have also been synthesised. The majority of the problems encountered with the synthesis of the peptide to PNA chimera have been overcome. A brief investigation into alternative protecting groups for the linker-peptide amine linkage should be all that is required to optimise the current synthetic route and provide multi-gram quantities of the chimera.

In summary novel methods have been developed to prepare complex conjugates of PNA and peptides which have been designed to facilitate the templated synthesis of proteins. The preparation of these molecules has required significant optimisation of some technically challenging chemistry and the development of new routes to chimeras of peptides and PNA. With the fundamental synthesis of these molecules established strong foundations have been developed for the examination of PNA directed ligation reactions.

4.4 References

[1] R. B. Andrade, O. J. Plante, L. G. Melean, P. H. Seeberger, *Org Lett.* **1999**, *1*, 1811.

[2] C. Mocquet, *University of Leeds* **2006**, *2007*.

[3] W. C. Chan, P. D. White, *Fmoc Solid Phase Peptide Synthesis*, Oxford University Press, **2000**.

[4] G. Haaiima, H. F. Hansen, L. Christensen, O. Dahl, P. E. Nielsen, *Nucleic Acids Res.* **1997**, *25*, 4639.

[5] X. Bu, G. Xie, C. W. Law, Z. Guo, *Tetrahedron Lett.* **2002**, *43*, 2419.

[6] A. Nagalingam, *University of Leeds* **2007**.

[7] B. D. Gildea, S. Casey, J. NacNeill, H. Perry-O'Keefe, D. Sorensen, J. M. Coull, *Tetrahedron Lett.* **1998**, *39*, 7255.

[8] M. Eriksson, L. Christensen, J. Schmidt, G. Haaima, L. Orgel, P. E. Nielsen, *Nouv J Chim* **1998**, *22*, 1055.

[9] D. Corey, www4.utsouthwestern.edu/coreylab/protocols/htm.

[10] C. Zorn, F. Gnad, S. Salmen, T. Herpin, O. Reiser, *Tetrahedron Lett.* **2001**, *42*, 7049.

[11] N. Thieriet, J. Alsina, E. Giralt, F. Guibe, F. Albericio, *Tetrahedron Lett.* **1997**, *38*, 7275.

[12] J. Burt, *University of Leeds* **2005**.

Experimental

Chapter 5

Chapter 5 - Experimental

5.1 General Experimental

THF was freshly distilled over sodium with benzophenone as an indicator, and DCM was distilled over calcium hydride. Ether refers to diethyl ether and petrol refers to petroleum ether (b.p. 40-60 °C). Solvents were removed under reduced pressure using a Büchi rotary evaporator and vacuubrand PC 2001 Vario pump. Samples were freed of remaining traces of solvents under high vacuum.

Flash column chromatography was carried out using silica (35-70 μm particles) according to the method of Still, Kahne and Mitra. Thin layer chromatography was carried out on commercially available pre-coated plates (Merck silica Kieselgel $60F_{254}$).

Proton and carbon NMR spectra were recorded on Bruker DPX 300 and DRX 500 Fourier transform spectrometers using an internal deuterium lock. Chemical shifts are quoted in parts per million downfield of tetramethylsilane and values of coupling constants in Hertz. Carbon NMR spectra were recorded with broad band proton decoupling and DEPT, HMQC and HMBC and pulse sequences were routinely used to aid assignment of spectra.

Melting points were determined using a Reichert hot stage apparatus and are uncorrected. Infrared spectra were recorded on a Perkin Elmer Spectrum One FT-IR spectrometer. All IR spectra for solid samples were measured neat using a diamond compression plate system; all other samples were measured as a thin film on KBr plates. Nominal mass spectrometry was performed on a Waters/ Micromass ZMD instrument using electrospray ionisation. Accurate masses were obtained by the staff of the school of chemistry on a Micromass LCT (ToF) or Bruker MicroTOF instrument using electrospray ionisation. LC-MS was performed on a Waters/ Micromass ZQ instrument fitted with a Waters XTerra C18 column (method F; MeCN:H_2O with 0.1% formic acid; 1.5 min, 0:100, 6.0 min, 30:70, 6.5 min, 95:5, isocratic to 9 min). Analytical HPLC was performed on a Dionex system and preparative HPLC on a Gilson system. Reverse phase solvents were acetonitrile and water containing 0.1% TFA.

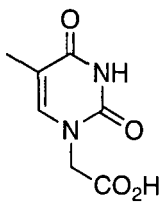
Fmoc Amino acids (Advanced ChemTech and Novabiochem) contained the following side chain protection: Arg (Pbf), Asn and Gln (Trt), Asp and Glu (*tert*-butyl ester,

OBu^t), Lys (N^{ε}-Boc), Thr (*tert*-butyl ether, Bu^t), Tyr (*tert*-butyl ether, Bu^t). Peptide elongations were performed manually using standard Fmoc solid phase peptide synthesis outlined by Chan and White using a vacuum tank attached to a water aspirator. 2-Chloro trityl resin, Wang resin and Rink Amide Novagel^{TM} resin were purchased from Novabiochem. DIC (Avocado), HOBt (Lancaster), HCTU and HATU (Novabiochem) were used without further purification. Anhydrous DMF and anhydrous acetonitrile (Aldrich) were purchased and stored in a SureSeal bottle under nitrogen. Sodium benzyloxide (Aldrich) was purchased as a 1.0M solution in toluene and stored in a sure seal under nitrogen. All other reagents were purchased from Aldrich.

^{1}H NMR splitting patterns are abbreviated as follows : (s : singlet, d : doublet, t : triplet, q : quartet, bs : broad singlet and m : multiplet). In the instance of rotomeric compounds, the compounds are assigned as either A or B where B is the major isomer.

5.2 Experimental

2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)acetic acid

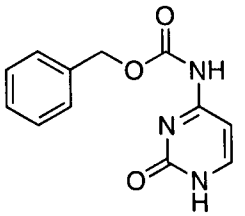


38

Methyl bromoacetate (15.2 mL, 159 mmol) was added dropwise to a vigorously stirred suspension of thymine (20.0g, 159mmol) and potassium carbonate (21.9 g, 159 mmol) in dry DMF (500 mL) at RT. The reaction was stirred for 20 h at RT and then inorganic salts were filtered from the reaction. The filtrate was evaporated to dryness and H_2O (150 mL) then 4M HCl (8 mL) were added to the crude residue and stirred for 30 min at RT. A colourless precipitate was recovered by filtration and washed with H_2O (3 x 80 mL). The solid was suspended in a mixture of H_2O (160 mL) and 2M NaOH solution (8 mL) and heated at reflux for 1 h, The reaction was cooled 0°C and acidified with 4M HCl solution (50mL) while stirring for 90min. A colourless precipitate was collected by filtration, washed with H_2O (3 x 100 mL) and dried under reduced pressure to give the title compound (15.96 g, 55%) as colourless amorphous crystals; Mp. 270-274°C

(decomp.); R_f 0.2 (EtOAc/MeOH/AcOH, 75:20:5); ^{1}H NMR (DMSO; 300 MHz): δ 7.49 (1H, s, T-6), 4.36 (2H, s, C$\underline{H}_2CO_2H$), 1.75 (3H, s, T-5CH_3); ^{13}C NMR (DMSO; 75 MHz); δ 170.0 ($CH_2\underline{C}O_2H$), 164.7 (T-4), 151.3 (T-2), 142.1 (T-6), 108.6 (T-5), 48.7 (T-1$\underline{C}H_2CO_2H$), 12.26 (T-5CH_3); *m/z* : 185 $[M+H]^+$; Data were in accordance with reported literature data.[1]

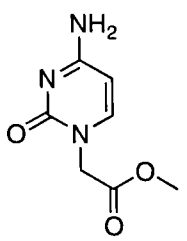
Benzyl 2-oxo-1,2-dihydropyrimidin-4-ylcarbamate



41

Benzyl chloroformate (5.20 mL, 36.0 mmol) was added dropwise to a stirred suspension of cytosine (2.0 g, 18.mmol) in pyridine (100 mL) at 0 °C. The colourless precipitate was allowed to warm to RT while stirring for 48 h. The reaction was evaporated to dryness to give a crude pink residue that was suspended in H_2O (20 mL) and acidified to pH 1 with 4M HCl solution. A light yellow precipitate was recovered by filtration and washed repeatedly with H_2O. The crude product was stirred as a suspension in EtOH (50 mL) and heated at reflux for 10 min. The title compound was recovered by filtration and dried under reduced pressure to give 1.54g (37%, lit. 56%)[1] of colourless amorphous crystals that were insoluble in common solvents; Mp. 268-271°C(decomp.); *m/z* :246 $[M+H]^+$;

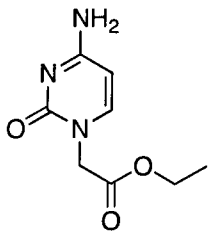
Methyl 2-(4-amino-2-oxopyrimidin-1(2H)-yl)acetate



43

Sodium hydride (1.41 g, 36.0 mmol) was added to a stirred suspension of cytosine (4.00 g, 36.0 mmol) in DMF (80 mL) at 0 °C. After 2 h methyl bromoacetate (3.41 mL, 36.0 mmol) was added dropwise to the grey slurry and the reaction was allowed to warm to RT while stirring for 48 h. The suspension was evaporated to dryness and the crude orange oil was triturated with ice-H_2O (80 mL) and stirred at 1-2 °C for 1 h. The crude product was recovered by filtration, recrystallised (50% aq. MeOH) and dried under reduced pressure to yield 3.69g (56%, lit. 64%)[2] of pink rhomboid crystals; Mp 225-6°C; R_f 0.55(EtOAc/MeOH, 1:1); ^{1}H NMR (DMSO; 300 MHz): δ 7.52 (1H, bs, C-4NH_2), 7.31 (1H, bs, C-4NH_2), 7.77 (1H, d, *J* 7.4, C-6), 5.90 (1H, d, *J* 7.4, C-5), 4.56 (2H, s, C$H_2$$CO_2CH_3$), 3.68 (3H, s, $CH_2CO_2CH_3$); ^{13}C NMR (DMSO; 75 MHz); δ 168.8 (CO_2CH_3), 163.3 (C-4), 152.3 (C-2), 148.4 (C-6), 94.1 (C-5), 52.6 ($CH_2CO_2CH_3$), 50.0 ($CH_2CO_2CH_3$); *m/z* : 184 $[M+H]^+$; Data were in accordance with literature data.[2]

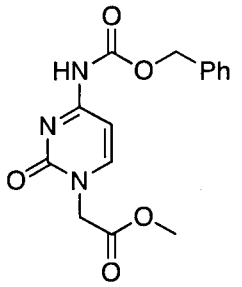
Ethyl 2-(4-amino-2-oxopyrimidin-1(2*H*)-yl)acetate



44

Sodium hydride (1.08 g, 45.0 mmol) was added to a stirred suspension of cytosine (5.00 g, 45.0 mmol) in DMF (100 mL) at 0 °C. After 2 h ethyl bromoacetate (5.00 mL, 45.0 mmol) was added dropwise to the grey slurry and the reaction was allowed to warm to RT while stirring for 14 h. The suspension was evaporated to dryness and the crude orange oil was triturated with ice-H_2O (80 mL) and stirred at 1-2 °C for 1 h. The crude product was recovered by filtration, recrystallised (50% aq. MeOH) and dried under reduced pressure to yield 3.42g (39%) of pink rhomboid crystals; Mp. 265°C (decomp.); R_f 0.60 (EtOAc/MeOH, 1:1); ^{1}H NMR (DMSO; 300 MHz): δ 7.54 (1H, d, *J* 7.2, C-6), 7.17 (1H, bs, C-4NH_2), 7.11 (1H, bs, C-4NH_2), 5.67 (1H, d, *J* 7.2, C-5), 4.41 (2H, s, C$H_2$$CO_2$Et), 4.11 (2H, q, *J* 7.1, CO_2C$H_2$$CH_3$), 1.19 (3H, t, *J* 7.1, CO_2CH_2CH_3); ^{13}C NMR (DMSO; 75 MHz); δ 169.9 ($CO_2C_2H_5$), 166.7 (C-4), 156.1 (C-2), 146.6 (C-6), 93.85 (C-5), 61.1 (CH_2CO_2Et), 50.2 ($CO_2CH_2CH_3$), 14.4 ($CO_2CH_2CH_3$); *m/z* : 198 $[M+H]^+$; Data were in accordance with literature data.[3]

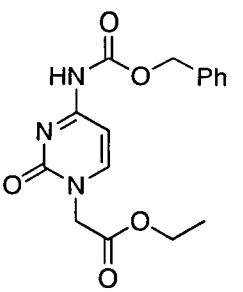
Methyl 2-(4-(benzyloxycarbonylamino)-2-oxopyrimidin-1(2*H*)-yl)acetate



45

DMAP (1.60 g, 13.1 mmol) was added to a stirred solution of benzyl chloroformate (1.96 mL, 13.1 mmol) in CH_2Cl_2 (12 mL) at 0 °C and cooled to –15 °C after a further 5 min stirring. Methyl 2-(4-amino-2-oxopyrimidin-1(2*H*)-yl)acetate (1.20g, 6.55mmol) was added slowly, portionwise and stirred for 15 min at –15°C. The reaction was allowed to warm to RT while stirring for 6 h then evaporated to dryness to give a colourless oil. The crude oil was dissolved in $CHCl_3$ (20 mL), washed once with each of 1M HCl (20 mL) and H_2O (25 mL) then evaporated to dryness. The colourless residual oil was triturated with ether to give 660 mg (32%, Lit. 77%)[3] of the title compound as colourless amorphous crystals; R_f 0.80 (EtOAc/MeOH, 1:1); Mp 137-8 °C; ^{1}H NMR (DMSO; 300 MHz): δ 8.06 (1H, d, *J* 7.3, C-6), 7.43-7.32 (5H, m, $CO_2CH_2C_6\underline{H}_5$), 7.04 (1H, d, *J* 7.3, C-5), 5.19 (2H, s, $CO_2C\underline{H}_2C_6H_5$), 4.63 (2H, s, $C\underline{H}_2CO_2CH_3$), 3.68 (3H, s, $CH_2CO_2C\underline{H}_3$);^{13}C NMR (DMSO; 75 MHz): δ 168.8 ($\underline{C}O_2CH_3$), 163.8 (C-4), 155.3 ($N\underline{C}O_2Bn$), 153.4 (C-2), 150.6 (C-6), 136.2 (Ph-1), 128.8 (Ph-2), 128.5 (Ph-3), 128.3 (Ph-4), 94.5 (C-5), 66.8 ($CO_2C\underline{H}_2Ph$), 52.6 ($CH_2CO_2\underline{C}H_3$), 50.8 ($\underline{C}H_2CO_2CH_3$); *m/z* : 318 $[M+H]^+$; Data were in accordance with literature data.[2]

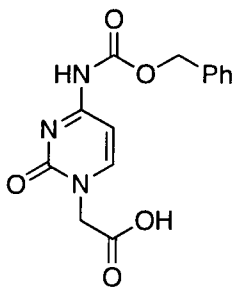
Ethyl 2-(4-(benzyloxycarbonylamino)-2-oxopyrimidin-1(2*H*)-yl)acetate



46

DMAP (4.66 g, 38.0 mmol) then benzyl chloroformate (5.45 mL, 38.0 mmol were sequentially dissolved in CH_2Cl_2 (40 mL) at –15 °C. After 15 min stirring, ethyl 2-(4-amino-2-oxopyrimidin-1(2*H*)-yl)acetate (3.72 g, 19.0 mmol) was added slowly, portionwise and stirred for 15 min at –15 °C. The reaction was allowed to warm to RT while stirring for 5 h then evaporated to dryness to give a pink solid. The crude product was ground to a fine powder and washed with ether before drying under reduced pressure to afford 4.41g (77%) of the title compound as pink amorphous crystals; Mp. 142-3 °C; R_f 0.77 (EtOAc/MeOH, 2:1);^{1}H NMR (DMSO; 300 MHz): δ 8.05 (1H, d, *J* 7.3, C-6), 7.30 (5H, bm, $CO_2CH_2C_6\underline{H}_5$), 7.05 (1H, d, *J* 7.3, C-5), 5.19 (2H, s, $CO_2C\underline{H}_2Ph$), 4.61 (2H, s, $C\underline{H}_2CO_2Et$), 4.14 (2H, q, *J* 7.1, $CO_2C\underline{H}_2CH_3$), 1.20 (3H, t, *J* 7.1, $CO_2CH_2C\underline{H}_3$); ^{13}C NMR (DMSO; 125 MHz): δ 167.2 ($\underline{C}O_2Et$), 162.7 (C-4), 155.4 ($N\underline{C}O_2Bn$), 152.1 (C-2), 148.7 (C-6), 134.9 (Ph-1), 128.7 (Ph-2), 128.4 (Ph-3), 128.3 (Ph-4), 95.1 (C-5), 68.1 ($CO_2\underline{C}H_2Ph$), 62.2 ($CO_2\underline{C}H_2CH_3$), 50.7 ($\underline{C}H_2CO_2Et$), 14.1 ($CO_2CH_2\underline{C}H_3$);*m/z* : 333 $[M+H]^+$; Data were in accordance with literature data.[3]

2-(4-(benzyloxycarbonylamino)-2-oxopyrimidin-1(2*H*)-yl)acetic acid

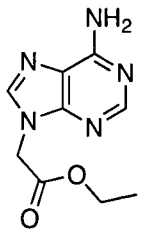


42

1M sodium hydroxide solution (17.5 mL) was added to a solution of ethyl 2-(4-(benzyloxycarbonylamino)-2-oxopyrimidin-1(2*H*)-yl)acetate (4.41 g, 13.9 mmol) dissolved in dioxane (110 mL) and was allowed to stir at RT for 16 h. Dioxane was removed from the reaction under reduced pressure evaporation and the resultant yellow solution was treated with 1M $KHSO_4$ until a colourless precipitate formed. The reaction was filtered to yield 3.41g (81%) of the title compound as colourless amorphous crystals after drying under reduced pressure; Mp. 272-5°C (decomp.); R_f 0.1(EtOAc/MeOH, 1:1); ^{1}H NMR (DMSO; 300 MHz): δ 8.04 (1H, d, *J* 7.3, C-6), 7.56-7.30 (5H, m, $CO_2CH_2C_6\underline{H}_5$), 7.02 (1H, d, *J* 7.3, C-5), 5.19 (2H, s, $CO_2C\underline{H}_2Ph$), 4.52 (2H, s, $C\underline{H}_2CO_2H$); ^{13}C NMR (DMSO; 75 MHz) 170.2 ($CH_2\underline{C}O_2H$), 164.1 (C-4), 155.8

(C-2), 153.9 ($\underline{C}O_2Bn$), 151.2 (C-6), 136.7 (Ph-1), 129.3 (Ph-2), 128.9 (Ph-3), 128.7 (Ph-4), 94.8 (C-5), 67.3 ($CO_2\underline{C}H_2Ph$), 51.3 ($\underline{C}H_2CO_2H$); *m/z* :304 $[M+H]^+$; Data were in accordance with literature data.[1, 2]

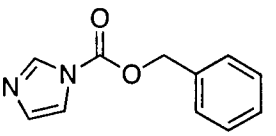
Ethyl 2-(6-amino-9*H*-purin-9-yl)acetate



48

Sodium hydride (3.5 g, 89 mmol) was added carefully in one portion to a stirred suspension of adenine (20.0 g, 148 mmol) in DMF (308 mL) at RT. After 30 min an additional 3.5 g of sodium hydride was added and the reaction and was allowed to stir for 3 h. The mixture was cooled to 0 °C and ethyl bromoacetate (18.1 mL, 163 mmol) was added dropwise slowly, ensuring that the reaction temperature did not exceed 30 °C. The mixture was allowed to warm to RT while stirring for 19 h and was then evaporated to dryness. H_2O (150 mL) was added and the mixture was stirred for 1h. A precipitate was recovered by filtration, washed with H_2O and recrystallised from EtOH (100 mL) to give 21.2 g (65%) of the title compound as colourless needles; Mp. 119-120 °C; R_f 0.53 (Et_2O/MeOH, 3:2); ^{1}H NMR (DMSO; 300 MHz): δ 8.12 (1H, s, A-2), 8.11 (1H, s, A-8), 7.29 (2H, bs, A-6NH_2), 5.06 (2H, s, $C\underline{H}_2CO_2Et$), 4.16 (2H, q, *J* 6.9, $CO_2C\underline{H}_2CH_3$), 1.20 (3H, t, *J* 6.9, $CO_2CH_2C\underline{H}_3$); ^{13}C NMR (DMSO; 75 MHz); δ 168.3 ($\underline{C}O_2Et$), 156.3 (A-6), 152.9 (A-2), 150.0 (A-4), 141.5 (A-8), 118.5 (A-5), 61.7 ($CO_2\underline{C}H_2CH_3$), 44.4 ($\underline{C}H_2CO_2Et$), 14.3 ($CO_2CH_2\underline{C}H_3$); *m/z* : 222 $[M+H]^+$; Data were in accordance with literature data.[3, 4]

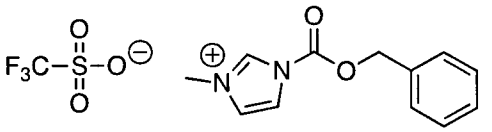
Benzyl 1*H*-imidazole-1-carboxylate



51

Benzylchloroformate (2.0 mL, 14.0 mmol) was added in one portion to a stirred solution of imidazole (1.91 g, 28.0 mmol) dissolved in toluene (23.5 mL) at 0 °C. A colourless precipitate formed immediately. After 5 min stirring the reaction was filtered and evaporated to dryness under reduced pressure to give 2.95 g of the title compound as a colourless oil. The product contained trace impurities and was used immediately as the crude material; ^{1}H NMR (DMSO; 300 MHz): δ 8.15 (1H, s, Im-2), 7.40 (5H, bs, $CO_2CH_2C_6\underline{H}_5$), 7.06 (2H, s, Im-4 and Im-5), 5.40 (2H, s, $CO_2C\underline{H}_2Ph$); ^{13}C NMR (DMSO; 75 MHz); δ 154.2 ($N\underline{C}O_2Bn$), 137.5 (Im-2), 134.4 (Ph-1), 131.1 (Im-4), 129.6 (Ph-2), 129.3 (Ph-3), 129.1 (Ph-4), 117.5 (Im-5); *m/z* : 203 $[M+H]^+$; Data were in accordance with literature data.[5]

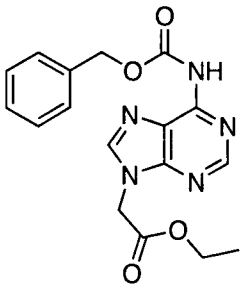
1-(benzyloxycarbonyl)-3-methyl-1*H*-imidazol-3-ium triflate



53

Methyl triflate (296 μL, 2.60 mmol) was added dropwise to a stirred solution of benzyl 1*H*-imidazole-1-carboxylate (0.50 g, 2.48 mmol) dissolved in CH_2Cl_2 (1.10 mL) at 0 °C. A colourless precipitate forms rapidly and additional CH_2Cl_2 (5.0 mL) was added to the reaction to aid stirring. After 5 min stirring at 0 °C the reaction was filtered and the filtrant was washed with ether and dried under reduced pressure to give 905 mg (99%) of the title compound as colourless amorphous crystals; Mp.78-80 °C(decomp.); ^{1}H NMR (DMSO; 300 MHz): δ 9.89 (1H, s, Im-2), 8.18 (1H, s, Im-4), 7.86 (1H, s, Im-5), 7.69 (2H, m, Ph-2), 7.47 (3H, m, Ph-3 and Ph-4), 5.56 (2H, s, $CO_2C\underline{H}_2Ph$), 3.90 (3H, s, Im-3$C\underline{H}_3$); ^{13}C NMR (DMSO; 75 MHz); δ 146.1 ($N\underline{C}O_2Bn$), 139.0 (Im-2), 136.1 (Im-4), 135.7 (Ph-1), 129.3 (Ph-2), 128.9 (Ph-3), 128.5 (Ph-4), 125.0 (Im-5), 120.3 (q, *J* 320, CF_3SO_3), 71.7 ($CO_2\underline{C}H_2Ph$), 36.8 (Im-3 $\underline{C}H_3$); *m/z* : 217 (M^+); Data were in accordance with literature data.[6]

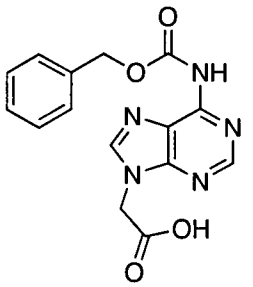
Ethyl 2-(6-(benzyloxycarbonylamino)-9*H*-purin-9-yl)acetate



49

1-(benzyloxycarbonyl)-3-methyl-1*H*-imidazol-3-ium triflate (3.43 g, 6.2 mmol) was added portionwise to a solution of ethyl 2-(6-amino-9*H*-purin-9-yl)acetate (1.38 g, 6.2 mmol) dissolved in DMF (20 mL) at 0 °C. The reaction was allowed to warm to RT while stirring for 18 h before quenching with saturated $NaHCO_3$ solution (15 mL). After 15 min stirring the reaction was partitioned with CH_2Cl_2 (30 mL) and washed with H_2O (30 mL), 1 M $KHSO_4$ solution (2 x 30mL) and saturated brine solution (30 mL). The reaction was evaporated to dryness and the resultant yellow oil was triturated with petrol to give a fine colourless precipitate that was collected by filtration. The solid was dried under reduced pressure to give 1.54g (70%) of the title compound as colourless amorphous crystals; Mp 133-134 °C; R_f 0.44 (Et_2O/MeOH, 9:1); ^{1}H NMR ($CDCl_3$; 500 MHz): δ 9.05 (1H, bs, A-6), 8.82 (1H, s, A-2), 8.02 (1H, s, A-8), 7.48 (2H, m, Ph-2) 7.46-7.39 (3H, m, Ph-3 and Ph-4), 5.35 (2H, s, $NCO_2C\underline{H}_2Ph$), 4.97 (2H, s, A-9 $C\underline{H}_2CO_2Et$), 4.31 (2H, q, *J* 7.1, $CO_2C\underline{H}_2CH_3$), 1.34 (3H, t, *J* 7.1, $CO_2CH_2C\underline{H}_3$); ^{13}C NMR ($CDCl_3$; 125 MHz); δ 167.1 ($\underline{C}O_2C_2H_5$), 153.4 (A-6), 152.0 (NCO_2), 151.4 (A-2 $N\underline{C}O_2Bn$), 150.0 (A-4), 143.6 (A-8), 135.7 (Ph-1), 129.3 (Ph-2), 129.0 (Ph-3), 128.9 (Ph-4), 121.8 (A-5), 68.2 ($CO_2\underline{C}H_2Ph$), 62.9 ($CO_2C\underline{H}_2CH_3$), 44.6 (A-9 $C\underline{H}_2$), 14.5 ($CO_2CH_2C\underline{H}_3$); *m/z* : 356 $[M+H]^+$; Data were in accordance with literature data.[4, 6]

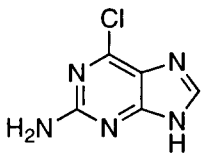
2-(6-(benzyloxycarbonylamino)-9*H*-purin-9-yl)acetic acid



54

2M Sodium hydroxide solution (39 mL) was added to a stirred solution of ethyl 2-(6-(benzyloxycarbonylamino)-9*H*-purin-9-yl)acetate dissolved in MeOH (39 mL) at RT. After 30 min stirring, the reaction was washed with CH_2Cl_2 (2 x 40 mL) and treated with 12 M HCl dropwise until a precipitate formed. The reaction was allowed to stir at RT for 18h and was filtered to recover 1.67g (70%) of the title compound as amorphous crystals after drying under reduced pressure; Mp.156-158°C; R_f 0.2 (EtOAc/MeOH, 1:1); ^{1}H NMR (DMSO; 300 MHz): δ 8.90 (1H, s, A-2), 8.74 (1H, s, A-8), 7.48-7.34 (5H, m, $NHCO_2CH_2C_6\underline{H}_5$), 5.28 (2H, s, $NHCO_2C\underline{H}_2C_6H_5$), 5.20 (2H, s, A-9 $C\underline{H}_2CO_2H$); ^{13}C NMR (DMSO; 75 MHz); δ 169.4 ($CH_2\underline{C}O_2H$), 153.4 (A-6), 152.5 ($NHCO_2Ph$), 150.0 (A-2), 149.1 (A-4), 146.0 (A-8), 136.7 (Ph-1), 129.3 (Ph-2), 129.0 (Ph-3), 128.8 (Ph-4), 120.3 (A-6), 67.8 ($NHCO_2\underline{C}H_2Ph$), 45.7 ($\underline{C}H_2CO_2H$); *m/z* : 328.1 $[M+H]^+$; Data were in accordance with literature data.[1, 4, 6]

6-chloro-9*H*-purin-2-amine

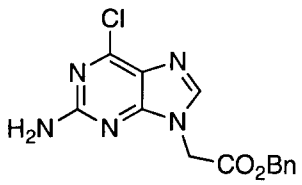


57

Phosphorus oxychloride (12.5 mL, 136 mmol) was added slowly to briskly stirred DMF at 0°C and allowed to warm to RT while stirring for 30 min. Oven-dried guanine (6.85 g, 45 mmol) was added slowly to the rapidly stirred bright yellow reaction and was heated at 90 °C for 5 h. The reaction was cooled to 0 °C and ice was added continuously during gentle agitation of the white and brown solid reaction mixture until only a black solution remained. The reaction was carefully adjusted to pH 8 with solid potassium carbonate and a yellow precipitate was recovered by filtration. The solid was dried

under reduced pressure and ground to a fine powder. The intermediate (10g) was suspended in stirred 12% acetic acid (100 mL) and heated at 70°C for 3h before cooling to RT. The suspension was filtered to recover a bright yellow precipitate. The intermediate was washed with H_2O, dried under reduced pressure and then stirred in 10% NaOH for 2 h at RT. The reaction was carefully neutralised at 0 °C with conc. HCl and was stirred for a further 1 h. The light yellow slurry was heated to 85 °C and cooled rapidly in an ice bath. The precipitate was collected by filtration to give the title compound (6.69g, 88%) as light yellow amorphous crystals; Mp.>300°C; R_f 0.55 (EtOAc);^{1}H NMR (DMSO; 300 MHz): δ 8.10 (1H, s, G-8), 6.77 (2H, bs, G-2NH_2); ^{13}C NMR (DMSO;75 MHz):160.1 (G-2), 155.4 (G-6), 149.1 (G-4), 141.9 (G-8), 123.3 (G-5); *m/z* : 170, 172 $[M+H]^+$.

Benzyl 2-(2-amino-6-chloro-9*H*-purin-9-yl)acetate

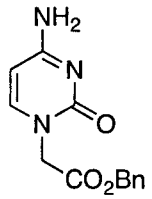


65

6-chloro-9H-purin-2-amine (12.67 g, 75 mmol) was heated to 85 °C in rapidly stirred DMF (127 mL) and potassium carbonate (15.4 g, 112 mmol) was added in one portion. After brief stirring the reaction was cooled to –5 °C and benzyl bromoacetate (4.23 mL, 25 mmol) was added dropwise slowly to the reaction, ensuring that the temperature did not exceed 0 °C. The reaction was stirred for 3 h at 0 °C, with further additions of benzyl bromoacetate (2 x 4.23 mL, 25 mmol) every sixty minutes. The reaction was allowed to warm to RT while stirring for 18 h. A light yellow precipitate was collected by filtration and washed with DMF before being added to vigorously stirred 1M HCl (67 mL). After 2 h stirring the yellow suspension was filtered, washed with H_2O (650 mL) and air-dried under suction on the sinter funnel. The solid was recrystallised from MeCN (125 mL) and dried thoroughly under reduced pressure to give 9.37g (38%) of the title compound as colourless needles. The mother liquor was concentrated and heated briefly at reflux in a minimal amount of MeCN to yield a further crop (2.36g, 10%) of the title compound; Mp. 187-8 °C; R_f 0.57 (EtOAc/Petrol, 1:1); ^{1}H NMR (DMSO; 300 MHz): 8.12 (1H, s, G-8), 7.37 (5H, m, $CO_2CH_2C_6\underline{H}_5$), 5.20 (2H, s, G-2 NH_2), 5.20 (2H, s, $CO_2C\underline{H}_2Ph$), 5.07 (2H, s, G-9$C\underline{H}_2$);^{13}C NMR (DMSO; 75 MHz);

168.5 ($\underline{C}O_2Bn$), 160.7 (G-2), 155.0 (G-6), 150.3 (G-4), 144.2 (G-8), 136.1 (Ph-1), 129.3 (Ph-2), 129.1 (Ph-3), 128.9 (Ph-4), 123.7 (G-5), 67.6 ($C\underline{H}_2Ph$), 44.8 (G-9$C\underline{H}_2$); *m/z* : 318 $[M+H]^+$; Data were in accordance with literature data.[3]

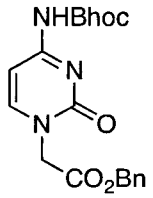
Benzyl 2-(4-amino-2-oxopyrimidin-1(2H)-yl)acetate



70

Cytosine (2.06 g, 18 mmol) and potassium *tert*-butoxide (2.32 g, 21 mmol) were heated in dry DMF (18 mL) at 100 °C for 2 h. The colourless suspension was cooled to 10 °C and benzyl bromoacetate (3.21 mL, 20 mmol) was added dropwise. The reaction was allowed to warm to RT while stirring for 18h and was then quenched with glacial acetic acid (1.18 mL). The pink suspension was evaporated to dryness. The residue was suspended in H_2O (20 mL) and stirred for 4 h and a pink precipitate was collected by filtration and washed with H_2O (4 x 30 mL). The solid was recrystallised from DMF and dried under reduced pressure to give the title compound (3.73g, 80%) as square plate crystals; Mp. 245-248 °C; R_f 0.20 (EtOAc/petrol, 1:1); ^{1}H NMR (DMSO; 300 MHz): δ 7.57 (1H, d, *J* 7.1, C-6), 7.37 (5H, m, $CH_2C_6\underline{H}_5$), 7.22 (1H, bs, C-4NH), 7.12 (1H, bs, C-4NH), 5.68 (1H, d, *J* 7.1, C-5), 5.16 (2H, s, $C\underline{H}_2Ph$), 4.51 (2H, s, C-1CH_2); ^{13}C NMR (DMSO; 75 MHz); δ 169.6 (CO_2Bn), 167.2 (C-4), 156.6 (C-2), 147.2 (C-6), 136.5 (Ph-1), 129.3 (Ph-2), 129.0 (Ph-3), 128.7 (Ph-4), 94.4 (C-5), 66.9 ($\underline{C}H_2Ph$), 50.8 (C-1CH_2); *m/z* : 260 $[M+H]^+$; Data were in accordance with literature data.[3]

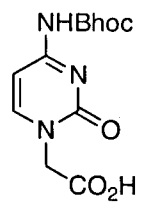
Benzyl 2-(4-(benzhydryloxycarbonylamino)-2-oxopyrimidin-1(2*H*)-yl)acetate



71

Benzyl 2-(4-amino-2-oxopyrimidin-1(2H)-yl)acetate (956 mg, 3.69 mmol) and carbonyldiimidazole (944 mg, 5.91 mmol) were suspended in dry DMF (7.5 mL) and stirred for 5 h. Benzhydrol (878 mg, 4.80 mmol) was added and the reaction was heated at 60°C for 1 h. Benzhydrol (67 mg, 0.48 mmol) was added each hour for a further two hours reaction was heated at 60 °C for 1 h after each addition. The reaction was stirred for 2 h at 60 °C and then evaporated to dryness. The crude oil was triturated with H_2O to give a colourless precipitate that was recrystallised from EtOH to give the title compound (1.42g, 82%) as colourless plates; Mp.133-135 °C; R_f 0.40 (EtOAc/petrol, 1:1); ^{1}H NMR (DMSO; 300 MHz): δ 11.07 (1H, s, C-4NH), 8.06 (1H, d, *J* 7.3, C-6), 7.46 (4H, d, *J* 7.1, Bhoc;Ph-2), 7.37 (9H, m, $CO_2CH_2C_6\underline{H}_5$ and Bhoc;Ph-3), 7.29 (2H, t, *J* 7.2, Bhoc;Ph-4), 6.98 (2H, d, *J* 7.3, C-5), 6.80 (1H, s, $CHPh_2$), 5.19 (2H, s, $CO_2C\underline{H}_2Ph$), 4.69 (2H, s, C-1CH_2); ^{13}C NMR (DMSO; 75 MHz); δ 168.7 CO_2Bn), 164.2 (C-4), 156.1 (NCO_2), 151.3 (C-2), 141.2 (C-6), 136.4 (Bhoc;Ph-1, Ph-1), 129.4 (Bhoc;Ph-2, Ph-2), 129.0 (Bhoc;Ph-3, Ph-3), 127.3 (Bhoc;Ph-4, Ph-4), 95.0 (C-5), 78.3 ($CHPh_2$), 67.2 (CO_2CH_2), 51.5 (C-1CH_2); *m/z* : 470 $[M+H]^+$; Data were in accordance with literature data.[3]

2-(4-(benzhydryloxycarbonylamino)-2-oxopyrimidin-1(2H)-yl)acetic acid

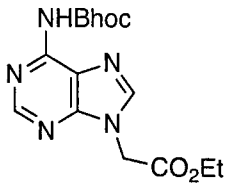


73

Benzyl 2-(4-(benzhydryloxycarbonylamino)-2-oxopyrimidin-1(2*H*)-yl)acetate (1.42 g, 3.03 mmol) was suspended in a 1:1:2:2 mixture of EtOH/H_2O/MeOH/MeCN (18.5 mL) and cooled to 0 °C. A solution of lithium hydroxide monohydrate (1.28g dissolved in 9.4mL of H2O) was added and the reaction was stirred until hydrolysis was complete,

as indicated by tlc. A solution of citric acid (2.85g dissolved in 14.2mL of H2O) was added resulting in immediate dissolution of the precipitate. The solution was chilled to 0°C and stirred for 2h. A colourless precipitate was recovered by filtration then suspended in a minimal amount of saturated brine solution and stirred for 6h. The precipitate was recovered by filtration and dried thoroughly to give the title compound (1.19g, 82%) as colourless amorphous crystals; Mp. 197-8°C; R_f 0.30 (EtOAc/MeOH, 7:3); ^{1}H NMR (DMSO; 300 MHz): δ 8.02 (1H, d, *J* 7.3, C-6), 7.45 (4H, d, *J* 7.2, Ph-2), 7.38 (4H, t, *J* 7.2, Ph-3), 7.29 (2H, t, Ph-4), 6.96 (1H, d, *J* 7.2, C-5), 6.81 (1H, s, $CHPh_2$), 4.52 (2H, s, C-1CH_2); ^{13}C NMR (DMSO; 75 MHz); δ 170.2 (CO_2H), 164.0 (C-4), 155.8 (NCO_2), 153.2 (C-2), 151.4 (C-6), 141.2 (Ph-1), 129.4 (Ph-2), 128.7 (Ph-3), 127.3 (Ph-4), 94.7 (C-5), 78.3 ($CHPh_2$), 51.5 (C-1CH_2); *m/z* : 380 $[M+H]^+$; Data were in accordance with literature data.[3]

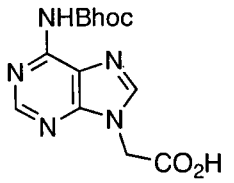
Ethyl 2-(6-(benzhydryloxycarbonylamino)-9H-purin-9-yl)acetate



74

Ethyl 2-(6-amino-9H-purin-9-yl)acetate (2.70 g, 12.2 mmol) and carbonyl diimidazole (2.93 g, 18.3 mmol) were heated in dry DMF (25 mL) at 100 °C for 3h and then cooled to 95 °C. Benzhydrol (3.37 g, 18.3 mmol) was added and the mixture was allowed to cool to RT while stirring for 16 h. The reaction was cooled to 0 °C and acetone (10 mL), then H_2O (50 mL) were added to the mixture. The reaction was stirred at 0 °C for 2 h then filtered to recover a yellow precipitate that was recrystallised from MeOH to give the title compound (3.28g, 63 %) as colourless needles; Mp. 88-89°C; R_f 0.75 (EtOAc); ^{1}H NMR (DMSO; 300 MHz): δ 10.98 (1H, s, A-6NH), 8.63 (1H, s, A-2), 8.46 (1H, s, A-8), 7.53 (4H, d, *J* 7.3, Ph-2), 7.38 (4H, t, *J* 7.3, Ph-3), 7.29 (2H, d, *J* 7.3, Ph-4), 6.83 (1H, s, $CHPh_2$), 5.20 (2H, s, A-9CH_2), 4.17 (2H, t, *J* 7.1, $\underline{C}H_2CH_3$), 1.21 (3H, q, *J* 7.1, CH_3); ^{13}C NMR (DMSO; 75 MHz); δ 168.5 (CO_2H), 152.9 (A-6), 152.6 (A-2), 151.9 (A-6N$\underline{C}O_2$Bn), 150.2 (A-4), 145.5 (A-8), 141.7 (Ph-1), 129.3 (Ph-2), 128.6 (Ph-3), 127.3 (Ph-4), 123.5 (A-5), 78.1 ($CHPh_2$), 62.4 ($C\underline{H}_2CH_3$), 45.1 (A-9CH_2), 14.8 (CH_3); *m/z* : 432 $[M+H]^+$; Data were in accordance with literature data.[3]

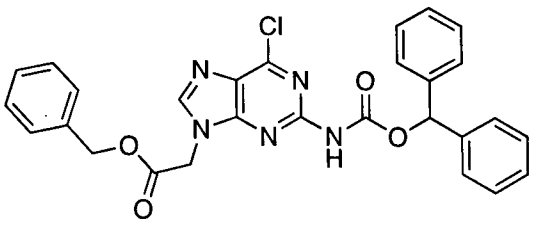
2-(6-(benzhydryloxycarbonylamino)-9H-purin-9-yl)acetic acid



75

Ethyl 2-(6-(benzhydryloxycarbonylamino)-9H-purin-9-yl)acetate (3.0 g, 6.96 mmol) was suspended in a mixture of MeCN/EtOH (1:1, 23mL) and stirred vigorously until dissolution occurred. H_2O (9.1 mL) was added slowly to the solution and the temperature was reduced to 0 °C before addition of lithium hydroxide monohydrate solution (4.68 g dissolved in 38 mL of H_2O). Upon precipitation of a colourless solid, citric acid solution (13.38g dissolved in 46 mL of H_2O) was added immediately. The reaction was stirred for 2 h at 0 °C and then filtered to give the title compound as its citrate salt. The solid was slurried in the minimum amount of saturated brine solution for 16h then filtered under reduced pressure to give the title compound (2.83 g, quantitative) as colourless amorphous crystals; Mp.109-111°C; R_f 0.15 (EtOAc/MeOH, 7:3); ^{1}H NMR (DMSO; 300 MHz): δ 10.89 (1H, s, A-6NH), 8.58 (1H, s, A-2), 8.39 (1H, s, A-8), 7.53 (4H, d, *J* 7.3, Ph-2), 7.37 (4H, t, *J* 7.3, Ph-3), 7.28 (2H, d, *J* 7.3, Ph-4), 6.82 (1H, s, $CHPh_2$), 4.80 (2H, s, A-9CH_2); ^{13}C NMR (DMSO; 75 MHz); δ 172.0 (CO_2H), 155.3 (A-6), 154.5 (A-2), 154.4 (A-6N$\underline{C}O_2$Bn), 152.2 (A-4), 148.5 (A-8), 144.1 (Ph-1), 131.7 (Ph-2), 130.8 (Ph-3), 129.6 (Ph-4), 125.8 (A-5), 74.3 ($CHPh_2$), 47.6 (A-9CH_2); *m/z* : 404 $[M+H]^+$; Data were in accordance with literature data.[3, 7]

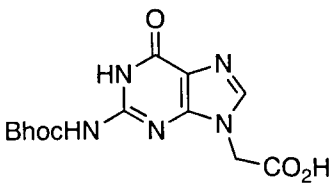
Benzyl 2-(2-(benzhydryloxycarbonylamino)-6-chloro-9*H*-purin-9-yl)acetate



76

Triphosgene (1.67 g, 5.6 mmol) was added in one portion to a stirred colourless supension of benzyl 2-(2-amino-6-chloro-9*H*-purin-9-yl)acetate (5.00 g, 15.7 mmol) in THF (60 mL) at 0 °C. After 1 h stirring, DIPEA (6.03 mL, 34.6 mmol) was added dropwise to the yellow reaction and stirred for 2 h. Benzhydrol (3.48 g, 18.8 mmol) was added and the reaction was allowed to warm to RT while stirring for 18 h. The orange reaction was quenched with EtOH (30 mL)and washed sequentially with 10% citric acid solution (50 mL) and saturated $NaHCO_3$ solution (50 mL). The mixture was dried ($MgSO_4$) and evaporated under reduced pressure to give an orange oil. Purification of the crude oil by chromatography (EtOAc/petrol, 1:4) gave the title compound (4.50 g, 55%) as colourless amorphous crystals; Mp. 180-181°C; R_f 0.25 (EtOAc/petrol, 1:4); ^{1}H NMR (DMSO; 300 MHz): δ 11.03 (1H, s, G-2NH), 8.52 (1H, s, G-8), 7.50 (4H, d, Bhoc;Ph-2), 7.39-7.28 (11H, m, Bhoc;Ph-3, Bhoc;Ph-4, $CO_2CH_2C_6H_5$), 6.82 (1H, s, Bhoc;CHPh$_2$), 5.22 (2H, s, $CO_2CH_2C_6H_5$), 5.19 (2H, s, G-9CH$_2$);^{13}C NMR (DMSO; 75 MHz); δ 168.1 (CO_2Bn), 154.0 (NCO_2CHPh$_2$), 153.1 (G-2), 151.7 (G-6), 150.0 (G-4), 147.6 (G-8), 136.0 (Bhoc;Ph-1, Ph-1), 129.3 (Bhoc;Ph-2 Ph-2), 129.1 (Bhoc;Ph-2, Ph-2), 128.8 (Bhoc;Ph-3, Ph-3), 128.8 (Bhoc;Ph-4, Ph-4), 127.3 (G-5), 77.9 (BhocCHPh$_2$), 67.7 (CH$_2$Ph), 45.3 (G-9CH$_2$); *m/z* : 523 $[M+H]^+$; Data were in accordance with literature data.[3]

2-(2-(benzhydryloxycarbonylamino)-6-oxo-1H-purin-9(6H)-yl)acetic acid

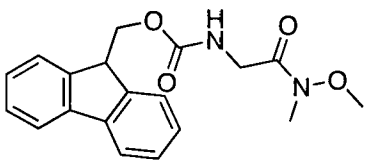


77

Sodium hydride (42 mg, 1.0 mmol) was added to dry THF (1.15 mL) and chilled to –78 °C. 3-Hydroxypropionitrile (72 μL, 1.0 mmol) was added to the suspension. The

reaction was warmed to 0 °C and stirred for 2.5 h. Benzyl 2-(2-(benzhydryloxycarbonylamino)-6-chloro-9H-purin-9-yl)acetate (100 mg, 0.21 mmol) was added to the reaction and the reaction was allowed to warm to RT while stirring for 48 h. The solvent was removed by evaporation and the crude yellow residue was dissolved in H_2O (0.66 mL). 20% citric acid solution (0.8 mL) was added to the stirred yellow solution. A colourless precipitate was collected by filtration and recrystallised (MeOH) to give the title compound (64 mg, 82%) as colourless needles; Mp. 171-2°C; Rf 0.18 (30% MeOH in EtOAc); ^{1}H NMR (DMSO; 300 MHz): δ 11.76 (1H, bs, CO_2H), 11.28 (1H, bs, G-2NH), 7.93 (1H, s, G-8), 4.76 (4H, d, *J* 7.2, Ph-2), 7.35 (4H, t, *J* 7.2, Ph-3), 7.30 (2H, t, *J* 7.2, Ph-4), 6.86 (1H, s, C<u>H</u>Ph_2), 4.82 (2H, s, G-9CH_2); ^{13}C NMR (DMSO; 75 MHz); δ 170.0 (CO_2H), 156.0 (N<u>C</u>O_2CHPh_2), 154.5 (G-2), 150.0 (G-6), 147.9 (G-4), 141.2 (G-8), 140.9 (G-5), 135.8 (Ph-1), 129.4 (Ph-2), 128.8 (Ph-3), 127.2 (Ph-4), 78.8 (<u>C</u>HPh_2), 45.6 (G-9CH); *m/z* : 420 $[M+H]^+$; Data were in accordance with literature data.[3, 7]

(9*H*-fluoren-9-yl)methyl 2-(methoxy(methyl)amino)-2-oxoethylcarbamate

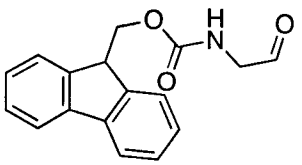


79

Isobutylchloroformate (2.20 mL, 16.7 mmol) was added dropwise to a mixture of 2-(((9*H*-fluoren-9-yl)methoxy)carbonylamino)acetic acid (5.0 g, 16.7 mmol) and N-Methylmorpholine (1.85 mL, 16.7 mL) stirred at –20 °C in THF (100 mL). After 10 min stirring, triethylamine (2.50 mL, 18.4 mmol) and a solution of *N,O*-dimethylhydroxylamine (1.65 g, 16.7 mmol) dissolved in DMF (40mL) were added sequentially to the reaction and stirred for 30 min at –20°C.The reaction was allowed to warm to RT and a colourless precipitate was removed by filtration. The filtrate was diluted with H_2O (166 mL) and acidified with 1M hydrochloric acid (16mL) before removing the THF component under reduced pressure evaporation. The aqueous mixture was adjusted to pH 8 by careful addition of 2M K_2CO_3 solution extracted with EtOAc (3 x 50 mL). The combined organic fractions were dried ($MgSO_4$) and evaporated under reduced ptressure to give the title compound (5.22 g, 92%) as a colourless syrup; R_f 0.25 (EtOAc/petrol, 1:1); ^{1}H NMR (DMSO; 300 MHz): δ 7.80 (2H,

d, *J* 7.4, Fmoc 1,8), 7.66 (2H, d, *J* 7.4, Fmoc 4,5), 7.44 (2H, t, *J* 7.4, Fmoc 2,7), 7.35 (2H, d, *J* 7.4, Fmoc 3,6), 5.66 (1H, bt, CO_2NH), 4.42 (2H, d, *J* 7.2, Fmoc 10), 4.29 (1H, t, *J* 7.2, Fmoc 9), 4.21 (2H, d, *J* 4.53, $C\underline{H}_2NHCO_2$), 3.77 (3H, s, $CONCH_3OC\underline{H}_3$), 3.26 (3H, s, $CONC\underline{H}_3OCH_3$); ^{13}C NMR (DMSO; 75 MHz); δ 166.1 ($CH_2\underline{C}O$), 156.7 (CO_2NH), 144.3 (Fmoc 1a,8a), 141.7 (Fmoc 4a,5a), 128.1 (Fmoc 2,7), 127.5 (Fmoc 3,6), 125.6 (Fmoc 4,5), 120.4 (Fmoc 1,8), 67.5 (Fmoc 10), 61.9 (OCH_3), 47.5 (Fmoc 9), 42.6 (NCH_3), 31.4 ($\underline{C}H_2CON$); *m/z* : 341$[M+H]^+$; Data were in accordance with literature data.[8]

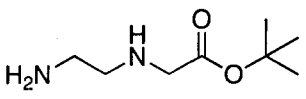
(9*H*-fluoren-9-yl)methyl 2-oxoethylcarbamate



80

Lithium aluminium hydride (615 mg, 16.1 mmol) was added to a stirred solution of (9H-fluoren-9-yl)methyl 2-(methoxy(methyl)amino)-2-oxoethylcarbamate (5.0 g, 14.7 mmol) dissolved in THF (80 mL) held at 0 °C. Upon cessation of effervescence, 1M $KHSO_4$ solution (36.5 mL)was added to the brown reaction mixture, resulting in solution becoming colourless with the formation of a grey precipitate and further effervescence being observed. The crude reaction mixture was extracted with EtOAc (3 x 50 mL) and the combined organics were washed with 1M HCl (2 x 50 mL) and 1M $NaHCO_3$ (50 mL) solutions. The organic fraction was dried ($MgSO_4$) and evaporated to give the title compound (3.95g, 95.9%) as a colourless foam; R_f 0.64 (EtOAc/petrol, 1:1); ^1H NMR (DMSO; 300 MHz): δ 9.70 (1H, s, $CH_2C\underline{H}O$), 7.80 (2H, d, *J* 7.4, Fmoc 1,8), 7.63 (2H, d, *J* 7.4, Fmoc 4,5), 7.45 (2H, t, *J*7.4, Fmoc 2,7), 7.36 (2H, d, *J*7.4, Fmoc 3,6), 5.51 (1H, bm, CO_2NH), 4.46 (2H, d, *J* 7.2, Fmoc 10), 4.28 (1H, t, *J* 7.2, Fmoc 9), 4.20 (2H, d, *J* 5.0, $C\underline{H}_2CHO$); ^{13}C NMR (DMSO; 75 MHz); δ 197.0 ($CH_2\underline{C}HO$), 156.7 (CO_2NH), 144.1 (Fmoc 1a,8a), 141.7 (Fmoc 4a,5a), 128.2 (Fmoc 2,7), 127.5 (Fmoc 3,6), 125.4 (Fmoc 4,5), 120.4 (Fmoc 3,6), 67.6 (Fmoc 10), 47.5 (Fmoc 9), 31.4 ($\underline{C}H_2CO$); *m/z* : 282.1 $[M+H]^+$. ; Data were in accordance with literature data.[8]

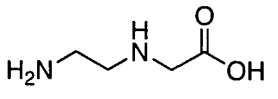
***Tert*-butyl 2-(2-aminoethylamino)acetate**



85

A solution of *tert*-butyl bromoacetate (4.0 mL, 27.1 mmol) dissolved in CH_2Cl_2 (30 mL) was added dropwise to a stirred solution of ethylene diamine (16.3 mL, 244 mmol) dissolved in CH_2Cl_2 (130 mL) at 0 °C. The reaction was allowed to warm to RT while stirring for 14 h before the crude mixture was washed with H_2O (3 x 30 mL) and the combined aqueous phases back-extracted with CH_2Cl_2 (2 x 30 mL). The combined organics were dried ($MgSO_4$) and evaporated to dryness to give the title compound quantitatively as a colourless gum; R_f 0.50 (MeOH/NH_3, 65:35); ^{1}H NMR ($CDCl_3$; 300 MHz): δ 3.18 (2H, s, $C\underline{H}_2CO_2{}^t$Bu), 2.67 (2H, t, *J* 5.9, $C\underline{H}_2CH_2NH_2$), 2.54 (2H, t, *J* 5.9, $CH_2C\underline{H}_2NH_2$), 1.45 (2H, bs, NH_2), 1.35 (9H, s, $C(CH_3)_3$); ^{13}C NMR (DMSO; 75 MHz); δ 172.3 (CO_2), 81.6 ($\underline{C}(CH_3)_3$), 52.6 ($C\underline{H}_2CO_2C_4H_9$), 52.0 ($C\underline{H}_2CH_2NH_2$), 42.2 ($\underline{C}H_2NH_2$), 28.5 ($C(\underline{C}H_3)_3$); *m/z* : 175.2$[M+H]^+$; Data were in accordance with literature data.[4]

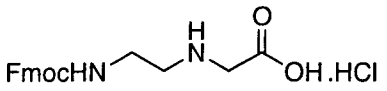
2-(2-aminoethylamino)acetic acid



88

Chloroacetic acid (10.0 g, 106 mmol) was added portionwise to freshly distilled ethylene diamine (48 mL, 719 mmol) at 5°C. The reaction was allowed to warm to RT while stirring for 48h. Excess ethylene diamine was removed under reduced pressure and the resultant yellow oil was triturated with DMSO (100mL) and stirred at RT for 24h. The colourless precipitate was recovered under reduced-pressure filtration and washed with ether. The crude product was recrystallised from aqueous EtOH (EtOH/H_2O, 2:1) to give the title compound (8.13g, 65%) as colourless needles; Mp. 153 °C; R_f 0.10 (MeOH); ^{1}H NMR (D_2O; 500 MHz): δ 3.23 (2H, s, $C\underline{H}_2CO_2H$), 2.99 (2H, t, *J* 6.1, $NHC\underline{H}_2CH_2$), 2.86 (2H, t, *J* 6.1, $NH_2C\underline{H}_2$); ^{13}C NMR (D_2O; 75 MHz); δ 178.2 ($\underline{C}O_2H$), 51.5 ($\underline{C}H_2CO_2H$), 46.4 ($NHCH_2$), 38.5 ($NH_2\underline{C}H_2$); *m/z* :119 $[M+H]^+$; Data were in accordance with literature data.[9]

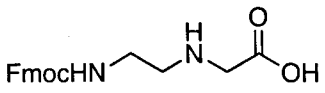
2-(2-(((9H-fluoren-9-yl)methoxy)carbonylamino)ethylamino)acetic acid Hydrochloride salt



89a

A solution of N-(9-Fluorenylmethoxycarbonyloxy)succinimide (16.6 g, 49.2 mmol) dissolved in DMF (100 mL) was added dropwise to a stirred solution of *N,O*-bis(trimethylsilyl)acetamide (48.7 mL, 197 mmol) and (2-aminoethylamino) acetic acid (5.53g, 46.9 mmol) dissolved in DMF (120 mL). After stirring at RT for 1 h the reaction was cooled to 0 °C. MeOH (50 mL) and H_2O (50 mL) were added concurrently to the reaction and stirred for 10 min. The reaction was allowed to warm to RT and 6M HCl (8.3 mL) was added slowly to the vigorously stirred reaction. After 10 min, the solvent was removed under reduced pressure evaporation and the resultant yellow oil was cooled to 0°C and triturated with CH_2Cl_2 (100 mL). The precipitate was stirred for 1 h at 0°C and additional CH_2Cl_2 (100 mL) was added gradually during a further 1 h of stirring. The precipitate was filtered rapidly and washed with (100 mL) CH_2Cl_2. The crude compound was added to warm (45°C) EtOAc (120 mL) and stirred thoroughly for 2 h. The suspension was filtered and the recovered solid was washed thoroughly with warm EtOAc and CH_2Cl_2 before drying under reduced pressure to give the title compound (15.3g, 83%) as amorphous colourless crystals; Mp. 137-8 °C; R_f 0.55 (2:2:1 n-butanol/H_2O/AcOH; ^{1}H NMR (DMSO; 300 MHz): δ 7.89 (2H, d, *J* 7.5, Fmoc 1,8), 7.69 (2H, d, *J* 7.5, Fmoc 4,5), 7.47 (1H, t, *J* 5.5, FmocNH), 7.42 (2H, t, *J* 7.5, Fmoc-2,7), 7.34 (2H, t, *J* 7.5, Fmoc-3,6), 4.33 (2H, d, *J* 7.2, Fmoc-10), 4.23 (1H, t, *J* 7.2, Fmoc-9), 3.52 (2H, s, CH_2CO_2H), 3.29 (2H, t, *J* 5.8, $FmocNHCH_2CH_2$), 2.95 (2H, t, *J* 5.8, $FmocNHCH_2$);^{13}C NMR (DMSO; 75 MHz); δ 173.2 (CO_2H), 156.6 (Fmoc;CO_2N), 144.1 (Fmoc-1a,8a), 141.1 (Fmoc-4a,5a), 127.9 (Fmoc-2,7), 127.4 (Fmoc-3,6), 125.5 (Fmoc-4,5), 120.5 (Fmoc-1,8), 65.8 (Fmoc-10), 47.1 (Fmoc-9), 46.9 (CH_2CO_2H), 46.7 ($FmocNHCH_2$), 40.2 ($FmocNHCH_2CH_2$); *M/z* : 341[M+H]$^+$; Data were in accordance with literature data.[3, 10]

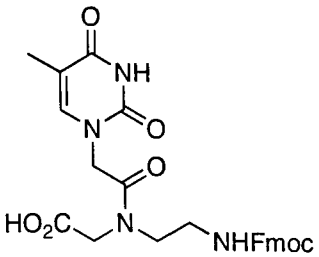
2-(2-(((9*H*-fluoren-9-yl)methoxy)carbonylamino)ethylamino)acetic acid



89

2-(2-(((9H-fluoren-9-yl)methoxy)carbonylamino)ethylamino)acetic acid Hydrochloride salt (5.0g, 12.6mmol) was suspended in 87mL MeOH and heated with stirring until dissolution occurred. The stirred solution was allowed to cool to RT and a solution of NaOH (478mg, 11.95mmol) dissolved in dry MeOH (105mL) was added dropwise over a period of 30 min, during which time a precipitate forms. H_2O (9mL) was added to the mixture and heating was applied until a clear solution was obtained and the reaction was cooled slowly to 0°C and stirred for 1h. The resulting colourless precipitate was recovered by vacuum filtration, washed /stirred thoroughly three times with 90% aqueous MeOH (3x20mL) and dried under vacuum in a desiccator for 48h. The title compound (3.06g, 71%) was obtained as colourless amorphous crystals; Mp. 136-8°C; R_f 0.55 (2:2:1 n-butanol/H_2O/acetic acid; ^{1}H NMR (DMSO; 500 MHz): δ 7.88 (2H, d, *J* 7.5, Fmoc 1,8), 7.68 (2H, d, *J* 7.5, Fmoc 4,5), 7.51 (1H, t, *J* 5.5, FmocNH), 7.34 (2H, t, *J* 7.5, Fmoc 2,7), 7.33 (2H, t, *J* 7.5, Fmoc 3,6), 4.31 (2H, d, *J* 7.2, Fmoc 10), 4.22 (1H, t, *J* 7.2, Fmoc-9) 3.25 (4H, m, $CH_2NHCH_2CO_2H$), 2.86 (2H, t, *J* 5.8, $FmocNHCH_2$); ^{13}C NMR (DMSO; 75 MHz); δ 168.3 (CO_2H), 157.0 (Fmoc;NCO_2), 144.6 (Fmoc 1a,8a), 141.5 (Fmoc 4a,5a), 128.4 (Fmoc 2,7), 127.9 (Fmoc 3,6), 126.0 (Fmoc 4,5), 120.9 (Fmoc 1,8), 66.3 (Fmoc 10), 47.1 (Fmoc 9), 50.6 (CH_2CO_2H), 47.5 ($FmocNHCH_2$), 38.1 ($FmocNHCH_2CH_2$); *m/z* : 341 $[M+H]^+$; Data were in accordance with literature data.[3, 10]

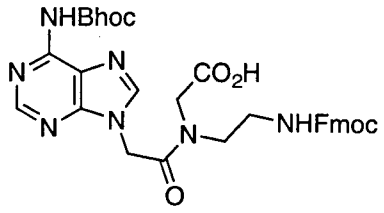
2-(N-(2-(((9H-fluoren-9-yl)methoxy)carbonylamino)ethyl)-2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)acetamido)acetic acid



37

2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)acetic acid (750 mg, 4.1 mmol) and NMM (895 μL, 8.2 mmol) were stirred together in MeCN (13.9 mL) for 10 min. The mixture was cooled to 0°C and added dropwise to vigorously stirred pivaloyl chloride (503 μL, 4.1 mmol) at 0°C. 2-(2-(((9H-fluoren-9-yl)methoxy)carbonylamino)ethylamino)acetic acid (1.39 g, 4.1 mmol) was suspended in a vigorously stirred 3:4 mixture of H_2O/MeCN and triethylamine was added dropwise until dissolution occurred. After 25 min the two solutions were combined and allowed to stir together for 30 min while warming to RT. The reaction was cooled to 0 °C and acidified with 3M HCl to pH 2 and allowed to warm to RT while stirring for 48 h. A colourless precipitate was recovered by filtration, washed thoroughly with H_2O and then ether before drying under reduced pressure to give the title compound (935 mg, 45%) as colourless amorphous crystals; Mp.197-8°C; R_f 0.55 (EtOAc/MeOH, 1:1); ^{1}H NMR (DMSO; 300 MHz; Rotomeric mixture (3:4): δ 11.33 (2H, s, CO_2H; R_A and R_B), 7.89 (4H, d, *J* 7.4, Fmoc-1,8; R_A and R_B), 7.68 (4H, d, *J* 7.4, Fmoc-4,5; R_A and R_B), 7.41 (4H, t, *J* 7.4, Fmoc-2,7; R_A and R_B), 7.34-7.26 (6H, m, Fmoc-3,6 and T-6; R_A and R_B), 4.65 (2H, s, T-1CH_2;R_B), 4.47 (2H, s, T-1CH_2;R_A), 4.35-4.20 (6H, m, Fmoc-9,10; R_A, R_B and C$\underline{H_2}CO_2H$;R_A), 3.98(2H, s, C$\underline{H_2}CO_2H$;R_B), 3.40 (2H, m, C$\underline{H_2}$NHFmoc;R_B), 3.40 (2H, m, C$\underline{H_2}$NHFmoc;R_B), 3.35-3.20 (4H, m, C$\underline{H_2}$NHFmoc;R_A and C$\underline{H_2}CH_2$NHFmoc;R_B), 3.10 (2H, q, *J* 5.8, C$\underline{H_2}CH_2$NHFmoc;R_A), 1.72 (3H, s, T-5CH_3);^{13}C NMR (DMSO; 75 MHz): δ 171.7 (CO_2H;R_A), 170.8 (CO_2H;R_B), 167.9 (T-1$CH_2\underline{C}$ON;R_A), 167.5 (T-1$CH_2\underline{C}$ON;R_B), 164.7 (T-4), 156.6 (Fmoc;CO_2N;R_B), 156.4 (Fmoc;CO_2N;R_A), 151.3 (T-2), 144.2(Fmoc-1a,8a), 142.4(T-6), 141.0 (Fmoc-4a,5a), 127.9 (Fmoc-2,7), 127.4(Fmoc-3,6), 125.4(Fmoc-4,5), 120.4(Fmoc-1,8), 108.5(T-5;R_A), 108.4(T-5;R_B), 65.8(Fmoc-10), 49.3 (C$\underline{H_2}$NHFmoc;R_A), 48.01 C$\underline{H_2}$NHFmoc;R_B), 47.0(Fmoc-9), 38.9 (C$\underline{H_2}CH_2$NHFmoc;R_B), 38.2 (C$\underline{H_2}CH_2$NHFmoc;R_A), 12.3 (T-5CH_3); *m/z* : 507 $[M+H]^+$; Data were in accordance with literature data.[3, 4]

2-(N-(2-(((9H-fluoren-9-yl)methoxy)carbonylamino)ethyl)-2-(6-(benzhydryloxycarbonylamino)-9H-purin-9-yl)acetamido)acetic acid

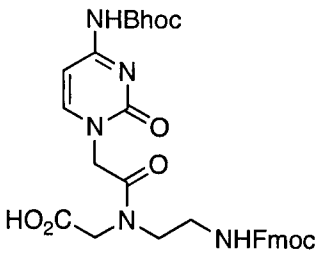


32

A mixture of 2-(6-(benzhydryloxycarbonylamino)-9H-purin-9-yl)acetic acid (608 mg, 1.5mmol) and NMM (185 μL, 3.0 mmol) in MeCN (5 mL) was cooled to –20°C and added dropwise to vigorously stirred pivaloyl chloride (206 μL, 1.5 mmoL). 2-(2-(((9H-fluoren-9-yl)methoxy)carbonylamino)ethylamino)acetic acid (572 mg, 1.5 mmol) was suspended in a vigorously stirred mixture of H_2O/MeCN (1:1, 10 mL) and triethylamine was added dropwise until dissolution occurred. After 20 min, the two solutions were combined and the reaction was allowed to warm to RT while stirring for 45 min. The reaction was evaporated to dryness, dissolved in the minimum amount of hot MeOH and added dropwise to 15 mL of rapidly stirred 20% citric acid solution. The precipitate was collected by filtration, washed and agitated thoroughly with brine on the sinter. A sample was dissolved in DMF and checked for purity by HP-LC. The solid was repeatedly treated to cycles of dissolution in hot MeOH and trituration by dropwise addition to stirred citric acid solution until all impurities had been removed. The precipitate was dried and analysed by NMR to determine if any PNA-citrate adduct had formed. If the citrate was present, the solid was ground to a fine powder and slurried with saturated brine solution for 36 hours. The slurry was filtered and washed repeatedly with H_2O. The precipitate was dried under reduced pressure to give the title compound (812 mg, 74%) as colourless amorphous crystals; Mp. 142-3 °C; R_f 0.50 (EtOAc/MeOH, 1:1); ^{1}H NMR (DMSO; 500 MHz, Rotomeric mixture (3:5)): δ 10.85 (2H, s, CO_2H; R_A and R_B), 8.55 (1H, s, A-2;R_A), 8.50(1H, s, A-2;R_B), 8.30(2H, s, A-8; R_A and R_B), 7.86 (4H, d, *J* 7.3, Fmoc 1,8; R_A and R_B), 7.66 (2H, d, *J* 7.3, Fmoc 4,5;R_B), 7.64 (2H, d, *J* 7.3, Fmoc 4,5;R_A), 7.51(8H, d, *J* 7.5, Bhoc;Ph-2; R_A and R_B), 7.37 (12H, m, Bhoc;Ph-3 and Bhoc;Ph-4; R_A and R_B), 7.26 (8H, m, Fmoc 2,3,6,7; R_A and R_B), 6.80 (2H, s, $CHPh_2$; R_A and R_B), 5.32 (2H, s, A-9CH_2;R_B), 5.14(2H, s, A-9CH_2;R_A), 4.36-4.27(2H, m, Fmoc-10;R_B and $C\underline{H}_2CO_2H$;R_A), 4.25-4.18 (3H, m, , Fmoc-10;R_A, $C\underline{H}_2CO_2H$;R_B and Fmoc-9), 3.52(2H, t, *J* 5.9, $C\underline{H}_2CH_2NHFmoc$;$R_A$), 3.34-3.32(4H, m, $C\underline{H}_2CH_2NHFmoc$;$R_B$ and $C\underline{H}_2NHFmoc$;R_B), 3.12 (2H, m, $C\underline{H}_2NHFmoc$;R_A);^{13}C NMR (DMSO; 125 MHz): δ 170.7 (CO_2H;R_A), 170.2 (CO_2H;R_B), 166.9(A-9$CH_2\underline{C}ON$;R_B), 166.5(A-9$CH_2\underline{C}ON$;R_A), 156.3(Fmoc;CO_2N;R_B), 156.0 (Fmoc;CO_2N;R_A), 152.2 (A-6),

151.4 (A-2), 151.0 (Bhoc;CO_2N), 149.1 (A-4), 146.6 (Fmoc-1a,8a;R_B), 146.4 (Fmoc-1a,8a;R_B), 145.2(A-8;R_A), 145.1(A-8;R_B), 143.8(Fmoc-4a and Fmoc-5a), 140.8 (Bhoc;Ph-1), 128.5(Bhoc;Ph-2), 127.6(Bhoc;Ph-3), 127.5(Fmoc-2 and Fmoc-7), 126.9 (Fmoc-3 and Fmoc-6), 126.4 (Bhoc;Ph-4), 125.1(Fmoc-4,5;R_A), 125.0 (Fmoc-4,5;R_B), 122.5 (A-5), 120.8(Fmoc-1 and Fmoc-8), 77.2($CHPh_2$), 65.4(Fmoc-10;R_B), 65.3(Fmoc-10;R_A), 47.6($\underline{C}H_2CO_2H$;R_A), 46.9($\underline{C}H_2CO_2H$;R_B), 46.7(Fmoc-9), 44.0($\underline{C}H_2CH_2NHFmoc$;$R_A$), 43.8($\underline{C}H_2CH_2NHFmoc$;$R_B$), 38.7($\underline{C}H_2NHFmoc$;$R_A$), 37.8($\underline{C}H_2NHFmoc$;$R_B$); *m/z* : 726 $[M+H]^+$; Data were in accordance with literature data.[3]

2-(N-(2-(((9H-fluoren-9-yl)methoxy)carbonylamino)ethyl)-2-(4-(benzhydryloxycarbonylamino)-2-oxopyrimidin-1(2H)-yl)acetamido)acetic acid

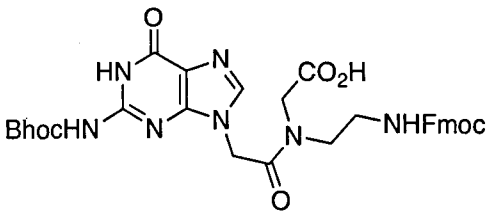


34

2-(4-(benzhydryloxycarbonylamino)-2-oxopyrimidin-1(2H)-yl)acetic acid (124 mg, 0.32 mmol) was suspended in MeCN (1 mL) and treated with NMM (72 μL, 0.64 mmol). The solution was added dropwise to pivaloyl chloride (40 μL) at –15 °C and stirred for 20 min. 2-(2-(((9H-fluoren-9-yl)methoxy)carbonylamino)ethylamino)acetic acid (111 mg, 0.32 mg) was suspended in a vigorously stirred mixture (3:7) of H_2O/MeCN and triethylamine was added dropwise until dissolution occurred. The two solutions were combined and the reaction was allowed to warm to RT while stirring for 30 min. The reaction was adjusted to pH 3 at 0 °C with 1M citric acid solution and was then allowed to warm to RT while stirring for 1 h. The reaction was evaporated to dryness, dissolved in the minimum amount of hot MeOH and added dropwise to 15 mL of rapidly stirred 20% citric acid solution. The precipitate was collected by filtration and washed on the sinter funnel with saturated brine solution. A sample was dissolved in DMF and checked for purity by HP-LC. If impurities were present, the precipitate was

dissolved in the minimum amount of hot DMF (80°C) and allowed to cool to RT. The slightly cloudy solution was filtered and the clear filtrate was added dropwise to a 15-fold volume of rapidly stirred ice-H_2O. A colourless precipitate was recovered by filtration and analysed by LC-MS to determine purity. The recovered was repeatedly treated with cycles of dissolution in hot MeOH and trituration by dropwise addition to stirred citric acid solution until all impurities had been removed. The precipitate was dried thoroughly and analysed by NMR to determine if a PNA-citrate adduct had formed. If citrate was present, the solid was ground to a fine powder and slurried with saturated brine solution for 36 hours. The slurry was filtered and washed repeatedly with H_2O and dried under reduced pressure to give the title compound (120 mg, 54%) as colourless amorphous crystals; Mp. 144-147 °C; R_f 0.60 (EtOAc/MeOH, 1:1); ^{1}H NMR (DMSO; 300 MHz, Rotomeric mixture (3:4)): δ 7.90-7.82 (6H, m, C-6 and Fmoc-1,8; R_A and R_B), 7.70-7.64 (4H, m, Fmoc-4,5; R_A and R_B), 7.45 (8H, d, *J* 7.3, Bhoc;Ph-2; R_A and R_B), 7.41-7.25 (22H, m, Bhoc;Ph-3,4 and Fmoc-2,3,6,7; R_A and R_B), 6.95 (1H, d, *J* 6.4, C-5;R_A), 6.91 (1H, d, *J* 6.4, C-5;R_B), 6.79 (2H, s, $CHPh_2$; R_A and R_B), 4.81 (2H, s, C-1CH_2;R_B), 4.62 (2H, s, C-1CH_2;R_A), 4.35-4.28(2H, m, Fmoc-10;R_B and C$\underline{H}_2CO_2H$;R_A), 4.27-4.18 (4H, m, Fmoc-10;R_A, C$\underline{H}_2CO_2H$;R_B and Fmoc-9; R_A and R_B), 3.44 (2H, m, C$\underline{H}_2CH_2NHFmoc$;$R_A$), 3.26-3.16 (4H, m, C$\underline{H}_2CH_2NHFmoc$;$R_B$ and C$\underline{H}_2NHFmoc$;R_B), 3.12 (2H, m, C$\underline{H}_2NHFmoc$;R_A); ^{13}C NMR (DMSO; 75 MHz): 171.1 (CO_2H;R_A) 170.7(CO_2H;R_B), 167.8 (C-1$CH_2\underline{C}ON$;R_A), 167.3 (C-1$CH_2\underline{C}ON$;R_B), 163.4 (C-4), 156.6(Fmoc;CO_2N;R_B), 156.4 (Fmoc;CO_2N;R_A), 155.3 (C-2), 152.7 (Bhoc;CO_2N), 151.2 (C-6), 144.2 (Fmoc-1a,8a), 141.0 (Fmoc-4a,5a), 140.7(Bhoc;Ph-1), 128.9(Bhoc;Ph-2), 128.2(Bhoc;Ph-3), 127.9(Fmoc-2 and Fmoc-7), 127.9(Fmoc-3 and Fmoc-6), 126.7 (Bhoc;Ph-4), 125.4(Fmoc-4,5), 120.4(Fmoc-2 and Fmoc-7), 95.1(C-5), 65.8 (Fmoc-10), 50.9 ($\underline{C}H_2CO_2H$;R_A), 49.8 ($\underline{C}H_2CO_2H$;R_B), 48.1($\underline{C}H_2CH_2NHFmoc$;$R_B$), 47.2($\underline{C}H_2CH_2NHFmoc$;$R_A$), 47.1 (Fmoc-9), 39.0($\underline{C}H_2NHFmoc$;$R_B$), 38.2($\underline{C}H_2NHFmoc$;$R_A$); *m/z* : 702 $[M+H]^+$; Data were in accordance with literature data.[3]

2-(*N*-(2-(((9*H*-fluoren-9-yl)methoxy)carbonylamino)ethyl)-2-(2-(benzhydryloxycarbonylamino)-6-oxo-1*H*-purin-9(6*H*)-yl)acetamido)acetic acid

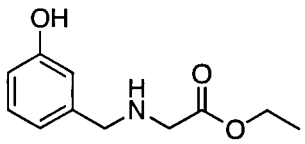


33

NMM (33 μL, 0.30 mmol) was added dropwise to 2-(2-(benzhydryloxycarbonylamino)-6-oxo-1H-purin-9(6H)-yl)acetic acid (64 mg, 0.15 mmol) suspended in a mixture of DMF:MeCN (2:3, 1.3 mL) and the resulting solution was stirred for 20 min. 2-(2-(((9H-fluoren-9-yl)methoxy)carbonylamino)ethylamino)acetic acid (52 mg, 0.15 mmol) was suspended in a mixture of H_2O/MeCN (2:3, 1.3mL) and triethylamine was added dropwise to the vigorously stirred mixture until dissolution occurred. The two solutions were combined at −15 °C then allowed to warm to RT while stirring for 15 min. The reaction was acidified to pH 4 with 20% citric acid solution and the reaction was evaporated to dryness. The crude residue was dissolved in the minimum amount of warm MeOH then added dropwise to vigorously stirred ice-H_2O. The resulting precipitate was collected by filtration and washed with saturated brine solution on the sinter. A sample was dissolved in DMF and checked for purity by LC-MS. If impurities were present, the precipitate was dissolved in the minimum amount of hot methanol and was added dropwise to a 15-fold volume of rapidly stirred ice-20% citric acid solution. A colourless precipitate was recovered by filtration and analysed by LC-MS to determine purity. The recovered was repeatedly treated with cycles of dissolution in hot MeOH and trituration by dropwise addition to stirred citric acid solution until all impurities had been removed. The precipitate was dried thoroughly and analysed by NMR to determine if a PNA-citrate adduct had formed. If citrate was present, the solid was ground to a fine powder and slurried with saturated brine solution for 36 hours. The slurry was filtered and washed repeatedly with H_2O and dried under reduced pressure to give the title compound (55mg, 49%) as colourless amorphous crystals; Mp. 221-232°C; R_f 0.40 (EtOAc/MeOH, 1:1);^{1}H NMR (DMSO; 300 MHz, Rotomeric mixture (3:5)): δ (2H, s, CO_2H; R_A and R_B), 7.87 (4H, d, *J* 7.3, Fmoc-1,8; R_A and R_B), 7.82 (1H, s, G-8;R_B), 7.81(1H, s, G-8;R_A), 7.67 (2H, t, *J* 7.3, Fmoc 4,5;R_B), 7.65 (2H, t, *J* 7.3, Fmoc 4,5;R_A), 7.45 (8H, d, *J* 7.0, Bhoc;Ph-2; R_A and R_B), 7.46-7.28 (20H, m, Bhoc;Ph-3,4; R_A and R_B and Fmoc-2,3,6,7; R_A and R_B), 6.86 (2H, s, $CHPh_2$; R_A and R_B), 5.11(2H, s, G-9CH_2;R_B), 4.94(2H, s, G-9CH_2;R_A), 4.35-4.33 (2H, m, Fmoc-10;R_B and $C\underline{H}_2CO_2H$;R_A), 4.32-4.20(4H, m, Fmoc-10;R_A and Fmoc-9; R_A and R_B), 4.01 (2H, s,

$C\underline{H}_2CO_2H;R_B$), 3.50(2H, m, $C\underline{H}_2CH_2NHFmoc;R_A$), 3.35-3.33(4H, m, $CH_2CH_2NHFmoc;R_B$ and $C\underline{H}_2NHFmoc;R_B$), 3.12 (2H, m, $C\underline{H}_2NHFmoc;R_A$); ^{13}C NMR (DMSO; 125 MHz): δ 171.6 ($CO_2H;R_B$), 170.7 ($CO_2H;R_A$), 166.8(G-9$CH_2\underline{C}ON;R_B$), 166.7(G-9$CH_2\underline{C}ON;R_A$), 156.7(Fmoc;$CO_2N;R_B$), 156.4 (Fmoc;$CO_2N;R_A$), 155.4 (G-2), 154.0 (Fmoc;CO_2N), 149.1 (G-6), 147.3 (G-4) 144.1(G-8), 141.1 (Fmoc-1a,8a), 140.0 (Fmoc-4a,5a) 140.3 (Bhoc;Ph-1), 128.9(Bhoc;Ph-2), 128.3(Bhoc;Ph-3), 127.9(Fmoc-2 and Fmoc-7), 126.9(Fmoc-3 and Fmoc-6), 127.3 (Bhoc;Ph-4), 125.0(Fmoc-4,5), 120.4 (G-5), 119.6(Fmoc-1 and Fmoc-8), 78.3($CHPh_2$), 65.7(Fmoc-10;R_B), 65.4(Fmoc-10;R_A), 47.5 ($C\underline{H}_2CH_2NHFmoc$), 47.4 (Fmoc-9), 44.8($\underline{C}H_2CO_2H;R_A$), 44.7($\underline{C}H_2CO_2H;R_B$), 37.8($\underline{C}H_2NHFmoc;R_A$), 37.4($\underline{C}H_2NHFmoc;R_A$); *m/z* : 742 $[M+H]^+$; Data were in accordance with literature data.[3]

Ethyl 2-(3-hydroxybenzylamino)acetate

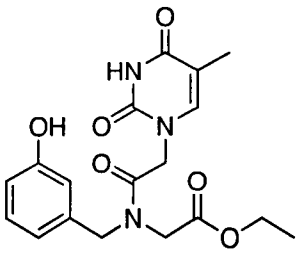


98

Ethyl glycinate (6.95 g, 50 mmol) and potassium hydroxide (2.80 g, 50.0 mmol) were stirred in EtOH (500 mL) and heated until dissolution. The reaction was cooled to RT and 3-hydroxybenzaldehyde (6.61g, 54.2 mmol) was added to the vigorously stirred solution. After 1 h stirring the reaction was placed in an ice bath and sodium borohydride (1.89 g, 50 mmol) was added portionwise slowly over 30 min. The reaction was allowed to stir for 30 min and was quenched by addition of H_2O (300 mL). The majority of MeOH was removed by evaporation and the aqueous mixture was extracted with EtOAc (5 x 100mL) and adsorbed onto silica. The crude product was purified by flash chromatography (EtOAc/petrol, 1:1) to give the title compound (4.6g, 47%) as colourless amorphous crystals; Mp 88-89°C; R_f : 0.55 (EtOAc); FT-IR ν_{max} 3449, 3301, 3048, 2936, 2885, 2841, 2596, 2841, 2596, 1729, 1586, 1486, 1378, 1222, 115, 871, 784, 695, 592; ^1H NMR (DMSO; 500 MHz): δ 9.28 (1H, s, Ph-3 OH), 7.08 (1H, t, *J* 7.5, Ph-5), 6.73 (1H, s, Ph-2), 6.69 (1H, d, *J* 7.5, Ph-6), 6.62(1H, d, *J* 8.0, Ph-4), 4.09 (2H, q, *J* 7.1, $C\underline{H}_2CH_3$), 3.62 (2H, s, Ph-3$C\underline{H}_2$), 3.28, (2H, s, $C\underline{H}_2CO_2Et$), 1.20 (3H, t, *J* 7.1, CH_3); ^{13}C NMR (DMSO; 75 MHz); δ 172.4 (CO_2Et), 158.0 (Ph-1), 141.9 (Ph-3), 129.4 (Ph-5), 118.6 (Ph-6), 115.2 (Ph-2), 114.0 (Ph-4), 60.1 ($\underline{C}H_2CH_3$), 52.4 (Ph-3CH_2),

49.7 ($\underline{C}H_2CO_2Et$), 14.5 ($C\underline{H}_2CH_3$); *m/z* 210 $[M+H]^+$: (calculated for $C_{11}H_{15}NO_3$ $[M+H]^+$ 209.1052, found 210.1138)

Ethyl 2-(N-(3-hydroxybenzyl)-2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)acetamido)acetate

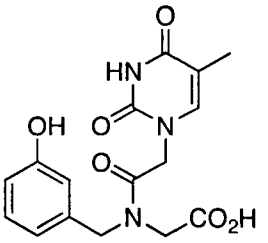


101

2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)acetic acid (626 mg, 3.4 mmol) and NMM (747 μL, 6.8 mmol) were stirred together in MeCN at RT for 15 min. The solution was cooled to 0 °C and added dropwise to pivaloyl chloride (417 μL, 3.4 mmol) and stirred vigorously for 30 min. A solution of methyl 2-(3-hydroxybenzylamino)acetate (711 mg, 3.4 mmol) dissolved in MeCN (2.15 mL) was added dropwise to the reaction and stirred for 30 min at RT. The mixture was evaporated to dryness and 6M HCl solution (5mL) was added to the crude residue and stirred for 5min. A colourless precipitate was recovered by filtration, washed with 2M HCl solution and dried via repeated azeotropic distillation of trace water with toluene to give the title compound (1.10 g, 86%) as colourless amorphous crystals; Mp. 221-222 °C; R_f 0.45 (4:1 EtOAc/MeOH); FT-IR ν_{max} 3179, 3039, 1743, 1701, 1638, 1591, 1466, 1413, 1345, 1279, 1233, 1212, 1151, 796, 712, 574; ^{1}H NMR (DMSO; 500 MHz, Rotomeric mixture (2:3)): 9.48 (1H, s, Ph-1OH;R_B), 9.39 (1H, s, Ph-1OH;R_A), 7.41 (1H, s, T-6;R_B), 7.36 (1H, s, T-6;R_A), 7.18 (1H, t, *J* 7.7, Ph-5;R_B), 7.11 (1H, t, *J* 7.7, Ph-5;R_A), 6.78 (1H, d, *J* 7.6, Ph-6;R_B), 6.74-6.70 (3H, m, Ph-6;R_A, Ph-2;R_A and Ph-4;R_A), 6.68-6.65(2H, m, Ph-2;R_B and Ph-4;R_B) 4.67 (2H, s, T-1CH_2;R_B), 4.60(2H, s, Ph-3CH_2;R_B), 4.59(2H, s, T-1CH_2;R_A) 4.60(2H, s, Ph-3CH_2;R_A), 4.10 (2H, s, CH_2CO_2Et;R_A), 4.01-3.98 (4H, m, $C\underline{H}_2CH_3$;R_A and R_B), 3.87 (2H, s, CH_2CO_2Et;R_B), 1.68 (3H, s, T-3CH_3;R_A), 1.67 (3H, s, T-3CH_3;R_B), 1.09 (6H, t, *J* 7.1, $CH_2C\underline{H}_3$; R_A and R_B);^{13}C NMR (DMSO; 75 MHz); 169.2 (CO_2Et;R_A), 169.1 (CO_2Et;R_B) 168.0 (T-1$CH_2\underline{C}ON$), 164.8 (T-4), 158.0(Ph-1;R_B) 157.9 (T-2;R_A), 151.3 (T-2) , 142.6 (T-6;R_B), 142.5 (T-6;R_A), 138.1 (Ph-3;R_A), 137.9 (Ph-3;R_B), 130.0 (Ph-5;R_B), 129.7 (Ph-5;R_A),

119.0 (Ph-6;R_A), 118.0 (Ph-6;R_B), 115.4 (Ph-4), 114.5 (Ph-2), 108.4 (T-5), 61.1 ($\underline{C}H_2CH_3$;R_A), 60.9 ($\underline{C}H_2CH_3$;R_B), 51.1 ($PhCH_2$;R_B), 50.0 ($PhCH_2$;R_A), 48.2 (T-1CH_2), 47.9 ($\underline{C}H_2CO_2Et$), 14.3 ($CH_2\underline{C}H_3$) , 12.2 (T-5CH_3); *m/z* $[M+H]^+$ 376.1: (calculated for $C_{18}H_{21}N_3NaO_6$ $[M+Na]^+$ 398.1328, found 398.1336

2-(N-(3-hydroxybenzyl)-2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)acetamido)acetic acid

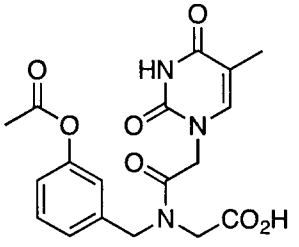


102

Methyl 2-(*N*-(3-hydroxybenzyl)-2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl) acetamido)acetate (375 mg, 1 mmol) was suspended in H_2O (3.5 mL) and cooled to 0 °C. A solution of LiOH.H_2O (715 mg dissolved in 5.3 mL of H_2O) was added and the mixture was stirred until hydrolysis was complete, as indicated by tlc. Citric acid solution (1.92 g dissolved in 6 mL of H_2O) was added and the reaction was stirred in an ice bath for 1 h and then 2 h at RT. A colourless precipitate was recovered by filtration and suspended in a minimal amount of saturated brine solution. The slurry was stirred at RT for 18h and then filtered and washed repeatedly with water . The recovered precipitate was dried under reduced pressure to give the title compound (335 mg, 97%) as colourless amorphous crystals; Mp. 268 °C(decomp.); R_f:0.40 (4:1, EtOAc/MeOH, 1-2 drops AcOH); FT-IR ν_{max} 3421, 2957, 2833, 2343, 1607, 1483, 1284, 1221, 1155, 908, 842, 781, 659, 471; ^{1}H NMR (D_2O; 500 MHz, Rotomeric mixture (2:3)): δ 11.32 (2H, bs, CO_2H; R_A and R_B), 9.51 (1H, bs, Ph-1OH;R_A), 9.43 (1H, bs, Ph-1OH;R_B), 7.41 (1H, s, T-6;R_B), 7.40 (1H, s, T-6;R_A), 7.17 (1H, t, *J* 8.0, Ph-5;R_B), 7.09 (1H, t, *J* 8.0, Ph-5;R_A), 6.75 (1H, d, *J*7.4, Ph-6;R_B), 6.71-6.69(3H, m, Ph-6;R_A, Ph-4;R_A and Ph-2;R_A), 6.68-6.50 (2H, m, Ph-4;R_B and Ph-2;R_B), 4.66 (2H, s, T-1CH_2;R_B), 4.56 (2H, s, T-1CH_2;R_A) 4.54 (2H, s, Ph-3CH_2;R_B), 4.44 (2H, s, Ph-3CH_2;R_A), 3.89 (2H, s, $C\underline{H}_2CO_2H$;R_A), 3.85 (2H, s, $C\underline{H}_2CO_2H$;R_B), 1.80 (3H, s, T-3CH_3;R_A), 1.77 (3H, s, T-3CH_3;R_B);^{13}C NMR (D_2O; 75 MHz); δ 170.6 (CO_2H;R_B), 170.3 (CO_2H;R_B), 167.9 (T-1$CH_2\underline{C}ON$), 164.8 (T-4), 158.0(Ph-1), 151.3 (T-2), 142.6 (T-6) 138.1 (Ph-3;R_B), 137.4

(Ph-3;R_B), 130.7 (Ph-5;R_B), 129.7 (Ph-5;R_A), 117.2 (Ph-6;R_A), 116.8 (Ph-6;R_B), 116.6 (Ph-4), 115.8 (Ph-4), 111.3 (Ph-2), 108.4 (T-5), 51.4($PhCH_2$), 50.9 ($\underline{C}H_2CO_2H$), 49.2 (T-1CH_2), 12.4 (T-5CH_3); *m/z* 348 $[M+H]^+$:(calculated for $C_{16}H_{17}N_3NaO_6$ $[M+Na]^+$ 370.1015, found 370.1021

2-(N-(3-acetoxybenzyl)-2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)acetamido)acetic acid

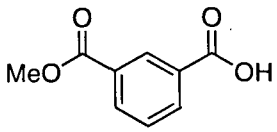


103

2-(*N*-(3-hydroxybenzyl)-2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)acetamido)acetic acid (100 mg, 0.28 mmol) was dissolved in NaOH solution (45 mg NaOH dissolved in 1mL of H_2O) and ice was added to the solution. Acetic anhydride (72 μL, 0.76 mmol) was added and the reaction was shaken vigorously by hand for 2 min. The reaction was acidified to pH 1 with 2M HCl solution then extracted with EtOAc (2 x 3mL). The combined organic phases were evaporated and dried by repeated azeotropic distillation of trace water with toluene to give the title compound (192 mg, 85%) as colourless amorphous crystals; Mp.148 °C; R_f: 0.46 (4:1 EtOAc/MeOH, 1-2 drops AcOH); FT-IR ν_{max} 3417, 2988, 2831, 2346, 2098, 1668, 1483, 1308, 1221, 1155, 1053, 908, 842, 659, 602; ^{1}H NMR (DMSO; 500 MHz, Rotomeric mixture (2:3)): δ 11.30 (2H, bs, CO_2H; R_A and R_B), 7.44 (2H, m, T-6;R_A and Ph-5;R_B), 7.38 (2H, m, T-6;R_B and Ph-5;R_A), 7.27 (1H, d, *J* 7.7, Ph-6;R_A), 7.17 (2H, m, Ph-6;R_B and Ph-2;R_B), 7.09 (1H, d, *J*7.7, Ph-4;R_A), 7.06 (2H, m, Ph-4;R_B and Ph-2;R_A), 4.70 (2H, s, T-1CH_2;R_A), 4.65 (2H, s, $PhCH_2$;R_A), 4.58 (2H, s, $PhCH_2$;R_B), 4.37 (4H, s, T-1CH_2; R_A and R_B), 4.15 ($C\underline{H}_2CO_2H$;R_A), 3.90 ($C\underline{H}_2CO_2H$;R_B), 2.28 (3H, s, $OCOCH_3$;R_A), 2.27 (3H, s, $OCOCH_3$;R_B), 1.77 (3H, s, T-5CH_3);^{13}C NMR (DMSO; 75 MHz); δ 170.7 (CO_2H) , 170.3 (Ph-1$\underline{C}O_2CH_3$), 169.5 (CH_3CO_2Ph-1), 168.2 ($CH_2\underline{C}ON$;R_B), 167.9 ($CH_2\underline{C}ON$;R_A), 164.8 (T-4), 151.3 (Ph-1;R_B), 151.1 (T-2), 150.8 (Ph-1;R_A), 144.1 (T-2), 142.6 (T-6), 139.1(Ph-3;R_A) 138.8 (Ph-3;R_B), 130.4 (Ph-5;R_B), 130.2 (Ph-5;R_A) 125.8(Ph-6;R_A), 125.3 (Ph-6;R_B), 121.8 (Ph-4), 121.2 (Ph-2), 108.5 (T-

5), 51.3 ($PhCH_2$), 48.6 (T-1CH_2), 48.0 ($\underline{C}H_2CO_2H$), 21.2 (Ph-3$CO_2\underline{C}H_3$ 12.4 (T-5CH_3); *m/z* 412 $[M+H]^+$ calculated for $C_{18}H_{19}N_3O_7$: 412.1120, found: 412.1121

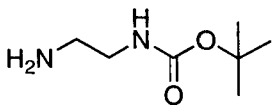
3-(methoxycarbonyl)benzoic acid



108

Dimethyl isophthalate (9.70 g, 60 mmol) was added to a solution of sodium hydroxide (2.3 g, 60 mmol) dissolved in MeOH (75 mL) and stirred at RT for 20 h. The stirred reaction was acidified to pH 1 with concentrated HCl at 0 °C and a colourless precipitate was formed. The reaction slurry was filtered and washed with MeOH to give a crude solid that was dissolved in EtOAc (300 mL) and washed with H_2O (10 x 25 mL). The organic fraction was evaporated to dryness to give a colourless residue that was dissolved in saturated Na_2CO_3 solution (50 mL) and washed with CH_2Cl_2 (2 x 25 mL). The aqueous fraction was acidified to pH 2 with concentrated hydrochloric acid to give a colourless slurry that was filtered and the recovered solid was dissolved in ethyl actate (300 mL) and dried ($MgSO_4$). Evaporation of the solution under reduced pressure gave the title compound (3.90 g, 48%) as colourless amorphous crystals; Mp 189-190 °C; R_f 0.75 (EtOAc, 1-2 drops AcOH); ^{1}H NMR (DMSO; 300 MHz): δ 8.48 (1H, s, Ph-2), 8.21- 8.16 (2H, m, Ph-4, Ph-6), 7.66 (1H, t, *J*7.8, Ph-5), 3.82 (3H, s, CH_3); ^{13}C NMR (DMSO; 75 MHz); δ 167.3 (CO_2H), 166.4 ($\underline{C}O_2Me$), 134.6 (Ph-1), 134.1 (Ph-3), 132.2 (Ph-4), 130.9 (Ph-6), 130.6 (Ph-2), 130.2 (Ph-5), 53.3 (CH_3); *m/z*: 180.3$[M+H]^+$; Data were in accordance with literature data.[11]

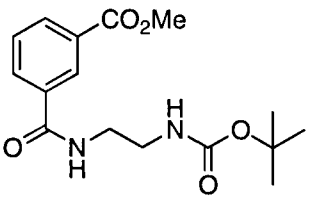
***Tert*-butyl 2-aminoethylcarbamate**



111

Boc Anhydride (874 mg, 4 mmol) dissolved in $CHCl_3$ (20 mL) was added dropwise to a stirred solution of ethylenediamine (28 mL, 40 mmol) dissolved in $CHCl_3$ (40 mL) at 0 °C over 15 min. The reaction was allowed to warm to RT while stirring for 30 min. The colourless opaque solution was washed with H_2O (1 x 10 mL), dried ($MgSO_4$) and evaporated to dryness to give the title compound (696 mg, quantitative) as a colourless oil; R_f 0.40 (CH_2Cl_2/MeOH, 1:1); ^{1}H NMR ($CDCl_3$; 500 MHz): δ 5.70 (1H, bs, NHCH$_2$CO$_2$^{t}Bu), 3.15 (2H, bm, NH$_2$CH$_2$CH$_2$), 2.78 (2H, t, *J*5.8, NH$_2$CH$_2$CH$_2$), 1.44 (9H, s, C(CH$_3$)$_3$); ^{13}C NMR (DMSO; 75 MHz); δ 156.4 (NCO_2), 78.4 (C(CH$_3$)$_3$), 44.5 (NH$_2$CH$_2$CH$_2$), 42.4 (NH$_2$CH$_2$CH$_2$), 29.1 (CH_3); *m/z* : 161 $[M+H]^+$; Data were in accordance with literature data.[12]

Methyl 3-(2-(tert-butoxycarbonylamino)ethylcarbamoyl)benzoate

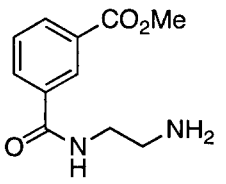


113

N-methylmorpholine (439 µL, 4 mmol) was added to a solution of 3-(methoxycarbonyl)benzoic acid (360 mg, 2 mmol) dissolved in MeCN (1.5 mL) and stirred for 15min at RT. The reaction was cooled to 0°C and pivaloyl chloride (245 µL, 2 mmol) was added dropwise to the mixture and stirred for 15 min. Tert-butyl 2-aminoethylcarbamate (320 mg, 2 mmol) was added to the reaction and allowed to stir for 25 min while the reaction warmed to RT. The solvent was removed under reduced pressure evaporation to give a crude residue that was taken up in saturated $NaHCO_3$ solution (5 mL) and extracted with EtOAc (3 x 5mL). The combined organics were washed with 10% citric acid solution (5 mL), dried ($MgSO_4$) and purified by chromatography (EtOAc/petrol, 1:1) to give the title compound (447 mg, 67%) as colourless amorphous crystals; Mp. 126-128 °C; R_f 0.58 (EtOAc/petrol, 1:1); ^{1}H NMR (DMSO; 500 MHz): δ 8.50 (1H, s, Ph-2), 8.20 (1H, d, *J* 7.8, Ph-4), 8.10 (1H, d, *J* 7.8, Ph-6), 7.56 (1H, t, *J* 7.8, Ph-5), 7.48 (1H, bs, ArCONH), 5.12 (1H, bs, NHBoc), 3.98 (3H, s, CO_2CH_3) 3.62 (2H, t, *J* 5.6, CONHCH$_2$), 3.46 (2H, t, *J* 5.6, CH$_2$NHBoc), 1.48 (9H, s, C(CH$_3$)$_3$); ^{13}C NMR (DMSO; 75 MHz): δ 167.1 (ArCONH), 166.8 (CO_2Me),

158.1 (NCO_2), 134.9 (Ph-3), 132.8 (Ph-4), 132.0 (Ph-6), 130.9 (Ph-1), 129.1 (Ph-5), 128.2 (Ph-2), 80.5 ($\underline{C}(CH_3)_3$), 52.7 ($CO_2\underline{C}H_3$), 42.7 ($ArCONH\underline{C}H_2$), 40.4 ($\underline{C}H_2NHCO_2{}^tBu$), 28.7 ($C(\underline{C}H_3)_3$); *m/z* : 323 $[M+H]^+$:(calculated for $C_{16}H_{22}N_2NaO_5$ $[M+Na]^+$ 345.1421, found 345.1412)

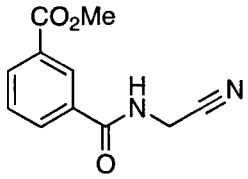
Methyl 3-(2-aminoethylcarbamoyl)benzoate



105

TFA (2 mL) was added dropwise to a solution of methyl 3-(2-(tert-butoxycarbonylamino)ethylcarbamoyl)benzoate (201 mg, 0.6 mmol) dissolved in CH_2Cl_2 (2 mL) at 0 °C and allowed to warm to RT while stirring for 2 h. The reaction was evaporated to dryness to give a crude residue that was taken up in saturated $NaHCO_3$ (5 mL) and extracted with EtOAc (3 x 5mL). The combined organics were dried (Na_2SO_4) and evaporated to give the title compound (115 mg, 53%) as a pale yellow oil; R_f 0.2 (Petrol/Acetone, 4:1); ^{1}H NMR (DMSO; 500 MHz): δ 8.50 (1H, s, Ph-2), 8.14 (1H, d, *J* 7.8, Ph-4), 8.09 (1H, d, *J* 7.8, Ph-6), 7.57 (1H, t, *J* 7.8, Ph-5), 3.93 (3H, s, CO_2CH_3), 3.62 (2H, t, *J* 5.6, $CONHC\underline{H}_2$), 3.46 (2H, t, *J* 5.6, $C\underline{H}_2NH_2$); ^{13}C NMR (DMSO; 75 MHz): δ 166.8 ($Ar\underline{C}ONH$), 166.6 ($\underline{C}O_2Me$), 135.4 (Ph-3), 132.9 (Ph-4), 132.7 (Ph-6), 130.6 (Ph-1), 129.8 (Ph-5), 128.8 (Ph-2), 53.2 ($CO_2\underline{C}H_3$), 41.5 ($ArCONH\underline{C}H_2$), 38.0 ($\underline{C}H_2NH_2$), 24.1 ($C(\underline{C}H_3)_3$); *m/z* : 223 $[M+H]^+$.

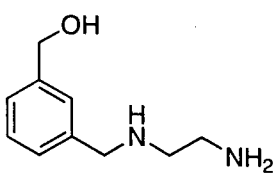
Methyl 3-(2-cyanoacetyl)benzoate



114

EDC (1.15 g, 5.5 mmol) was added to a solution of 3-(methoxycarbonyl)benzoic acid (900 mg, 5.0 mmol), hydroxybenzotriazole (740 mg, 5.5 mmol) and triethylamine (1.40 mL, 10.0 mmol) dissolved in DMF (15 mL) and stirred for 1 h at RT. A solution of aminoacetonitrile hydrochloride (463 mg, 5.0 mmol) and triethylamine (0.70 mL, 5.0 mmol) dissolved in DMF (5.0 mL) was added to the reaction and stirred for 16 h at RT. The reaction was evaporated to dryness under reduced pressure at 50 °C and the crude residue was partitioned between H_2O (100 mL) and EtOAc (50 mL). The aqueous phase was extracted with EtOAc (3 x 50 mL). The combined organics were washed with 75mL each of 10% $KHSO_4$ solution, saturated sodium bicarbonate solution and saturated brine solution, dried (Na_2SO_4) then evaporated to dryness. The residue was purified by chromatography (EtOAc/petrol, 1:1) to give 780 mg (72%) of the title compound as colourless amorphous crystals; Mp. 114-5°C; Rf 0.65 (EtOAc/petrol, 1:1); ^{1}H NMR ($CDCl_3$; 500 MHz): δ 7.32 (1H, t, *J* 7.5, Ph-5), 7.31 (1H, s, Ph-2), 7.24 (2H, d, *J* 7.5, Ph-4/Ph-6), 4.62 (4H, s, CH_2); ^{13}C NMR ($CDCl_3$; 75MHz): 167.0 ($\underline{C}OCH_2$), 166.6 ($\underline{C}O_2Me$), 133.6 (Ph-6), 133.3 (Ph-1), 132.5 (Ph-4), 130.9 (Ph-3), 129.5 (Ph-5), 128.3(Ph-2), 116.6 (CN), 65.4 (CH_2); *M/z* : 219.2 $[M+H]^+$: (calculated for $C_{11}H_{10}N_2NaO_3$ $[M+Na]^+$ 241.0584, found 241.0577)

(3-((2-aminoethylamino)methyl)phenyl)methanol)

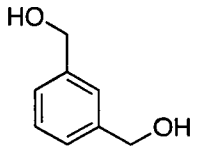


107

(3-(chloromethyl)phenyl)methanol (156 mg, 1mmol) was added dropwise *via* syringe pump (1mL per hour) to stirred EDA (667 µL, 10 mmol) at 0-5 °C. The reaction was allowed to warm to RT over 18h and was then evaporated to dryness under reduced

pressure (1 mBar) with heating to give the title compound (143 mg, 79%) as a yellow syrup; ^{1}H NMR (DMSO; 300 MHz): δ 7.28-7.21 (4H, m, Ph-2,4,5,6), 4.48 (2H, s, Ph-1C$\underline{H}_2$OH), 3.75 (2H, s, Ph-3CH_2), 2.81 (2H, t, *J* 5.9, C$\underline{H}_2$$CH_2NH_2$), 2.81 (2H, t, *J* 5.9, CH_2C$\underline{H}_2$$NH_2$); ; ^{13}C NMR (DMSO; 75 MHz): 142.7 (Ph-1), 140.7 (Ph-3), 128.1 (Ph-5), 126.4 (Ph-6), 126.4 (Ph-4), 125.0 (Ph-2), 63.1 (Ph-1CH_2), 53.3 (Ph-3CH_2), 48.9 ($\underline{C}H_2CH_2NH_2$), 40.1 $CH_2\underline{C}H_2NH_2$); *m/z* : 181 $[M+H]^+$: (calculated for $C_{10}H_{17}N_2O$ $[M+H]^+$ 181.1335, found 181.1329).

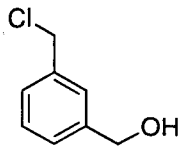
1,3-phenylenedimethanol



120

Sodium borohydride (76 mg, 2.0 mmol) was added to a stirred solution of isophthaldehyde (134 mg, 1.0 mmol) dissolved in MeOH (2.0 mL) at RT. After 24 h of stirring the reaction was cooled to 0 °C and 2M sodium hydroxide solution (2.0 mL) was added to the reaction. The reaction was allowed to warm to RT over 20 min and the MeOH was removed by reduced pressure evaporation from the mixture. The aqueous solution was extracted with EtOAc (3 x 5.0 mL) and the combined organics were dried ($MgSO_4$) and evaporated under reduced pressure to give 130 mg (94%) of the title compound as a pale yellow oil; R_f 0.65 (EtOAc/petrol, 1:1); ^{1}H NMR ($CDCl_3$; 500 MHz): δ 7.32 (1H, t, *J* 7.5, Ph-5), 7.31 (1H, s, Ph-2), 7.24 (2H, d, *J* 7.5, Ph-4/Ph-6), 4.62 (4H, s, CH_2); ^{13}C NMR ($CDCl_3$; 125 MHz): 141.6 (Ph-1, Ph-3), 129.1 (Ph-4, Ph-6), 126-6 (Ph-2), 125.9 (Ph-5), 65.4 (CH_2); *M/z* : 139 $[M+H]^+$.

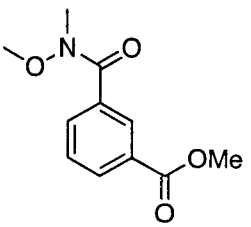
(3-(chloromethyl)phenyl)methanol



121

Concentrated hydrochloric acid (8.0 mL) was added to a stirred solution of 1,3-phenylenedimethanol (2.20 g, 15.9 mmol) dissolved in toluene (79.0 mL) at RT. The reaction became an intense red colour upon addition of the acid and this slowly changed to yellow as the reaction was stirred for 24 h at RT. The reaction was repeatedly washed with aliquots of sodium hydrogen carbonate solution (5.0 mL) until there was no further effervescence observed and then extracted with CH_2Cl_2 (3 x 10.0 mL). The combined organics were dried ($MgSO_4$) and evaporated under reduced pressure to give 1.97 g (79%) of the title compound as a pale yellow oil; R_f 0.75 (EtOAc/petrol, 1:1); ^{1}H NMR ($CDCl_3$; 500 MHz): ^{1}H NMR ($CDCl_3$; 500 MHz): δ 7.36-7.34 (2H, m, Ph-2, Ph-5), 7.33-7.31 (2H, m, Ph-4, Ph-6), 4.71 (2H, s, CH_2Cl), 4.59 (2H, s, CH_2OH);^{13}C NMR ($CDCl_3$; 125 MHz): 140.7 (Ph-1, Ph-3), 128.2 (Ph-6), 127.6 (Ph-4), 126.3 (Ph-2), 126.2 (Ph-2), 65.4 (CH_2OH), 45.4 (CH_2Cl); *m/z* : 157.0, 159.1$[M+H]^+$; Data were in accordance with literature data.[13]

Methyl 3-(methoxy(methyl)carbamoyl)benzoate

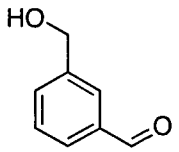


133

Isobutylchloroformate (132 μL, 1 mmol) was added dropwise to a stirred solution of 3-(methoxycarbonyl)benzoic acid (180 mg, 1 mmol) and N-methylmorpholine (110 μL, 1 mmol) dissolved in THF (6 mL) at –20°C. After 10 min, triethylamine (153 μL, 1.1 mmol) and then a solution of *N,O*-dimethylhydroxylamine hydrochloride (98 mg, 1 mmol) dissolved in DMF (2.4 mL) were added dropwise to the reaction and allowed to warm to RT while stirring for 30 min. A colourless precipitate was removed by filtration and the filtrate was diluted with H_2O (10 mL) and acidified with 1M hydrochloric acid (1mL) before removing the THF component under reduced pressure

evaporation. The aqueous mixture was extracted with EtOAc (5 x 20 mL) and the combined organics were dried (Na_2SO_4) and to give the title compound (166 mg, 60%) as a colourless syrup; R_f 0.44 (EtOAc/petrol, 1:1); ^{1}H NMR (DMSO; 300 MHz): δ 8.35 (1H, s, Ph-2), 8.13 (1H, d, *J* 7.8, Ph-6), 7.88 (1H, d, *J* 7.8, Ph-4), 7.52 (1H, t, *J* 7.8, Ph-5), 3.93 (3H, s, CO_2CH_3), 3.55 (3H, s, $NCH_3OC\underline{H}_3$), 3.38 (3H, s, $NC\underline{H}_3OCH_3$);^{13}C NMR (DMSO; 75 MHz): 169.2 ($\underline{C}O_2CH_3$), 166.8 ($\underline{C}ON(CH_3)OCH_3$), 134.8 (Ph-1), 132.9 (Ph-4), 131.9 (Ph-6), 130.4 (Ph-3), 129.7 (Ph-2), 128.6 (Ph-5), 61.5 ($CON(CH_3)O\underline{C}H_3$), 52.6 ($CO_2\underline{C}H_3$), 33.9 ($CON(\underline{C}H_3)OCH_3$); *m/z* : 224 $[M+H]^+$.

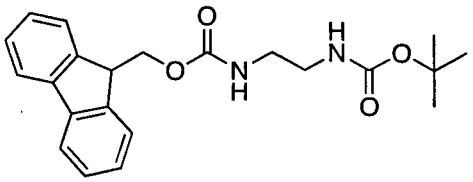
3-(hydroxymethyl)benzaldehyde



131

Sodium borohydride (150 mg, 3.9 mmol) was added in one portion to a stirred slurry of isophthalaldehyde (1.5 g, 11.2 mmol) suspended in THF (10 mL) at 0 °C and allowed to warm to RT over 1 h. The slurry initially dissolved to give a rose coloured solution that effervesced and then become yellow and opaque. The reaction was evaporated to dryness and the crude residue was treated with 1M sodium hydroxide solution (10 mL), 1M HCl solution (10 mL) and extracted with EtOAc (3 x 10mL). The combined organics were dried ($MgSO_4$), concentrated under reduced pressure and adsorbed on silica. The crude mixture was purified by chromatography (EtOAc/petrol, 1:1) to give the title compound (733mg, 49%) as a yellow oil; R_f 0.75 (EtOAc/petrol, 1:1); ^{1}H NMR ($CDCl_3$; 500 MHz): δ 9.90 (1H, s, ArCHO), 7.80 (1H, s, Ph-2), 7.72 (1H, d, *J* 7.5, Ph-6), 7.57 (1H, d, *J* 7.5, Ph-4), 7.46 (1H, t, *J* 7.5, Ph-5), 4.70 (2H, s, CH_2);^{13}C NMR ($CDCl_3$; 125 MHz); δ 192.3 (CHO), 141.9(Ph-1), 136.6 (Ph-4), 132.8 (Ph-6), 129.2(Ph-3), 127.2 (Ph-2), 128.2(Ph-5), 64.9(CH_2); *m/z* : 136.2 $[M+H]^+$; Data were in accordance with literature data.[14]

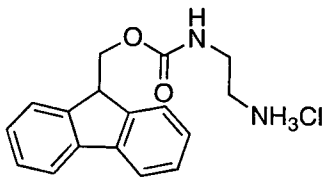
***N-Tert*butoxycarbonyl *N'*-((9H-fluoren-9-yl)methylcarbamate) ethane 1,2-diamine**



134

A solution of FmocOSuccinimide (1.41 g, 4.2mmol) in MeCN (8.4 mL) was added to a solution of *tert*-butyl 2-aminoethylcarbamate (670 mg, 4.2 mmol) dissolved in H_2O (4.2 mL) at pH 9.0. The reaction was stirred at RT for 10 min and filtered to recover a colourless solid. The filtrant was washed with H_2O and MeCN and then dissolved in a mixture of $CHCl_3$/MeOH (8.4 mL, 1:1) and co-evaporated with toluene (5.0 mL). The residue was dried under reduced pressure to afford the title compound (816 mg, 53%) as colourless amorphous crystals; Mp.158-159°C; R_f: 0.65 (EtOAc/Petrol, 1:1); ^{1}H NMR ($CDCl_3$; 500 MHz): δ 7.75 (2H, d, *J* 9.0, Fmoc-1,8), 7.60 (2H, t, *J* 9.0, Fmoc-4,5), 7.40 (2H, t, *J* 9.0, Fmoc-2,7), 7.30 (2H, t, *J* 9.0, Fmoc-3,6), 4.41 (2H, d, *J* 7.8, Fmoc-10), 4.18 (1H, t, *J* 7.8, Fmoc-9), 3.28-3.24 (4H, m, C$\underline{H_2}$C$\underline{H_2}$), 1.45 (9H, s, $C(CH_3)_3$); ; ^{13}C NMR (DMSO; 125 MHz): 156.5 (CO_2NH), 144.0 (Fmoc-1a,8a), 142.4 (Fmoc-4a,5a), 127.9 (Fmoc-2,7), 127.0 (Fmoc-3,6), 125.0 (Fmoc-4,5), 119.3 (Fmoc-1,8), 67.0 (Fmoc-10), 47.3 (Fmoc-9), 41.8 ($CH_2NHFmoc$), 40.3 (CH_2NHBoc), 29.0 ($C(\underline{C}H_3)_3$); *m/z* : 383.3 $[M+H]^+$; Data were in accordance with literature data.[15]

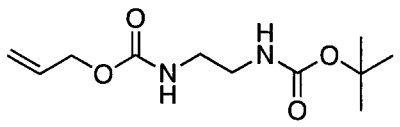
(9*H*-fluoren-9-yl)methyl 2-aminoethylcarbamate hydrochloride salt



132

Acetyl chloride (235 µL, 3.0 mmol) was dissolved in MeOH (1.22 mL) at 0°C and allowed to stir for 30 min, The reaction was allowed to warm to RT *N-tert* butoxycarbonyl *N'-(*(9H-fluoren-9-yl)methylcarbamate) ethane 1,2-diamine (383 mg, 1.0mmol) was added to the solution. The reaction was stirred for 5h and then evaporated to dryness. The crude foam was used immediately without further purification; *m/z* : 283 $[M+H]^+$;

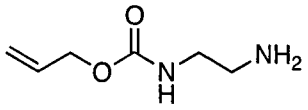
N-*tert*-butyl carbamate N'-allyl carbamate ethylenediamine



136

A solution of allylchloroformate (17.65 mL, 162 mmol) dissolved in CH_2Cl_2 (800 mL) was added dropwise to a stirred solution of tert-butyl 2-aminoethylcarbamate (25.91 g, 162 mmol) and triethylamine (22.5 mL, 162 mmol) dissolved in CH_2Cl_2 (800 mL) at 0 °C. The reaction was allowed to warm to RT while stirring for 24 h and was washed with 10% citric acid solution (3 x 50 mL), saturated $NaHCO_3$ solution (2 x 50 mL) and dried (Na_2SO_4). The organic fraction was evaporated to dryness to give the title compound (37.49 g, 89 %) as colourless amorphous crystals; Mp. 111-2°C; R_f 0.83 (EtOAc/petrol, 1:1); ^{1}H NMR (DMSO; 300 MHz): δ 5.92 (1H, m, Allyl-2), 5.37 (1H, bs, NH), 5.26 (1H, d, *J* 17.2, Allyl-3^T), 5.22 (2H, d, *J* 10.7, Allyl-3^C), 5.05 (2H, d, *J* 5.2, Allyl-1), 3.29 (4H, bm, CH_2CH_2), 1.45 (9H, s, $C(CH_3)_3$); ^{13}C NMR (DMSO; 75 MHz); δ 157.1 ($NH\underline{C}O_2Allyl$), 156.9 ($NH\underline{C}O_2{}^tBu$), 133.2 (Allyl-2), 118.1 (Allyl-3), 79.9 ($\underline{C}(CH_3)_3$), 64.8 (Allyl-1), 41.8 ($AllocNH\underline{C}H_2$), 41.0 ($Boc\ NH\underline{C}H_2$); *m/z* : 245 $[M+H]^+$. Data were in accordance with literature data.[16]

Allyl 2-aminoethylcarbamate

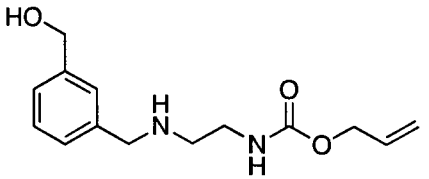


137

Acetyl chloride (235 μL, 3.0 mmol) was dissolved in MeOH (1.22 mL) at 0°C and allowed to stir for 30 min, The reaction was allowed to warm to RT and *N*-tert-butyl carbamate *N*'-allyl carbamate ethylenediamine (244 mg, 1.0mmol) was added to the solution. The reaction was stirred for 5 h and then evaporated to dryness. The crude residue was taken up in saturated $NaHCO_3$ solution (3 mL) and extracted with EtOAc (5 x 3mL). The combined organics were dried (Na_2SO_4) and evaporated to dryness to give the title compound (142 mg, 99%) as a pale yellow foam; R_f 0.65 (EtOAc/MeOH, 1:1); ^{1}H NMR (D_2O; 500 MHz): δ 5.93 (1H, m, Allyl-2), 5.26 (1H, d, *J* 17.2, Allyl-3^T), 5.22

(2H, d, *J* 10.5, Allyl-3^{C}), 4.51 (2H, d, *J* 4.6, Allyl-1), 3.38 (2H, t, *J* 5.5, AllocNHC$\underline{H}_2$), 3.07 (2H, t, *J* 5.5, $NH_2C\underline{H}_2$); ^{13}C NMR (DMSO; 75 MHz); δ 156.3 ($\underline{C}O_2N$), 133.7 (Allyl-2), 117.4 (Allyl-3), 64.8 (Allyl-1), 38.8 (AllocNH$\underline{C}H_2$), 38.2 ($\underline{C}H_2NH_2$); *m/z* : 145 $[M+H]^+$; Data were in accordance with literature data.[16]

Allyl 2-(3-(hydroxymethyl)benzylamino)ethylcarbamate

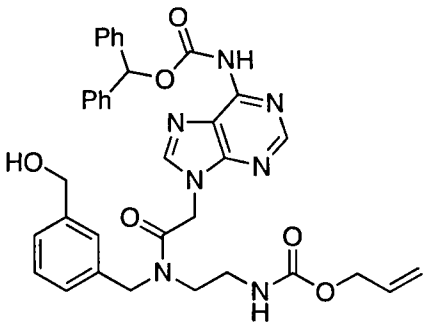


139

A solution of allyl 2-aminoethylcarbamate (5.41 g, 30 mmol) dissolved in saturated $NaHCO_3$ solution (10 mL) was added to 3-(hydroxymethyl)benzaldehyde (3.72 g, 27 mmol) and stirred at RT for 3 h. The reaction was evaporated to dryness and a cold (0-5 °C) solution of $NaBH_4$ (1.14 g. 30 mmol) dissolved in H_2O (30 mL) was added to the residue. The reaction was allowed to stir for 30 min while warming to RT and was then chilled to 0 °C and acidified to pH 0 with 2M HCl. The black suspension was neutralised and filtered through a thin pad of celite and the filtrate was evaporated to dryness. MeOH (50 mL) was added to the colourless residue and stirred vigorously at RT until a colourless slurry had formed. The slurry was filtered and the recovered insoluble material was extracted with MeOH (2 x 50 mL) in an identical manner. The combined organic extracts were evaporated to dryness and the residue was heated at reflux in MeCN briefly with the dropwise addition of H_2O until colourless crystals had precipitated. The colourless insoluble crystals were removed by filtration and the filtrate evaporated to give a pale yellow oil that was purified via ion-exchange chromatography (method x) and then flash silica chromatography (EtOAc/MeOH, 7:3) to give the title compound (6.43 g, 90%) as a colourless oil; R_f 0.35 (EtOAc/MeOH, 7:3); ^{1}H NMR (DMSO; 300 MHz): δ 7.36-7.22 (4H, m, Ph-2, Ph-4, Ph-5, Ph-6), 5.92 (1H, ddt, *J* 17.3, 11.8, 5.6, allyl-2), 5.48 (1H, bs, NHAlloc), 5.32 (1H, d, *J* 17.3, allyl-3^{T}), 5.23 (1H, d, *J* 11.7, allyl-3^{C}), 4.67 (2H, s, C$\underline{H}_2$OH), 4.57 (2H, d, *J* 5.6, allyl-1), 3.79 (2H, s, ArC$\underline{H}_2$NH), 3.29 (2H, dt, *J*5.9, C$\underline{H}_2$NHAlloc), 2.78 (2H, t, *J* 5.6, C$\underline{H}_2$CH$_2$NHAlloc); ^{13}C NMR (DMSO; 75 MHz); δ 156.9 (NH$\underline{C}O_2$), 141.9 (Ph-3), 139.9 (Ph-1), 133.3 (Allyl-2), 129.1 (Ph-2), 127.9 (Ph-5), 127.2 (Ph-6), 126.3 (Ph-4), 118.1 (allyl-3), 66.0

(CH_2OH), 65.3 (allyl-1), 53.7 (CH_2NH), 48.7 ($C\underline{H}_2NHAlloc$), 40.8 ($C\underline{H}_2CH_2NHAlloc$); *m/z* : 265.2 $[M+H]^+$: (calculated for $C_{14}H_{21}N_2O_3$ $[M+H]^+$ 265.1558, found 265.1547).

Benzhydryl 9-(2-((2-(allylcarboxyamino)ethyl)(3-(hydroxymethyl)benzyl)amino)-2-oxoethyl)-9H-purin-6-ylcarbamate

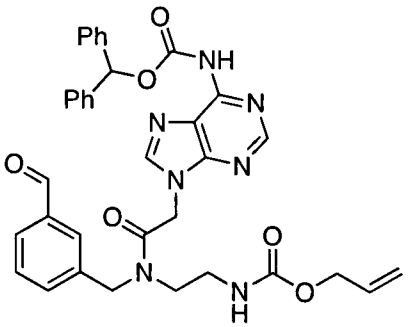


140

NMM (623 μL, 5.7 mmol) was add to a suspension of 2-(6-(benzhydryloxycarbonylamino)-9*H*-purin-9-yl)acetic acid (1.15 g, 2.85 mmol) and stirred at RT until a solution was obtained. The reaction was cooled to –15°C and pivaloyl chloride (349 μL, 2.85 mmol) was added dropwise and allowed to stir for 30 min at –15 °C. A solution of allyl 2-(3-(hydroxymethyl)benzylamino)ethylcarbamate (502 mg, 1.90 mmol) dissolved in 50% aqueous MeCN (25 mL) was added to the reaction and stirred for 30 min while the reaction was allowed to warm to RT. The yellow solution was evaporated to dryness and the crude residue was taken up in a minimal volume of $MeOH/CH_2Cl_2$ (7.5:92.5) and adsorbed onto silica. Purification by chromatography (MeOH/EtOAc, 3:7) gave of the title compound (1.18g, 96%) as a colourless oil; R_f 0.30 (EtOAc); FT-IR ν_{max} 3692, 2939, 1754, 1711, 1662, 1615, 1588, 1538, 1469, 1215, 1157, 987, 732, 642;^{1}H NMR (CD_2Cl_2; 300 MHz, Rotomeric mixture (4:5)): δ 9.43 (2H, bs, Bhoc;CO_2NH; R_A and R_B), 8.61 (2H, s, A-2; R_A and R_B), 8.05 (1H, s, A-8;R_A), 8.00(1H, s, A-8;R_B), 7.45 (8H, d, *J* 7.3, Bhoc;Ph-2; R_A and R_B), 7.40-7.19 (18H, m, Bhoc;Ph-3, Bhoc;Ph-4, Ph-2, Ph-5, Ph-6; R_A and R_B), 7.16-7.08 (1H, m, Ph-4;R_A and Ph-4;R_B), 6.93 (2H, s, $CHPh_2$; R_A and R_B), 6.01-5.81 (3H, m, NHAlloc;R_A and allyl-2; R_A and R_B), 5.57 (1H, bs, NHAlloc;R_B), 5.26 (2H, d, *J* 17.3, allyl-3^T; R_A and R_B), 5.19 (2H, d, *J* 11.4, allyl-3^T; R_A and R_B), 5.13 (2H, s, A-9CH_2;R_A), 4.99 (2H, s, A-9CH_2;R_B), 4.69-4.50 (12H, m, $C\underline{H}_2OH$, $PhC\underline{H}_2N$ and allyl-1; R_A and R_B), 3.47 (4H, m, $CH_2NHAlloc$; R_A and R_B), 3.30 (4H, $C\underline{H}_2CH_2NHAlloc$; R_A and R_B);^{13}C NMR

(CD_2Cl_2; 75 MHz): 167.7 (A-9CH_2CON;R_A), 167.0 (A-9CH_2CON;R_B), 157.5 (alloc;NCO_2;R_A), 157.2 (alloc;NCO_2;R_B), 153.2(A-2), 152.3(A-6;R_A), 152.2(A-6;R_B), 151.1(A-4), 150.0(Bhoc;NCO_2), 145.2(A-8;R_B), 145.0(A-8;R_A), 143.5(Bhoc;Ph-1), 140.9(Ph-1), 137.6 (Ph-3;R_B), 136.5(Ph-3;R_A), 133.9(allyl-2;R_B), 133.7(allyl-2;R_B), 129.9 (Ph-4;R_B), 129.4 (Ph-4;R_A), 129.3 (Bhoc;Ph-2), 128.7(Bhoc;Ph-3) 127.7(Bhoc;Ph-4 and Ph-5), 127.3 (Ph-6;$R_{A/B}$), 127.0 (Ph-6;$R_{A/B}$), 126.2 (Ph-2;$R_{A/B}$) 125.4(Ph-2;$R_{A/B}$), 121.9 (A-5), 118.2 (allyl-3;R_A), 117.8 (allyl-3;R_B), 79.3($CHPh_2$), 66.4(allyl-1;R_A), 66.1(allyl-1;R_B), 65.2(CH_2OH), 52.2 ($PhCH_2$;R_A), 50.0($PhCH_2$;R_B), 47.7(CH_2NHAlloc;R_B), 46.9(CH_2NHAlloc;R_A), 45.2(A-9CH_2;R_B), 44.9(A-9CH_2;R_A), 39.8($\underline{C}H_2CH_2$NHAlloc;R_A), 39.7($\underline{C}H_2CH_2$NHAlloc;R_B); *m/z* 650 $[M+H]^+$:(calculated for $C_{35}H_{35}N_7NaO_6$ $[M+Na]^+$ 672.2524, found 672.2541)

Benzhydryl 9-(2-((2-(allylcarboxyamino)ethyl)(3-formylbenzyl)amino)-2-oxoethyl)-9H-purin-6-ylcarbamate



144

A solution of benzhydryl 9-(2-((2-(allylcarboxyamino)ethyl)(3-(hydroxymethyl)benzyl)amino)-2-oxoethyl)-9*H*-purin-6-ylcarbamate (365 mg, 0.56 mmol) and Dess-Martin periodinane (369 mg, 0.84 mmol) dissolved in CH_2Cl_2 (4.0 mL) was stirred at RT for 30 min. The reaction was poured into a vigorously stirred mixture of saturated $Na_2S_2O_3$ and saturated Na_2CO_3 solutions (16.0 mL, 1:1) and stirred at RT for 30 min. The quenched reaction was extracted with dichloromethane (2 x 10.0 mL) and the combined organics were dried (Na_2SO_3) and evaporated to dryness. Purification of the crude solid with chromatography (EtOAc) gave 217 mg (60%) of the title compound as colourless amorphous crystals; Mp. 82-3°C; R_f 0.40 (EtOAc); FT-IR ν_{max} 3689, 2940, 1754, 1703, 1665, 1614, 1587, 1467, 1214, 988, 732, 700, 642;^{1}H NMR ($CDCl_3$; 500 MHz, Rotomeric mixture (4:5)): δ 10.04 (1H, s, CHO;R_A), 9.89 (1H, CHO;R_B), 8.78 (2H, bs, Bhoc;NH; R_A and R_B), 8.73 (1H, s, A-2;R_B), 8.70 (1H, s,

A-2;R_A), 8.09 (1H, s, A-8;R_B), 8.06 (1H, s, A-8;R_A), 7.85 (1H, d, *J* 7.4, Ph-6;R_A), 7.79 (1H, d, *J* 7.4, Ph-6;R_B), 7.79 (1H, s, Ph-2;R_B), 7.74 (1H, s, Ph-2;R_A), 7.60 (1H, t, *J* 7.4, Ph-5;R_A), 7.56 -7.47 (2H, m, Ph-5;R_B, Ph-4;R_B, Ph-4;R_A), 7.42 (8H, d, *J* 7.3, Bhoc;Ph-2; R_A and R_B), 7.35 (8H, t, *J* 7.3, Bhoc;Ph-3; R_A and R_B), 7.30 (4H, t, *J* 7.3, Bhoc;Ph-4; R_A and R_B), 7.00 (2H, s, Bhoc;$CHPh_2$; R_A and R_B) 5.88 (2H, ddt, *J* 17.0, 11.3, 5.3, allyl-2; R_A and R_B), 5.88 5.73 (1H, bs, NHAlloc;R_B), 5.31 (1H, bs, NHAlloc;R_A), 5.28 (1H, d, *J* 17.3, allyl-3^T; R_B), 5.24 (1H, d, *J* 17.3, allyl-3^T; R_A), 5.17 (3H, m, allyl-3^C ; R_A, R_B and A-9CH_2;R_A), 5.02 (2H, s, A-9CH_2;R_B), 4.82 (2H, s, PhCH_2;R_A), 4.69 (2H, s, PhCH_2;R_B), 4.56 (2H, d, *J* 5.3, allyl-1;R_B), 4.54 (2H, d, *J* 5.3, allyl-1;R_A), 3.58 (4H, m, C$\underline{H_2}$NHAlloc; R_A and R_B), 3.44 (2H, m, C$\underline{H_2}CH_2$NHAlloc;R_B), 3.40 (2H, m, C$\underline{H_2}CH_2$NHAlloc;R_A);^{13}C NMR ($CDCl_3$; 125 MHz): δ 191.9(CHO;R_A), 191.6(CHO;R_B), 166.8(A-9$CH_2\underline{C}$ON;R_A), 166.3(A9$CH_2\underline{C}$ON;R_B), 156.7 (Alloc;NCO_2), 152.8 (A-2), 151.5(A-6;R_B), 151.4(A-6;R_A), 150.1(Bhoc;NCO_2), 149.1 (A-4), 144.0 (A-8;R_B), 143.8 (A-8;R_A), 139.6 (Bhoc;Ph-1), 137.7 (Ph-3;R_A), 137.2 (Ph-1;R_B), 136.8 (Ph-3;R_B), 136.7 (Ph-1;R_A), 134.1(Ph-4;R_B), 132.7 (Allyl-2;R_A), 132.4 (Allyl-2;R_B), 132.1(Ph-4;R_A), 130.1(Ph-5;R_A), 130.0(Ph-6;R_A), 129.6 (Ph-5;R_B, Ph-6;R_B), 128.5 (Ph-2;R_A, Bhoc;Ph-3), 128.0 (Bhoc;Ph-4), 127.3 (Bhoc;Ph-2), 126.8 (Ph-2;R_B), 121.3 (A-5;R_A), 121.2 (A-5;R_B), 118.1(allyl-3;R_B), 117.1 (allyl-3;R_A), 78.7(Bhoc;$CHPh_2$), 65.9 (allyl-1;R_B), 65.6(allyl-1;R_A), 51.9 (PhCH_2;R_A), 49.1(PhCH_2;R_B), 46.9 (CH_2NHAlloc;R_A), 46.5 (CH_2NHAlloc;R_B) 44.1(A9-CH_2;R_A), 43.9(A9-CH_2;R_B), 39.1($\underline{C}H_2CH_2$NHAlloc;R_B), 38.8 ($\underline{C}H_2CH_2$NHAlloc;R_A);); *m/z* 648 $[M+H]^+$:(calculated for $C_{35}H_{33}N_7NaO_7$ $[M+Na]^+$ 670.2385, found 670.2374).

5.3 General Solid Phase Synthesis Protocols

5.3.1Resin Loading

Method A : 2-Chlorotrityl Chloride Resin

2-Chlorotrityl chloride resin (26 mg, 36 μMol) was swelled in CH_2Cl_2 for 10 min and isolated by filtration. A solution of T-PNA monomer (72 μMol) and DIPEA (180 μMol) in CH_2Cl_2 (2.0 mL) was added to the dry resin and the mixture was agitated for 1 h. The resin was filtered and washed repeatedly with DMF (5 × 2.0 mL × 2 min), (MeOH/DIPEA/ CH_2Cl_2, 1:1:8) (2 × 2.0 mL × 2 min), and DMF (5 × 2.0 mL × 2 min) with orbital agitation. The resin was treated with a solution of 20% piperidine dissolved in DMF (5 × 2.0 mL × 2 min) and washed repeatedly with DMF (5 × 2.0 mL × 2 min), CH_2Cl_2 (5 × 2.0 mL × 2 min) and MeOH (5 × 2.0 mL × 2 min) and dried over KOH pellets under reduced pressure.

Method B : Rink Amide Novogel ™

Rink Amide Novogel ™ resin (30 mg, 15 μMol) was swelled in DMF for 1 h and isolated by filtration. The resin was treated with a solution of 20% piperidine dissolved in DMF (5 × 2.0 mL × 2 min) and washed repeatedly with DMF (5 × 2.0 mL × 2 min). A positive Kaiser test confirmed successful deprotection. HATU (70 μMol) was added to a solution of C-PNA monomer (75 μMol) and PNA base solution (150 μMol) dissolved in DMF (1.0 mL) and agitated for 1 min before being added to the dry resin. The mixture was agitated for 1 h. The resin was isolated by filtration, washed repeatedly with DMF (5 × 2.0 mL × 2 min), CH_2Cl_2 (5 × 2.0 mL × 2 min) and MeOH (5 × 2.0 mL × 2 min) and dried over KOH pellets under reduced pressure.

5.3.2 Manual Solid Phase Peptide synthesis

General Swelling protocol

The resin was swelled in DMF for 1 h, isolated by filtration and treated with a solution of 20% piperidine in DMF (5 × 2.0 mL × 2 min) with gentle orbital agitation. After repeated washes with DMF (5 × 2.0 mL × 2 min), a positive Kaiser test confirmed successful Fmoc deprotection.

Fmoc Amino Acid Coupling

Method C: Single coupling with HCTU activation

Fmoc amino acid (75 μMol) and HCTU (67 μMol) were dissolved in DMF (1.0 mL). DIPEA (150 μMol) was added and the solution was added to the dry resin (15 μMol). The mixture was gently agitated for 1 h and the resin was isolated by by filtration and washed with DMF (5 × 2.0 mL × 2 min). A negative Kaiser test signified a complete reaction. The resin was treated with a solution of 20% piperidine in DMF (5 × 2.0 mL × 2 min) and then washed repeatedly with DMF (5 × 2.0 mL × 2 min). A positive Kaiser test confirmed successful Fmoc deprotection.

Method D: Double coupling with HCTU activation

Fmoc amino acid (75 μMol) and HCTU (67 μMol) were dissolved in DMF (1.0 mL). DIPEA (150 μMol) was added and the solution was added to the dry resin (15 μMol). The mixture was gently agitated for 1 h and the resin was isolated by by filtration and washed with DMF (5 × 2.0 mL × 2 min). The process was repeated. A negative Kaiser test signified a complete reaction. The resin was treated with a solution of 20% piperidine in DMF (5 × 2.0 mL × 2 min) and then washed repeatedly with DMF (5 × 2.0 mL × 2 min). A positive Kaiser test confirmed successful Fmoc deprotection.

Method E : Kaiser Colour Test

A small number of beads were collected and washed with EtOH. The beads were treated with two drops each of phenol (80% in EtOH, x/v), potassium cyanide (0.001 M, 1 mL, in pyridine, 98 mL) and ninhydrin (5% in EtOH, w/v). The beads were heated at 120 °C for 3 min and the colour change was observed; A blue solution indicated a positive test, indicating the presence of an NH_2 group.

5.3.3 Manual solid phase PNA synthesis

Method F: Single coupling with HATU activation

PNA monomer (75 μMol) and HATU (67 μMol) were dissolved in DMF (1.0 mL). DIPEA (150 μMol) was added and the solution was added to the dry resin (15 μMol). The mixture was gently agitated for 1 h and the resin was isolated by by filtration and washed with DMF (5 × 2.0 mL × 2 min). A negative Kaiser test signified a complete reaction. The resin was treated with a solution of 20% piperidine in DMF (1 × 2.0 mL × 3 min) and then agitated vigourously and repeatedly with DMF (2 × 2.0 mL × 1 min). The resin was treated once more with 20% piperidine in DMF (1 × 2.0 mL × 3 min) and

again agitated vigourously and repeatedly with DMF (6 × 2.0 mL × 1 min). A positive Kaiser test confirmed successful Fmoc deprotection.

Method G: Double coupling with HATU activation

PNA monomer (75 μMol) and HATU (67 μMol) were dissolved in DMF (1.0 mL). DIPEA (150 μMol) was added and the solution was added to the dry resin (15 μMol). The mixture was gently agitated for 1 h and the resin was isolated by by filtration and washed with DMF (5 × 2.0 mL × 2 min). The process was repeated. A negative Kaiser test signified a complete reaction. The resin was treated with a solution of 20% piperidine in DMF (1 × 2.0 mL × 3 min) and then agitated vigourously and repeatedly with DMF (2 × 2.0 mL × 1 min). The resin was treated once more with 20% piperidine in DMF (1 × 2.0 mL × 3 min) and again agitated vigourously and repeatedly with DMF (6 × 2.0 mL × 1 min). A positive Kaiser test confirmed successful Fmoc deprotection.

5.3.4 Manual solid phase linker integration methods

Method H: Linker 2 Attachment; Reductive amination

Dried peptidyl resin (15 μMol) was swelled in CH_2Cl_2 (2.0 mL) with orbital agitation for 30 min. The resin was drained and a solution of benzhydryl 9-(2-((2-(allylcarboxyamino)ethyl)(3-formylbenzyl)amino)-2-oxoethyl)-9H-purin-6-ylcarbamate (1.1 equiv.) dissolved in CH_2Cl_2 (2.0 mL) was added to the dry resin and gently agitated for 1 h. AcOH (5.0 equiv) and $NaBH(OAc)_3$ were added sequentially to the mixture and agitated gently for 3 h, with venting of the reaction tube every 15 min. The resin isolated by filtration and repeatedly washed with CH_2Cl_2 (5 × 2.0 mL × 2 min), MeOH (5 × 2.0 mL × 2 min), DMF (5 × 2.0 mL × 2 min) and CH_2Cl_2 (5 × 2.0 mL × 2 min). A negative Kaiser test and analytical LC-MS indicated that the reaction was complete. The resin was isolated by filtration and a solution of Boc anhydride (10 equiv.) dissolved in CH_2Cl_2 was added to the resin and agitated gently for 1 h. The resin was isolated by filtration and washed repeatedly with CH_2Cl_2 (5 × 2.0 mL × 2 min) and MeOH (5 × 2.0 mL × 2 min) and dried over KOH pellets under reduced pressure.

Method I: Linker 2 Deprotection; Alloc deprotection

Dried peptidyl resin (15 μMol) was swelled in CH_2Cl_2 (2.0 mL) and gently agitated for 30 min. The resin was isolated by filtration and a solution of $Pd(PPh_3)_4$ (3 μMol) and DABCO (75 μL) dissolved in CH_2Cl_2 was added to the dry resin and agitated for 3 h. The resin was isolated by filtration and washed repeatedly with each of CH_2Cl_2 (3 × 2.0 mL × 2 min), DMF (3 × 2.0 mL × 2 min), DIPEA/DMF (5:95, v/v)(3 × 2.0 mL × 2 min), diethyldithiocarbamic acid disodium salt/DMF (0.5:99.5, w/v) (3 × 2.0 mL × 2 min), DMF (3 × 2.0 mL × 2 min), CH_2Cl_2 (3 × 2.0 mL × 2 min), DMF (3 × 2.0 mL × 2 min). A positive Kaiser test indicated a complete reaction.

Method J: Linker 1 Attachment

2-(N-(3-acetoxybenzyl)-2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)acetamido)acetic acid (75 μMol) and HATU (67 μMol) were dissolved in DMF (1.0 mL). DIPEA (150 μMol) was added and the solution was added to the dry resin (15 μMol). The mixture was gently agitated for 1 h and the resin was isolated by by filtration and washed with DMF (5 × 2.0 mL × 2 min). A negative Kaiser test signified a complete reaction.

Method K: Linker 1 Deprotection; Phenyl Acetate Cleavage

Hybrid derivitised resin was swelled in DMF for 1 h, isolated by filtration and washed repeatedly with CH_2Cl_2 (3 × 2.0 mL × 2 min). A mixed solution of (0.5M sodium methoxide dissolved in $MeOH/CH_2Cl_2$, 1:9 v/v (1 mL) was added to the dry resin and agitated for 2 h. The resin was isolated by filtration and washed repeatedly with each of $MeOH/CH_2Cl_2$ (1:4, v/v)(3 × 2.0 mL × 2 min), MeOH (1 × 2.0 mL × 2 min), THF (3 × 2.0 mL × 2 min) and CH_2Cl_2 (3 × 2.0 mL × 2 min). LC-MS analysis indicated a complete reaction.

5.3.5 Peptide and Chimera Cleavage/Deprotection

Method L : Mild cleavage of 2-chlorotrityl resin

The dried derivitised resin (20 μMol) was treated with a solution of AcOH/TFE/CH_2Cl_2 (1:1:8, v/v/v) (2.0 mL) and agitated gently for 1 h. The resin was rinsed with TFE/CH_2Cl_2 (1:9, v/v) (2 × 2 mL). The combined filtrates were diluted with hexane (8 mL) and concentrated under reduced pressure.

Method M : Cleavage/deprotection of acid-labile resin

The dried derivitised resin (15 μMol) was treated with a cleavage cocktail of TFA/TIS/H_2O (95.0:2.5:2.5) for 1 h with gentle agitation. The resin was rinsed with TFA (2 × 2 mL) and the combined filtrates were concentrated under reduced pressure and the resulting yellow oil added dropwise to ice-cold Et_2O. The ether was decanted away from the colourless precipitate and dried under a stream of N_2 before being analysed by LC-MS and purified.

5.3.6 Chimera/Model Hybrid Synthesis

PNA to Peptide chimera CTCAG-Linker 1-OGlyNHFmoc (174)

Rink resin (15.0 μMol) was swelled as according to the general swelling protocol. PNA synthesis proceeded using the single coupling procedure **F** except after the G-PNA coupling which required a double coupling (procedure **G**). Linker 1 was coupled and deprotected as per procedures **J** and **K**. Fmoc glycine was coupled onto the unprotected hybrid using the standard peptide single coupling protocol. The resin was cleaved using the standard cleavage protocol **M**. Purification by preparative HPLC gave 15.4 mg (53%) of the title compound as a colourless powder; *m/z* 981 $[M+H]^+$:(calculated for $C_{86}H_{99}N_{33}O_{23}$ $[M+2H]^{2}$+ 981.8790, found 981.8802)

Peptide to PNA chimera Arg-Tyr-Thr-Gly-Linker 2-CTGAGNHFmoc

Wang resin (15.1 μMol) preloaded with Fmoc arginine(Pbf) was swelled as according to the general swelling protocol. The tetramer was assembled with using the single

coupling procedure **J** with the following exceptions: Tyrosine was preactivated as the pentafluorophenol ester and a coupling reagent was not required; Coupling of threonine required the double coupling procedure **K**. Linker attachment and deprotection proceeded as according to the standard procedures **H** and **I**. PNA synthesis was carried out without intervening analysis using single coupling procedure **G**. Cleavage of the resin using standard cleavage protocol **M** gave 28.0 mg of a crude colourless solid; *m/z* 981 $[M+H]^+$:(calculated for $C_{107}H_{130}N_{45}O_{24}$ $[M+H]^+$ 2430.0307, found 2430.0312)

5.3.7 Miscellaneous Chromatography methods

Reverse-Phase Chromatographic Purification of A-PNA monomer 32

Three 600mg Maxi-Clean™ SPE pre-packed large-pore C18 cartridges (Alltech) were assembled into a single contiguous column. The column was activated by flushing with MeOH (15mL, flow rate 3mL per min) and then H_2O (15mL, flow rate 3mL per min). Crude A-PNA monomer **32** (20mg) was suspended in H_2O(500μL) and DMF (600μL) was added. The suspension was agitated with gentle heating until the crude material had fully dissolved and was then carefully loaded onto the column (injection rate 1mL per min). The column was flushed with water (2.20mL, double the loading volume of the sample) to wash away the loading solvent from the crude sample prior to elution. The column was eluted with a H_2O/MeCN gradient at a flowrate of 1mL per min with an initial eluent ratio of (H_2O/MeCN, 9:1). The eluent was collected in 1mL fractions and the gradient was stepped with a 10% increase in MeCN after every 2mL of eluent collected, up to a maximum gradient of (H_2O/MeCN, 3:7); a total of 8mL of the final eluent was collected. Fractions containing pure **32** (as determined by tlc and LC-MS analysis) were combined and evaporated to give 5mg of a single compound that was confirmed by spectroscopic analysis to be **32.**

Anion-exchange chromatographic purification of Linker 2 intermediate 139

A 10.0 g prepacked SCX column was mounted on a vacuum tank and activated by elution with methanol (50 mL) at a flow rate of 5 mL per min. A solution of crude amine **139** (500 mg) dissolved in a minimum volume of methanol was loaded onto the column at a flow rate of 1 mL per min. The column was eluted with methanol (25 mL) at flow rate of 1 mL per minute and then 15 mL of ammonia saturated methanol at a

flow rate of 3 mL per min; the ammonia-containing fractions were collected as 3 x 5 mL fractions. Fractions containing pure **139** (as determined by LC-MS analysis) were combined and evaporated to give 453mg of a single compound that was confirmed by spectroscopic analysis to be **139.** The SCX column was regenerated by flushing first with a solution of saturated ammonia in methanol (until all base line material had been removed) and then with an equal volume of methanol; the column was reused to purify further batches of material in the same manner.

5.4 References

[1] K. L. Dueholm, M. Egholm, C. Behrens, L. Christensen, H. F. Hansen, T. Vulpius, K. H. Petersen, R. H. Berg, P. E. Nielsen, O. Buchardt, *J. Org. Chem.* **1994**, *59*, 5767.

[2] C. Schwergold, G. Depecker, C. Di Giorgio, N. Patino, F. Jossinet, B. Ehresmann, R. Terreux, D. Cabrol-Bass, R. Condom, *Tetrahedron* **2002**, *58*, 5675.

[3] J. M. Coull, M. Egholm, R. P. Hodge, M. Ismail, S. B. Rajur, (Perseptive Biosystems, Inc., USA). Application: WO, **1996**, p. 138 pp.

[4] S. A. Thomson, J. A. Josey, R. Cadilla, M. D. Gaul, C. F. Hassman, M. J. Luzzio, A. J. Pipe, K. L. Reed, D. J. Ricca, et al., *Tetrahedron* **1995**, *51*, 6179.

[5] B. E. Watkins, J. S. Kiely, H. Rapoport, *JACS* **1982**, *104*, 5702.

[6] G. Haaiima, H. F. Hansen, L. Christensen, O. Dahl, P. E. Nielsen, *Nucleic Acids Res.* **1997**, *25*, 4639.

[7] F. Debaene, N. Winssinger, *Org. Lett.* **2003**, *5*, 4445.

[8] F. Diness, J. Beyer, M. Meldal, *QSAR & Combinatorial Science* **2004**, *23*, 130.

[9] H. E. Gallo-Torres, E. P. Heimer, C. Witt, O. N. Miller, J. Meienhofer, C. I. Cheeseman, *Can. J. Physiol. Pharmacol.* **1984**, *62*, 319.

[10] G. D. Breipohl, E. D. Uhlmann, J. D. Knolle, (Hoechst A.-G., Germany). Application: EP, **1995**, p. 31 pp.

[11] S. Umbreen, M. Brockhaus, H. Ehrenberg, B. Schmidt, *Eur. J. Org. Chem.* **2006**, 4585.

[12] A. Eisenfuhr, P. S. Arora, G. Sengle, L. R. Takaoka, J. S. Nowick, M. Famulok, *Biorg. Med. Chem.* **2003**, *11*, 235.

[13] H. Dodziuk, O. M. Demchuk, W. Schilf, G. Dolgonos, *J. Mol. Struct.* **2004**, *693*, 145.

[14] Y. Kobayashi, A. Fukuda, T. Kimachi, M. Ju-Ichi, Y. Takemoto, *Tetrahedron* **2005**, *61*, 2607.

[15] A. Boeijen, J. van Ameijde, R. M. J. Liskamp, *J. Org. Chem.* **2001**, *66*, 8454.

[16] K.B. Jenson, T.M. Braxmeier, M. Demarcus, J.G. Frey, J.D. Kilburn, *Chem. Eur. J.* **2002**, *8*, 1300.