Synthesis and Incorporation of a Trishomocubane Amino Acid into Short Peptides

by

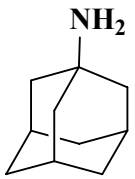
Samuel Jali
B.Sc. (Hons)

Submitted in fulfilment of the academic
requirements for the Degree of
Master of Science in the
School of Chemistry,
University of KwaZulu-Natal,
Durban

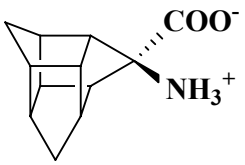
January 2006

Abstract

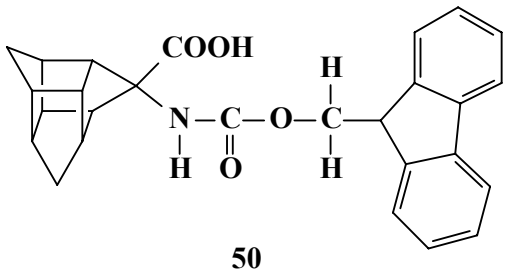
Cage compounds have attracted pharmaceutical and biological interest amongst others as anti-Parkinson agents. The serendipitous observation of the activity of 1-aminoadamantane **1** in Parkinsonian patients against selected viruses i.e. Herpes simplex Type I & II and Influenza A2-Asian viruses/Taiwan has increased the interest in cage compounds. This study involves the synthesis of the cage amino acid **14**. Due to the insolubility of pentacyclo-[6.3.0.$0^{2,6}$.$0^{3,10}$.$0^{5,9}$]-undecane (trishomocubane) amino acid **14** in both polar and nonpolar solvents, including DMSO (d_6), the synthesis of Fmoc-tris amino acid **50** was required for analysis. The Fmoc derivative of trishomocubane amino acid was also useful for controlled* coupling of the cage amino acid **14** to short peptides. The synthesis of the Fmoc-tris amino acid fluoride derivative is described as well as that of the tri-peptide (Ala-Ala-Ala). The incorporation of the Fmoc-tris amino acid fluoride in a tetra-peptide Ala-Ala-Ala-tris and in a hepta-peptide Ala-Ala-Ala-tris-Ala-Ala-Ala will also be presented.



1



14



50

A computational chemistry project was undertaken using density functional theory (B3LYP) at the 6-31+G(d) level of theory, so as to enhance the understanding of the mechanism of esterification. Methanol, acetyl chloride and acetic acid were used in the model for simplicity. Four membered ring transition states were obtained with both acetyl chloride and acetic acid. A six membered ring transition state is facilitated by the selective use of one methanol molecule from the solvent. Both a concerted and step-wise mechanism are presented.

* Addition of N-protected amino acid ensures the controlled coupling of one amino acid segment to the peptide chain.

Preface

The experimental work described in this dissertation was carried out in the School of Chemistry, University of KwaZulu-Natal, Durban, from February 2004 to December 2005, under the supervision of Dr. Hendrik G. Kruger.

These studies represent original work by the author and have not otherwise been submitted in any form for any degree or diploma to any tertiary institution. Where use has been made of the work of others it is duly acknowledged in the text.

_____

Samuel Jali

_____ day of _____ 2006

As the candidates supervisor I have approved this dissertation for submission.

Signed:_____ Name:_____ Date:_____

Table of contents

Page No.

List of Figures

List of Schemes

List of Tables

List of Abbreviations

Abu	α-aminobutyric acid
Aib	α-aminoisobutyric acid
Ala	Alanine
Asp	Aspartic acid
BBB	Blood Brain Barrier
BLYP	Becke, Lee, Yang and Parr
Boc_2O (t-Boc)	di-*tert*-butyl dicarbonate
$CDCl_3$	Deuterated trichloromethane
^{13}C NMR	Carbon-13 Nuclear Magnetic Resonance Spectroscopy
CPU	Central processing units
CNS	Central nervous system
D	Dextrotatory
Dbg	Dibutylgylcine
DCC	N,N-Dicyclohexylcarbodiimide
DFT	Density functional theory
DIPCDI	N,N-diisopropylcarbodiimide
DMAP	4-dimethyl-aminopyridine
DMF	N,N-Dimethylformamide
$DMSO(d_6)$	Deuterated dimethyl sulfoxide
Dpg	Dipropylglycine
E	Energy
EI	Electron impact
Eqn	Equation
ESR	Electron Spin Resonance
Et_3N	Triethylamine
eV	Electron volts
FAB	Fast atom bombardment
Fmoc(-Cl)	9-fluorenylmethyl (chloro)formate
FmocAM	p-[(R,S)-α–[1–(9H–fluoren–9-yl)methoxy-formamido]-2,4-dimethyl-benzyl]-phenoxyacetic acid
g	Grams
Glu	Glutamic acid
Gly	Glycine
Gly-Gly	Glycine-glycine
G03	Gaussian 03
GTO	Gaussian type orbitals
ΔG	Change in Gibbs free energy
H	Hamiltonian operator
^1H NMR	Proton Nuclear Magnetic Resonance Spectroscopy
HATU	*O*-(7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium-hexafluorophosphate
HBTU	2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium-hexafluorophosphate

H_f^o	Heat of formation
HF	Hartree-Fock
HOBt	1-Hydroxybenzotriazole
HOMO	Highest occupied molecular orbital
HPLC	High performance liquid chromatography
ΔH	Change in enthalpy
hν	Electromagnetic radiation
IR	Infrared spectroscopy
IRC	Intrinsic reaction co-ordinates
L	Levorotatory
LCAO	Linear combination of atomic orbitals
LUMO	Lowest unoccupied molecular orbital
MBHA	*p*-methylbenzhydrylamine
MEP	Minimum energy path
m/z	mass per electron
MP2	Second order Møller-Plesset
Mp	Melting point
NCSA	National Computational Science Alliance
Nle	Norleucine
NRF	National Research Foundations
Nva	Norvaline
p	Para
p-TS	*p*-Toluenesulphonic acid
PAL	5-(4-(9-fluorenylmethyloxycarbonyl) aminomethyl-3,5-dimethoxyphenoxy)-valeric acid (Peptide amide linker)
PCU	Pentacyclo-[5.4.0.$0^{2,6}$.$0^{3,10}$.$0^{5,9}$]-undecane
PCM	Polarisable continuum model
PES	Potential energy surface
R	*Rectus*
RHF	Restricted Hartree-Fock
S	*Sinister*
SPPS	Solid Phase Peptide Synthesis
S_N2	Nucleophilic substitution
STO	Slater type orbitals
SCRF	Onsager self-consistent reaction field
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
TLC	Thin layer chromatography
TNBS	2,4,6-trinitrobenzenesulfonic acid
Tris	Trishomocubane
TS	Transition state
TS-4-X	Four membered ring transition state
TS-6-X	Six membered ring transition state

UHF	Unrestricted Hartree-Fock
UKZN	University of KwaZulu-Natal
USA	United States of America
US	United States
UV	Ultra-violet
$\mathbf{a}_o$	Radius of the cavity
Ψ	Wave function
Å	Ångstrom
α	Alpha
ϕ	Atomic orbitals
β	Beta
ρ	Density
δ	Delta
^oC	Degrees Celsius
ν_{max}	Frequency of maximum absorption
γ	Gamma
ω	Omega
%	Percent
ψ	Psi
π	Pi
ϕ	Phi
σ	Sigma
ζ	Zeta

Acknowledgements

I would like to acknowledge those who have guided my path through the field, especially Dr H.G. Kruger for allowing me to join his research group and giving me freedom to explore throughout this study, I have learnt a lot. Dr T. Govender and Mr S.P. Mdluli are also acknowledged for their mentorship and guidance throughout this study. I also thank Prof. G.A. Kolawole from the University of Zululand for his support and motivation to continue with my honours studies. I am particularly grateful to those who have dedicated their time to read and comment upon draft copies of various chapters including Mr J. Paraskevopoulos and Mr W. Waudo. Many others too numerous to specify, who have assisted me to make this work possible are also acknowledged. My discussions with Mr J. Paraskevopoulos and Dr G.E.M. Maguire have been particularly fruitful, not only to this study but in extending my chemistry knowledge. The following sponsors are well acknowledged for their financial support and making it possible for this study to be undertaken: the National Research Foundations (NRF), the University of KwaZulu-Natal, the National Computational Science Alliance (USA) under grant number CHE010004 (NCSA IBM P690) and Sasol for upgrading our computational facilities at the University of KwaZulu-Natal on which some work was undertaken.

This work is dedicated to my family and friends for supporting and encouraging me to continue with my studies.

Chapter 1

Introduction

In recent years, saturated polycyclic "cage" molecules have attracted special attention from organic chemists.[1,2,3] Most saturated cage molecules contain considerable strain energy due to unusually long carbon-carbon σ-bonds† in the cage framework and unusual C-C-C bond angles that deviate significantly from the normal ~109.5°. The adamantane "cage" is an exception with a carbon skeleton close to the perfect diamond structure.[4,5] The strain energy in other cage moieties expresses itself through unusual patterns of chemical reactions.

Polycyclic "cage" hydrocarbons are of interest in three areas of applied research:

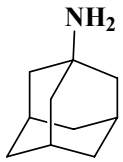
- As a potential new class of high energy or high-density fuels due to their highly strained structures.[1,6,7] Cage compounds such as 6,6,10,10-tetranitropentacyclo-[5.3.0.$0^{2,6}$.$0^{3,10}$.$0^{5,9}$]-decanes are of interest as a potential new class of explosives and propellants.[1,8,9]
- Pharmaceutical interest.[4,10,11,12]
- A relatively new field involving applications of chiral cage compounds as catalysts in asymmetric synthesis.[13,14,15,16]

Pharmacological interest in cage compounds abounded after it was serendipitously observed that amantadine (1-aminoadamantane) **1** was active as an anti-Parkinson agent when used to treat patients against influenza A_2 Asian viruses. Other amino functionalised cage compounds like **2** and **3** have shown activity against selected viruses i.e. Herpes simplex Type I & II and Influenza A_2. Two amino derivatives of (D_3)-trishomocubane **3a** and **3b** were evaluated for their *in vivo* and *in vitro* activities and compared favourably with amantadine **1**. One of the derivatives **3a** exhibited a great improvement on amantadine when used to protect mice against influenza A_2.[17]

Parkinson's disease was also known as (*paralysis agitans*) before it was described in 1817 by the London physician James Parkinson.[18] It is a slowly progressive malady of the sixth and later decades of life, and James Parkinson's description of this disease was that it is "involuntary tremulous motion, with lessened muscular power, in parts not in action and even when supported; with propensity to bend the trunk forwards, and to pass from a walking to a running pace, the senses and intellects being uninjured."

---

† A σ-bond is formed by head-on overlap of two p-orbitals, the shared electrons are centered about the axis between the two nuclei.



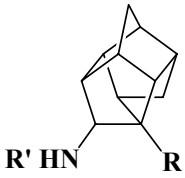
1

Figure 1: Amantadine (1-Aminoadamantane)

The interest of the organic group at UKZN[19] in cage molecules was stimulated by the biological and pharmaceutical properties of amino functionalised cage compounds such as **2** and **3**:

- as potential antiviral agents.
- as anti-Parkinson agents mimicking the activity of amantadine **1**.[4,20]

2



3a R' = R = CH_2CH_3
3b R' = H, R = CH_2CH_3

Figure 2: PCU and Trishomocubane amines

Cage compounds have therefore found wide applications in drugs, since it was discovered that the incorporation of the cage skeleton enhances the transport of the drug across cell membranes, the blood brain barrier (BBB) and into the central nervous system (CNS).[4,21] Drugs modified to include cage moieties experience an increased affinity for lipophilic regions in receptor molecules due to the hydrophobic/lipophilic nature of the cage skeleton. It has been suggested that the increase in activity observed for cage-modified drugs is a direct result of an increase in drug delivery.[4,21] A further advantage, conferred by the steric bulk of the cage skeleton, is the resistance of cage-modified drugs to metabolic degradation. This increases the *in vivo* life time of the drug and reduces the frequency of drug administration to the patient.[21,22] The research group at UKZN is currently using three different cage structures, namely pentacyclo-[5.4.0.$0^{2,6}$.$0^{3,10}$.$0^{5,9}$]-undecane (PCU, **4**), a more stable isomer[23] with the trivial name of trishomocubane **5**[24] and the norbornane **6**‡ /camphor **7** skeleton.

The trishomocubane skeleton is intrinsically chiral, while the PCU skeleton is a (*meso*) compound that can be made chiral through asymmetric synthesis. The camphor skeleton **7** is also a chiral compound while norbornane **6** is achiral (*meso*).

‡ Norbornane and camphor are strictly speaking not classical cage compounds, but are structurally related to the PCU skeleton.

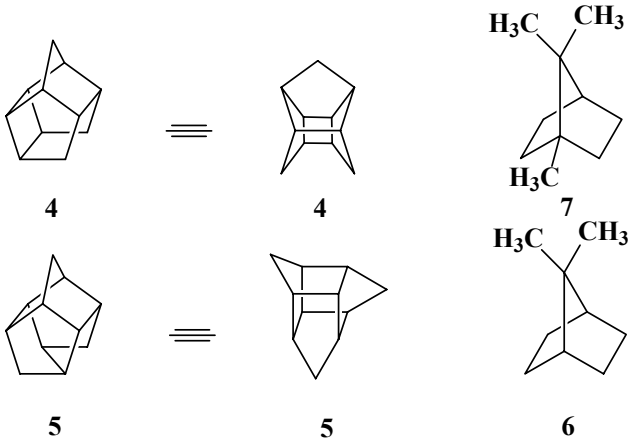
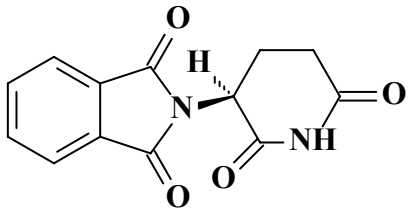


Figure 3: Cage skeletons

Chirality is an important phenomenon in biochemistry and in the pharmaceutical industry.[25,26] The effect of chirality is observed in drugs, where a biologically active chiral compound (drug) interacts with a receptor site which is chiral. Enantiomers of the drug interact differently and produce different effects. Thalidomide **8** (a drug prescribed for morning sickness) is a good example, where both its enantiomers have a sedative effect but the (-)-enantiomer **8** causes foetal deformity.

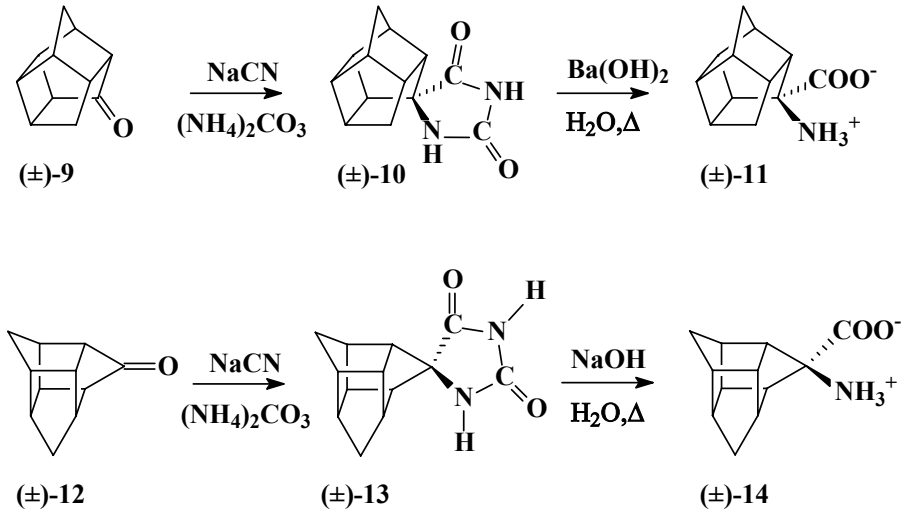


8

Figure 4: (-)-Thalidomide

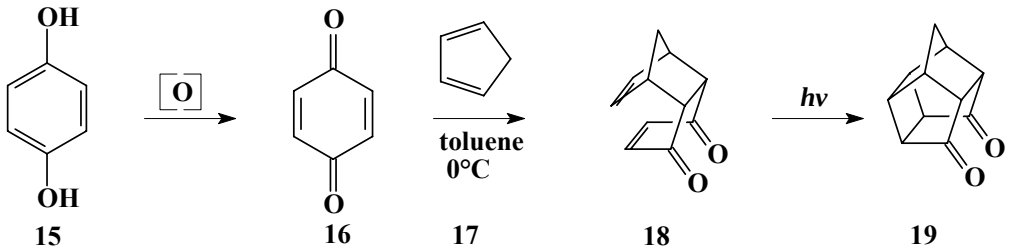
Sessler *et al.*[27] have demonstrated that (*S*)-amino acids are transported with greater affinity than the (*R*)-enantiomers in a lasalocid conjugate carrier. Fujii *et al.*[28] have shown that the opioid peptide, dermorphin, [Try-(*R*)-Ala-Phe-Gly-Try-Pro-Ser-NH_2] loses its activity when (*R*)-Ala is substituted by the (*S*)-Ala enantiomer.

One of the interests of the organic research group at UKZN is to incorporate unnatural cage α-amino acids into peptides. The pentacyclo-[5.4.0.$0^{2,6}$.$0^{3,10}$.$0^{5,9}$]-undecane amino acid **11**[29] and the pentacyclo-[6.3.0.$0^{2,6}$.$0^{3,10}$.$0^{5,9}$]-undecane amino acid (trishomocubane amino acid, **14**)[30,31] were synthesized from the corresponding monoketones **9** and **12** via hydantoin formation (**Scheme 1**).[29,30] Note that the detailed syntheses of **9**, **12** and **14** will be presented in **Chapter 2**.



Scheme 1: Synthesis of the cage amino acids

The starting material for both monoketones **9** and **12** is readily obtained from the Diels-Alder reaction between *p*-benzoquinone **16** and cyclopentadiene **17**. The intramolecular [2 + 2] photocyclization reaction[32] of 5,8-methano-4a,5,8,8a-tetrahydro-1,4-naphthoquinone (adduct, **18**) yields Cookson's diketone **19**.[33,34] Cookson's diketone **19** has formed the foundation for many studies about the chemistry of pentacycloundecyl compounds.



Scheme 2: Synthesis of the adduct (18) and the Cookson's diketone (19)

The incorporation of the cage residue into non-natural peptides was investigated using molecular mechanics and (*ab initio*) computational methods. It has been demonstrated[35,36,37,38,39] that amino acids such as **11** and **14** induce an α-helix and strong β-turn effects in the peptide chain. β-turns[35,36,37,40] play an important role in globular proteins from both structural and functional points of views. Amongst the forces that control protein folding, β-turns comprise about 25% of all residues in proteins and are the most important motif for protein folding to produce their compact structure.[41] A polypeptide chain cannot fold into a compact structure without the component of turns. The turns (β-turns) usually occur on the exposed surface of proteins and hence probably represent antigenic sites or involve molecular recognition.[42,43] β-turns provide useful information to define template structures for the design of new molecules such as drugs and pesticides. Therefore the prediction of β-turns in the secondary structure of proteins is an important element of computational analysis.[36,43]

Kruger's research group has started a computational chemistry project[44] to develop a computational model for esterification reactions. Singh started an ambitious project to model the esterification reaction of the PCU amino acid **11**, but it turned out that the development of a simpler model is required before the project can be concluded. The trishomocubane amino acid **14** has previously been esterified,[45] but the PCU amino acid **11** seems to be more resistant towards esterification.[45,46]

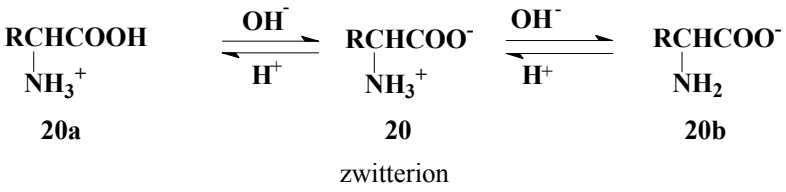
This investigation had the following objectives:

- To synthesize the trishomocubane amino acid **14** in good yield.
- To collaborate in a project to establish a procedure for the incorporation of the trishomocubane amino acid into a short peptide.
- To develop a computational model that describes the mechanism of esterification of acids and acid chlorides with methanol. This is required to explain (in future) the difference in reactivity of the two cage amino acids (**11** and **14**) towards esterification, as mentioned above.

Chapter 2

Synthesis of (D_3)-trishomocubane amino acid and its derivatives

Amino acids are organic compounds which contain both carboxylic acid groups (-COOH) and amino groups (-NH_2). Alpha (α) amino acids are those with both the amino and carboxylic acid groups linked to the same α-carbon atom. Amino acids are amphoteric compounds (**Scheme 3**) due to the acidic nature of the carboxylic acid group and the basic nature of the amino group.[47,48] The R-groups play an important role by contributing to the amphipathic character of the amino acids. The R-groups make the amino acid more hydrophobic/lipophilic and the charged groups make it hydrophilic. The interaction between hydrophobic residues in an aqueous environment enables the stability and molecular self-assembly of amphipathic peptides.



Scheme 3: The structure of amino acids is pH dependent.[49]

At physiological pH most natural amino acids are in the zwitterion form, making the amino acids more soluble in blood. The natural occurring amino acids are made of C, H, O and N except cysteine and methionine which contain sulphur.[50] Amino acids have attained a wide variety of commercial applications as ingredients in food and animal feed and as chiral building blocks for pharmaceuticals. There is an increasing interest in optically pure (*R*)- and (*S*)-α-amino acids as precursors for semisynthetic antibiotics, new herbicides, insecticides, and physiologically active peptides.[51] Amino acids are chiral[§] molecules which, when in solution, induce (±)-rotation of plane-polarised light. In nature amino acids appear predominantly in the (*S*)-form except for cysteine, the achiral glycine (Gly) and various non-common amino acids found in sea organisms and plants. These (*R*)-amino acids are in many cases used in various types of chemical defense mechanisms.[52] Asymmetric synthesis of these amino acids produces racemates. The (-)-rotation which is counter-clockwise by definition is called "levorotatory" rotation whereas the (+)-rotation is clockwise and is called "dextrotatory" rotation.[53]

§ The word *chiral* comes from the Greek word meaning *handedness*. A *chiral* compound has a stereogenic centre (carbon with four different groups according to a simplified organic definition).

Chirality is demonstrated by the two alanine **21** enantiomers below. Alanine can be obtained in two forms (*S*)-Ala-**21** and (*R*)-Ala-**21** which are mirror images. The two mirror images have similar physical and chemical properties in the absence of an external chiral influence. This implies that (*S*)-Ala-**21** and (*R*)-Ala-**21** have the same melting point, solubility, chromatographic retention time, infrared (IR) and nuclear magnetic resonance (NMR) spectra. They differ physically only in that they rotate plane-polarized light to an equal degree but in opposite directions.[25,53]

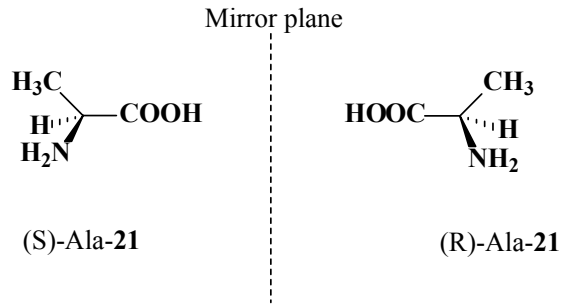
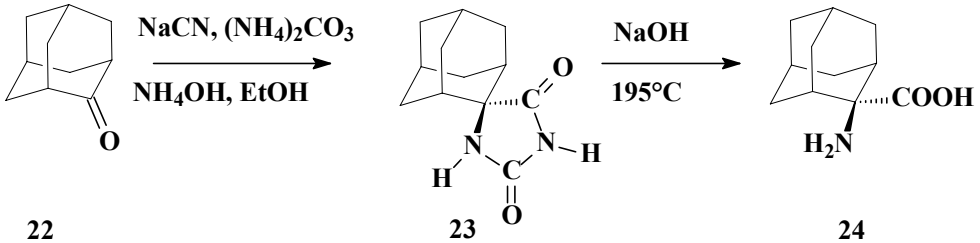


Figure 5: (*R*) and (*S*)-alanine

In recent years there has been an increased interest in the preparation of novel unnatural amino acid derivatives to serve as building blocks in drugs.[29,54] The work reported in this dissertation is based on artificial (non-proteinogenic) amino acids that are of great importance in pharmaceutical chemistry. Based on the enhanced pharmaceutical properties of drugs attached to cage moieties,[4,21] cage amino acids have potential as a new class of bioactive compounds. It was predicted[35,36,38] that cage amino acids could be used in the field of peptide synthesis, where the hydrophobic side chains could potentially assist the peptide to cross the blood brain barrier.[20,55] Synthesis of unnatural amino acids is important not only for the production of potentially useful compounds but also for the preparation of structural analogues, since in many cases only small quantities of active species are isolated from natural sources.[54,56]

The strategy for the synthesis of the two amino acids **11** and **14** was adopted from the method used to obtain 2-aminoadamantane-2-carboxylic acid **24**, (**Scheme 4**).[57]



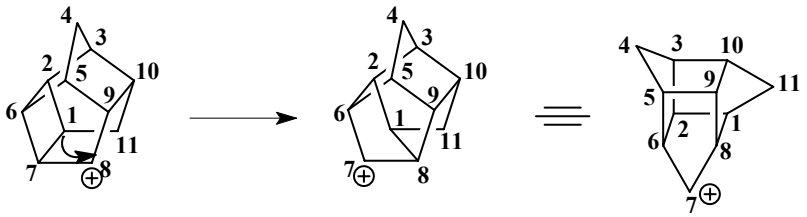
Scheme 4: Synthesis of (2-aminoadamantane-2-carboxylic acid) (24)

Trishomocubane amino acid (tris-amino acid, **14**) was first synthesized by Govender and Raasch *et al*. It was first characterized as the Fmoc derivative due to its insolubility in organic solvents.

As indicated in **Chapter 1**, the starting material for the tris-amino acid **14** is Cookson's diketone **19**. The synthesis of **19** *via* a Diels-Alder reaction is presented in **Scheme 2**. *p*-Benzoquinone **16** is recrystallized from dichloromethane/diethyl ether and dissolved in toluene/benzene. Freshly cracked cyclopentadiene **17** is then added to yield 5,8-methano-4a,5,8,8a-tetrahydro-1,4-naphthoquinone (adduct, **18**).

It was believed that the adduct **18** underwent photocyclisation *via* a $[\pi 2_s+\pi 2_s]$ cyclisation "suprafacially" to yield the diketone **19**.[33,58] Recently Marchand *et al.* used (*ab initio*) computational studies to demonstrate that the cyclisation of the adduct probably occurs via a stepwise diradical mechanism. The mechanism proceeds through a triplet excited state rather than a suprafacial $[\pi 2_s+\pi 2_s]$ cyclisation. Kent *et al.* also used computational modelling (force field calculations) to demonstrate that the PCU skeleton **25** is a kinetic product which can rearrange to the thermodynamically more stable trishomocubane **26**.

Force field calculations were used to demonstrate that the trishomocubane skeleton is ~10 kcal.mol^{-1} more stable than the PCU moeiety.[22,23,59] Kent *et al.* used the reported values by Schleyer *et al.* and Allinger *et al.* to argue that the PCU skeleton can be used as an excellent starting material for the synthesis of trishomocubane related compounds, following a 1,2 alkyl shift of the intermediate secondary cation (**Scheme 5**). They also calculated the stabilities of various carbocations and showed that a carbocation at C8 or C11 is more stable than any other possibility. This is also the only position of the carbocation on cage **25** which does not give rise to a structure with a three or four membered ring. A cation at C9 may not be prohibitively strained but all possible 1,2 alkyl shifts lead to a four membered ring product. A discussion of the computational studies executed by Raasch at UKZN on the rearrangement of the PCU alcohol **31** to trishomocuban-4-ol **27** will be presented towards the end of this chapter.



25 **26** **26**

Scheme 5: Rearrangement of PCU skeleton to trishomocubane

The intrinsically chiral trishomocubanone **12** is a key member and a useful intermediate for the synthesis of amino related compounds, as the carbonyl group provides a convenient "handle" for optical resolution and for the introduction of other functional groups.[60]

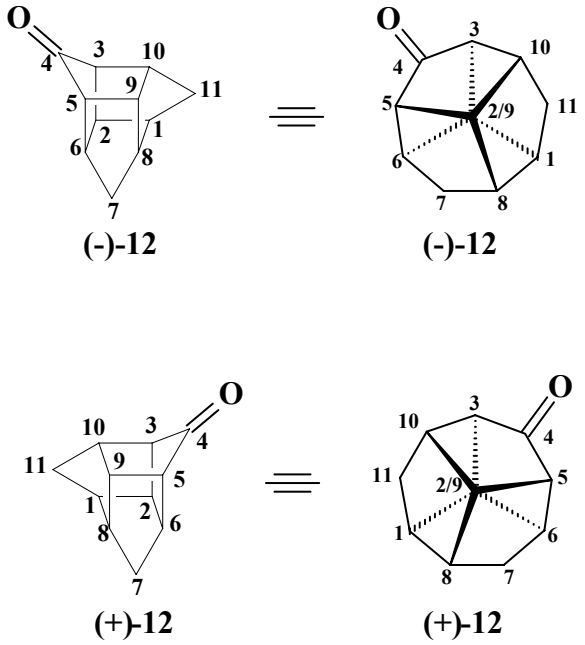


Figure 6: (-) and (+)-Trishomocubanone

Breakage of a single bond within a molecule is required to convert one stereogenic center into another.[26,61] This principle is not applicable to the intrinsically chiral trishomocubane skeleton. The chirality of a molecule is sometimes ascribed to the presence of a chiral axis. The same principle is applied to the (D_3)-trishomocubane skeleton, where the chiral axis is a C_3 axis through C2-C9. The result is that the trishomocubane skeleton resembles either a left handed or right handed propeller (Figure 6).

Nakazaki *et al.*[62] have managed to isolate the (+)-ketone **12** enantiomer. The (-)-**12** enantiomer was also successfully isolated by Mueller *et al.*[63] For normal molecules without D_3 symmetry one would expect that addition of a nucleophile to the carbonyl carbon of the trishomocubane monoketone (±)-**12** would lead to two pairs of diastereomers (Figure 7).

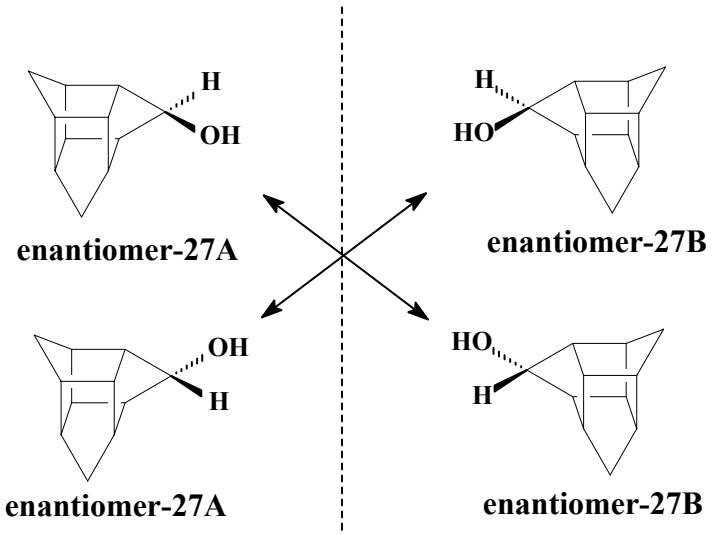
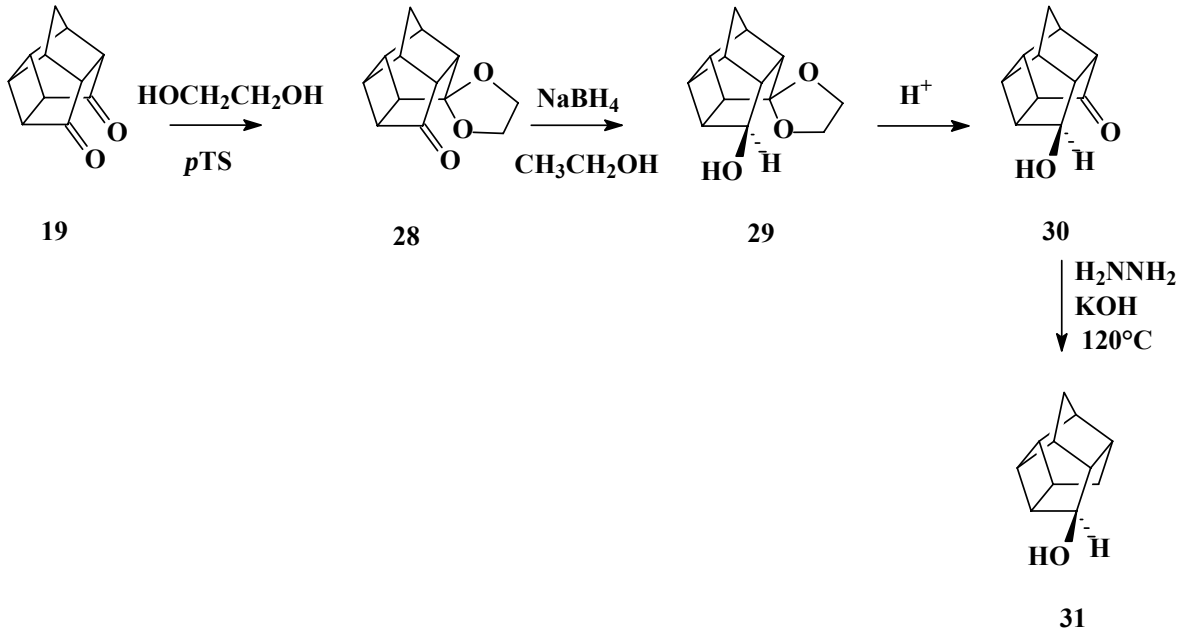


Figure 7: Alcohol "diastereomers" of the quaternary substituted carbon of trishomocubane

There are essentially two experimental approaches to the rearrangement of the PCU skeleton to trishomocubanone **12**.

The first route (**Scheme 6** and **Scheme 7**) was reported by Dekker *et al.*[64] and the second method (**Scheme 8**) was reported by Kent *et al*. The first method starts with the hydroxyketone **30** which is converted to an alcohol **31**. The alcohol can then be rearranged as will be demonstrated later.

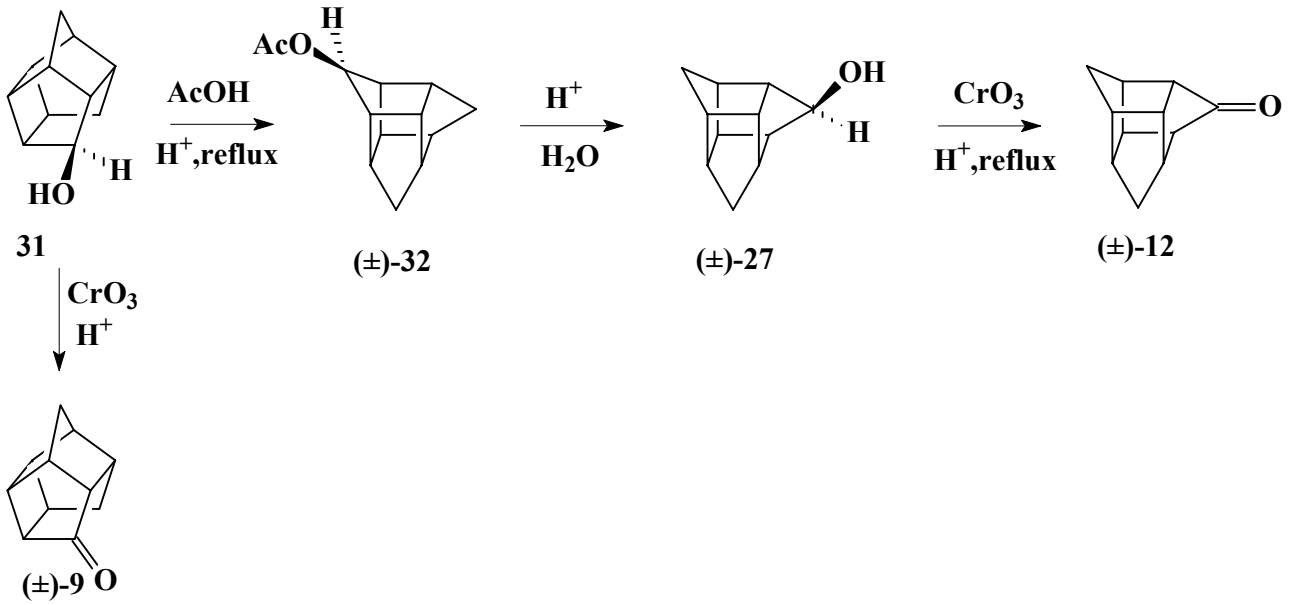
Synthesis of the alcohol **31** from the diketone **19** involves protection of the free ketone via ketal formation for the purpose of mono-ketone reduction. Mono protection is achieved by refluxing the dione **19** in toluene/benzene in the presence of *p*-toluenesulfonic acid and 1,2-ethanediol in a Dean and Stark apparatus to yield the mono-ketal (pentacyclo-[5.4.0.0^{2,6}.0^{3,10}.0^{5,9}]-undecane-8,11-dione-mono-ethylene ketal **28**).[65] Even with excess 1,2-ethanediol only one side of the diketone can be protected due to the steric hindrance imposed by the cage moiety.[64,65]



Scheme 6: Synthesis of pentacyclo-[5.4.0.0^{2,6}.0^{3,10}.0^{5,9}]-undecane-8-ol (31)

The mono ketal **28** is then reduced using sodium borohydride (or lithium aluminium hydride) in ethanol to yield the hydroxy-ketal **29** (11-hydroxypentacyclo-[5.4.0.0^{2,6}.0^{3,10}.0^{5,9}]-undecane-8-one-ethylene ketal).[29,65] Deprotection of the hydroxy-ketal **29** is achieved by hydrolysis of the ketal in 10% (v/v) hydrochloric acid to yield the hydroxy-ketone **30**. Reduction of the alcohol using the Huang-Minlon method gives the PCU alcohol **31** which has a distinct camphor-like smell.[30,64,69]

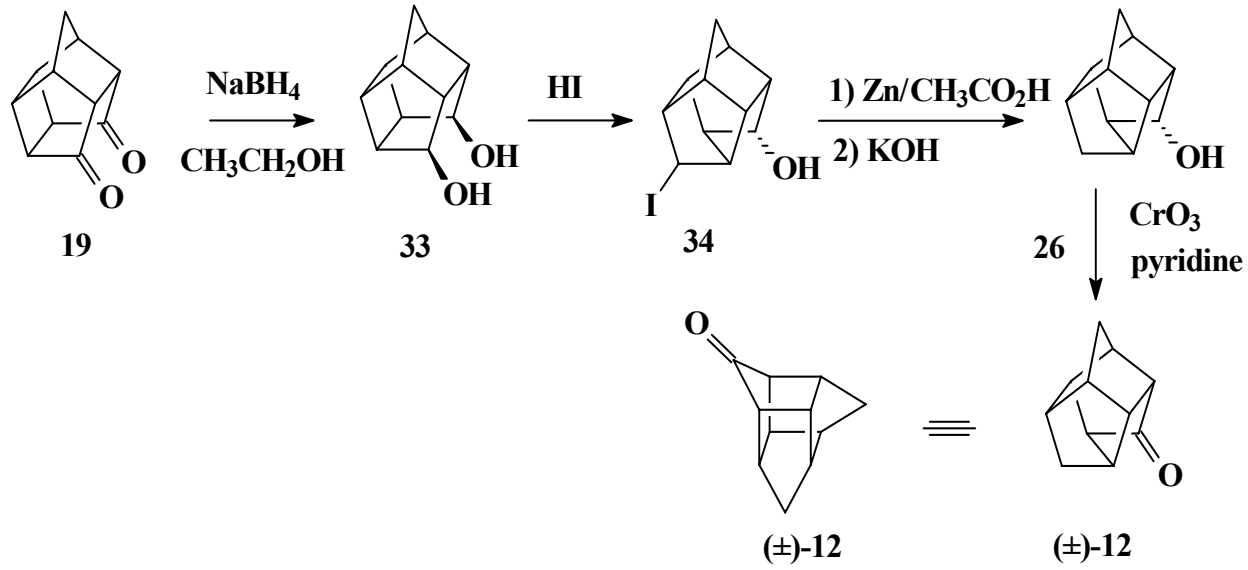
Rearrangement of pentacyclo-[5.4.0. 0^{2,6}.0^{3,10}.0^{5,9}]-undecane-8-ol **31** to pentacyclo-[6.3.0.0^{2,6}.0^{3,10}.0^{5,9}]-undecane-4-ol **27** is achieved by refluxing in glacial acetic acid and concentrated sulphuric acid to yield (D_3)-trishomocuban-4-acetate **32** as an orange oil **(Scheme 7)**.



Scheme 7: Synthesis of pentacyclo-[5.4.0.^{2,6}.0^{3,10}.0^{5,9}]-undecane-8-one (9) and (D_3)-trishomocuban-4-one (12)

The trishomocuban-4-acetate (orange oil, **32**) is then hydrolysed by stirring in methanol and potassium carbonate to yield the white solid trishomocuban-4-ol **27** in good yield (~72%). An overall yield of (~ 45%) trishomocuban-4-ol **27** was obtained relative to **19**. Trishomocubanone **12**[23,60] was obtained by refluxing the trishomocuban-4-ol **27** overnight using a Jones oxidation. An overall yield of ~37% trishomocubanone **12**, relative to the dione **19** was obtained.

The second synthetic route (**Scheme 8**) reported by Kent *et al.* for the synthesis of (D_3)-trishomocubanone **12** also starts from the pentacyclic diketone **19**.



Scheme 8: Kent's method for the synthesis of trishomocubanone

The reduction of **19** using sodium borohydride in ethanol readily produced the *endo*-*endo*-diol **33** in good yield (~75%). There are various reducing agents that can be used to obtain **33** in good yield

e.g. lithium aluminum hydride in THF, 2-propoxide in 2-propanol and $NaAl(C_2H_5)_2H_2$. Kent *et al.* wrongly reported that all reducing agents favour the *endo-endo* isomer over the *exo-exo* isomer. $LiAlH_4$ reduction of the dione yields the *exo-endo*-alcohol as the major product.[66,67]

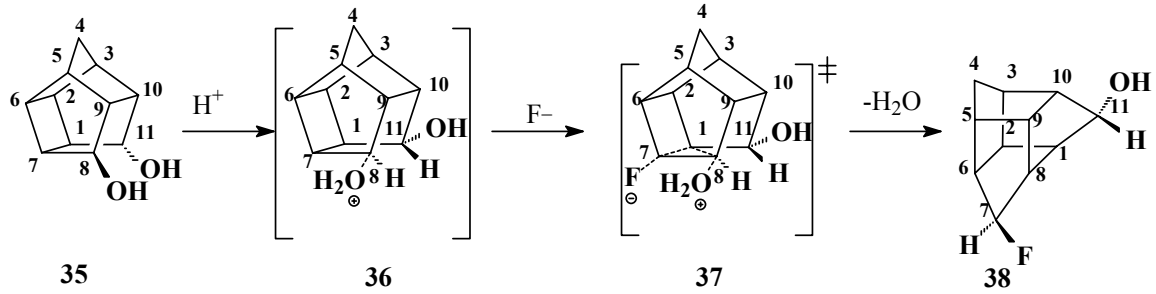
The rearrangement of the diol **33** into the trishomocubyl system can be effected with acid under milder conditions. The diol **33** can be treated with aqueous HI for 2h at 100°C to give the iodotrishomocubyl alcohol **34**.[23,64,68] Deiodination of **34** is achieved with zinc in acetic acid and potassium hydroxide to yield trishomocuban-4-ol **27**, which can then be oxidized with CrO_3 and pyridine to yield the monoketone **12**.

The method by Kent *et al.* has three advantages over the method reported by Dekker *et al.* It is shorter, all reagents required are readily available and preparation appears to be simple. Although the first method reported by Dekker *et al.* is longer, it makes use of simple chemistry and produces reasonable yields [~43% overall yield for trishomocuban-4-ol **27** and ~37% for trishomocubanone **12**]. The method by Kent gives an overall yield of ~44% trishomocubanone **12** mainly due to the poor yield of **34**. Dekker's method is preferred in our laboratory, mainly because it is well established and the shorter method appears to be less consistent in our hands.

Raasch[69] has used computational modelling at the RHF/3-21+G level of theory to enhance the understanding of the rearrangement of the PCU diol to trishomocubane. He showed that the intramolecular S_N2 rearrangement reaction to fluoro-trishomocubanol **38** can only proceed from either the *exo-endo* PCU diol **35** or the *exo-exo* PCU diol. The *endo-endo* diol **33** can only proceed via a S_N1 mechanism which is energetically much less preferred.

Rearrangement of the diol[23,64,68] requires refluxing in HI at 100°C for 2h to yield **33** (**Scheme 8**). Hydrofluoric acid was used in the computational investigation by Raasch so as to reduce the size of the calculations. It is apparent that good basis sets for iodine are problematic.[70] The computational model involves the rearrangement of **35** *via* the protonation of the 8-*exo*-hydroxy-11-*endo*-hydroxy PCU **35** with HF to yield 7-*exo*-fluoro-11-*endo*-hydroxy-trishomocubane **38** (**Scheme 9**). A scan calculation was performed to pull the fluorine ion closer to C7 until a maximum was reached. The transition state **37** involves the breaking of the C1-C7 bond of the PCU skeleton and the formation of a C1-C8 bond for the cage rearrangement to trishomocubane. The fluorine ion is facilitating the rearrangement where C1 eventually performs an intramolecular S_N2 attack on C8, so as to cleave

off water. Only one enantiomer** was used although the synthesis produces a racemate (**Scheme 9**).



35 **36** **37** **38**

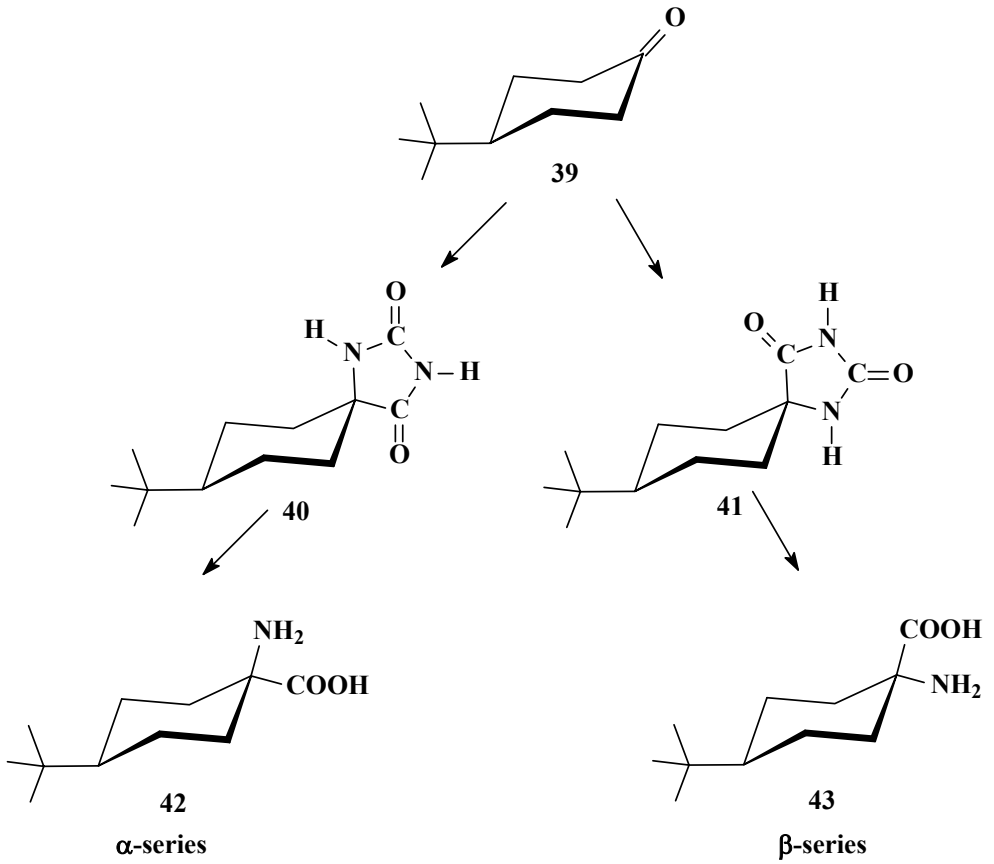
Scheme 9: Proposed mechanism for the *exo-endo*-PCU diol rearrangement to 7-*exo*-fluoro-11-*endo*-hydroxy-trishomocubane

The transition state with the *exo-endo* PCU diol was proven to be ~2.6 kcal/mol lower than the corresponding *exo-exo* PCU diol transition state. The fact that the *exo-endo* diol is kinetically preferred to the *exo-exo* diol implies that the *exo-endo* diol **35** can be used as the preferred starting material. It also implies that $LiAlH_4$ is perhaps a better choice for the reduction of the dione as *exo-endo*-alcohol is the major product,[66,67] while reduction with $NaBH_4$ produces the undesired *endo-endo*-alcohol as the major product.

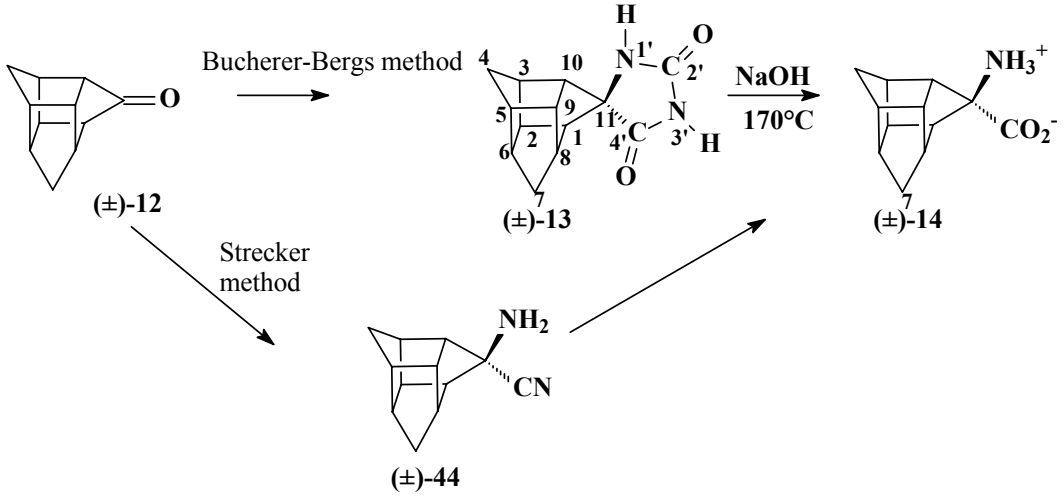
Trishomocubanone **12**[23,60] was converted to the trishomocubane hydantoin **13** by treatment with NaCN, $(NH_4)_2CO_3$, NH_4OH and CH_3CH_2OH at 100°C (Bucherer-Bergs[71] method - **Scheme 11**). Trishomocubanone **12** has D_3 symmetry and has a theoretical advantage over the PCU monoketone since only two hydantoin isomers can potentially be formed, compared with four possible hydantoin isomers in the case of PCU monoketone **9**. The Strecker or Bucherer-Bergs method can be used to obtain the same amino acid (**Scheme 11**). However, Munday[72] (**Scheme 10**) found that the Bucherer-Bergs reaction of 4-*tert*-butylcyclohexanone **39** gave predominantly one isomeric hydantoin (**40**, α-isomer), accompanied by only a trace amount of the second isomer (**41**, β-isomer).†† Hydrolysis of these hydantoins **40** and **41** furnished the two isomeric amino acids **42** and **43** (**Scheme 10**). While the amino acid of the α-series is the major product of the Bucherer-Bergs route, only the amino acid of the β-series can be obtained from the amino nitrile furnished by the Strecker route.

** "Note that enantiomers have similar properties, as well as identical energies"

†† Non symmetric cyclohexanones form hydantoin diastereomers with the Strecker synthesis and the Bucherer-Bergs method.



Scheme 10: Munday's comparison of the Strecker and Bucherer-Bergs reactions



Scheme 11: Bucherer-Bergs method for synthesis of amino acid

The same argument is expected in both trishomocubanone **12** and the PCU monoketone **9**. Evidence of the successful synthesis of the hydantoin **13** was attributed to the presence of two carbonyl absorption peaks at 1772 and 1726 cm^{-1} and an N-H absorption peak at 3212 cm^{-1} in the IR spectrum. The ^{1}H NMR spectrum shows two D_2O exchangeable peaks. One at 10.54 ppm (imide N3'-H) which was shifted downfield since it is positioned between two electron withdrawing carbonyl groups.[29,30] A second peak was observed at 7.88 ppm and is due to the amide N1'-H.

The ^{13}C NMR spectrum exhibited two carbonyl carbon peaks, at 178 and 152 ppm. The peak at 152 ppm was due to the urea carbonyl group (C2'). The eight methine carbons of the cage are found between 39.57 and 55.67 ppm. The presence of the imide (N3'-H) and the amide (N1'-H) peaks at 10.54 and 7.88 ppm in ^{1}H NMR proves the successful synthesis of trishomocubane hydantoin. The appreance of two carbonylic peaks at 178 (C4') and 152 ppm also proved the conversion of a ketone carbonyl to a hydantoin ring.

As a result of the unique D_3 symmetry of the trishomocubane cage skeleton, only two enantiomers [(+)-**13** and (-)-**13**] (see **Figure 8**) of the hydantoin **13** are obtained as a racemate.

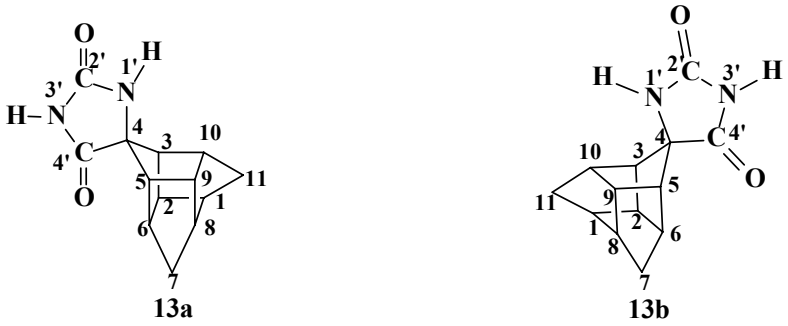
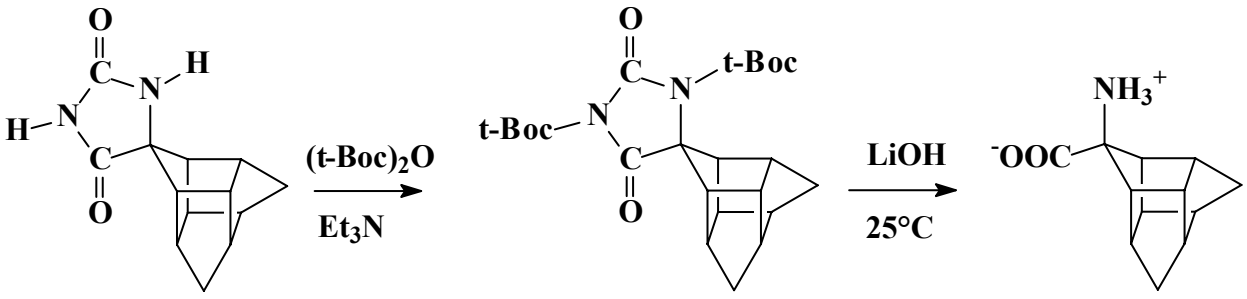


Figure 8: Trishomocubane hydantoin (13)[62,63]

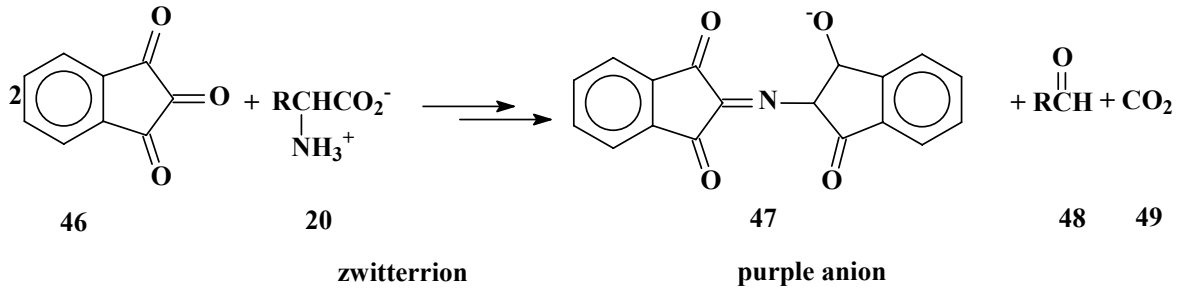
Due to the poor yield of amino acid **14** (~61%) obtained by the sodium hydroxide hydrolysis of the hydantoin at 170°C, a further investigation by Raasch *et al.* led to the development of a new method for hydrolysis, using bis-t-Boc protection of the hydantoin to yield **45** (**Scheme 12**).



Scheme 12: Bis-t-Boc method for the synthesis of amino acid

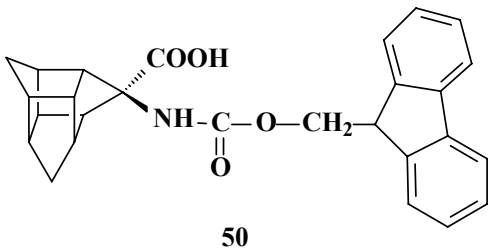
The two N-H groups of the hydantoin **13** were protected by di-*tert*-butyl dicarbonate (Boc_2O) in the presence of a base, triethylamine (Et_3N), following Rebek's method[73] for the hydrolysis of lactams and secondary amides.[74] Since the carbonyl groups of the t-Boc protected amides and imides are more susceptible to nucleophilic attack during hydrolysis, the amino acid **14** is obtained in quantitative yield. The reason for this more facile hydrolysis may be due to release of steric strain and the fact that nitrogen is converted into a better leaving group.[30,69,73,74]

A successful synthesis of the bis-t-Boc tris-hydantoin **45** was confirmed by the disappearance of the two N-H peaks at 7.88 and 10.56 ppm in the proton NMR spectrum and the existence of four carbonyl peaks at 1724, 1751, 1767 and 1801 cm^{-1} in the infrared spectrum. The ^{13}C NMR spectrum displayed four carbonyl carbon peaks, at 145.79, 149.02, 149.67and 169.58 ppm and were identical to the results previously reported by Raasch *et al*. The bis-t-Boc hydantoin hydrolysis produced > 95% yield of tris-amino acid **14**.[45,73] The trishomocubane amino acid **14** is isolated at pH 6.5 as a zwitterion.[29,30,73,74] Due to the poor solubility of the cage amino acid in organic solvents [even in DMSO(d_6)], only two analytical tools were used to characterize the underivatised amino acid **14**. The first test (tool) is the Ninhydrin test (Kaiser test)[75] which gives a violet-blue colour due to the presence of the free $–NH_2$ group (**Scheme 13**).



Scheme 13: Reaction of an amino acid (20) with ninhydrin

The second analytical tool was infrared (IR) spectroscopy. Evidence for the successful synthesis of the amino acid was the appearance of a broad absorption peak for the carboxylic acid -OH at 3436 cm^{-1}, and a sharp absorption peak at 3198 cm^{-1} for the $-NH_2$ group. A carbonyl absorption peak at 1623 cm^{-1} is also characteristic of zwitterion amino acids (**Scheme 3**)[76]. The NMR characterisation of **14** was performed as the Fmoc derivative **50**.[30,45]



50

Figure 9: Fmoc-trishomocubane amino acid

The protected amino acid **50** is not only useful for the characterisation of the tris-amino acid but it is also useful in solid phase peptide synthesis (SPPS) for the controlled coupling of **14** with other residues. Evidence for the successful synthesis of **50** was confirmed by infrared spectroscopy; a broad -OH absorption peak was observed at 3417 cm^{-1}, a sharp -NH absorption peak at 3358 cm^{-1} and carbonyl peaks at 1744 and 1711 cm^{-1}. The NMR data obtained were identical to those

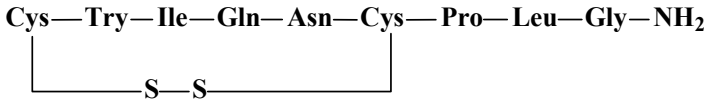
reported by Raasch *et al*. In the ^{1}H NMR spectrum of **50** the presence of four peaks from 1.47-3.40 ppm represents the four cage methine protons (-CH) and the peaks between 1-1.43 ppm represents the four methylene groups (-CH_2).[30,77,78,79] The multiplet between 7.00 and 8.00 ppm represents the aromatic protons of the Fmoc group. A broad peak between 10.0 and 12.5 ppm is typical of a carboxylic acid -OH. Note that the –OH peak between 10 and 12.5 ppm is only visible when DMSO(d_6) is used as a solvent but not when deuterated chloroform is used.

Two carbonyl peaks were present at 174 ppm and 155 ppm in the ^{13}C NMR spectrum. The methylene carbon peaks were observed at 31.7-32.7 ppm, the methine carbon peaks at 40.0-80.0 ppm, and the aromatic carbons were registered between 120.0-142.0 ppm. Note that the integration values of the aromatic protons are a bit higher than expected due to an overlap with the solvent peak. Further evidence for the successful synthesis of **50** was a negative ninhydrin test. In the mass spectrum the observation of a molecular ion $[M+H]^+$ = 428 m/z peak was taken as further confirmation of the successful synthesis of **50**.

Chapter 3

Peptide Chemistry

Peptide synthesis is an art, the practice of which nowadays requires its own book.[80] Peptides are compounds consisting of two or more amino acids which are linked together by amide bonds (peptide bonds). The peptide bond is of central importance in biochemistry and protein chemistry,[81,82,83] as it provides a linkage between amino acids in peptides and proteins. Peptides which contain 2 to 15 residues are called oligopeptides and those which contain more than 10 residues are called polypeptides/proteins. Oligopeptides have commonly been used as ligands and/or in drugs with high affinity and specificity for particular receptor sites.[84] Non-proteinogenic amino acids are vital in the future of peptide-related drugs.[29,46,85] It is widely believed that the incorporation of conformationally constrained amino acids is a good strategy for the design and development of selective drugs. Peptides play an important role in modern drug design and synthesis. Analysts have estimated the peptide market to be worth 12 billion US dollars and many people now consider peptides the drugs of the future. The term "peptide" was first proposed by Emil Fischer in 1901, and the first peptide, synthesised in 1901, was the dipeptide Gly-Gly (glycine-glycine) made by treating dioxopiperazine with mineral acid. The term peptide is similar to "peptones", which are products of protein degradation by pepsin. Peptides have very attractive advantages as they can be made highly specific and their potency can be increased, by simple amino acid substitution. Synthetic peptide chemistry reached its climax after the synthesis of the neurohypophyseal hormone oxytocin **51** in 1953 by V. du Vigneaud.[42,48,86]



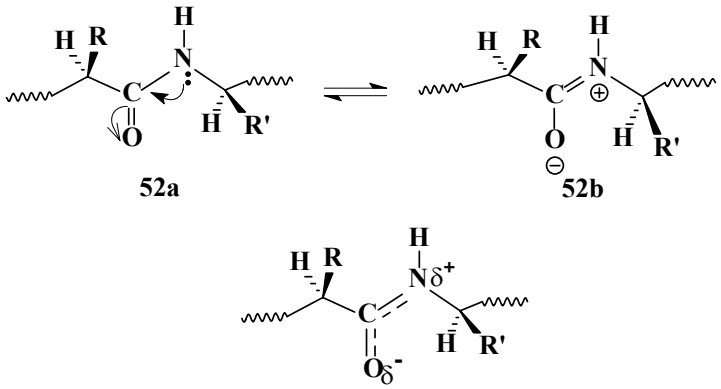
51

Figure 10: Oxytocin

This hormone has many functions but only few of them will be presented in this report. It is involved in the facilitation of birth and breastfeeding as well as in bond formation between people (trust). The hormone **51** is sometimes called the hormone of love, since it facilitates the emotional bond between a mother and child.[87]

In general a peptide bond is formed by nucleophilic attack of the N-atom of the second component on the carbonyl C-atom of the first. Crystallographic studies have shown that the amide bond (-CO-NH-) is shorter than the normal C-N single bond, due to the mesomeric nature of the amide bond which restricts the free rotation about the C-N bond.[88] The amide group is "flat" with the nitrogen

sp^2 hybridized. The hydrogen in N-H is normally in a *trans*-position with respect to the carbonyl oxygen **52** as this minimizes the steric hindrance between the R-groups of the two adjacent amino acid segments. Proline is the only amino acid where the N-H hydrogen is in a *cis* position.

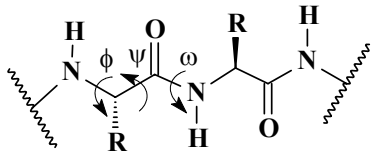


52a **52b**

52

Figure 11: Delocalization in peptide backbone

The backbone conformation of a peptide can be described by three torsion angles, ω (omega), ψ (psi) and ϕ (phi) illustrated in (Figure 12, **53**). The partial double bond character of the amide bond ensures that the C, O, N and H atoms of the amide bond lie in a plane (see **52** in Figure 11 above), with ω usually close to 180° except for proline, where ω is close to 0°. [48,88,89]

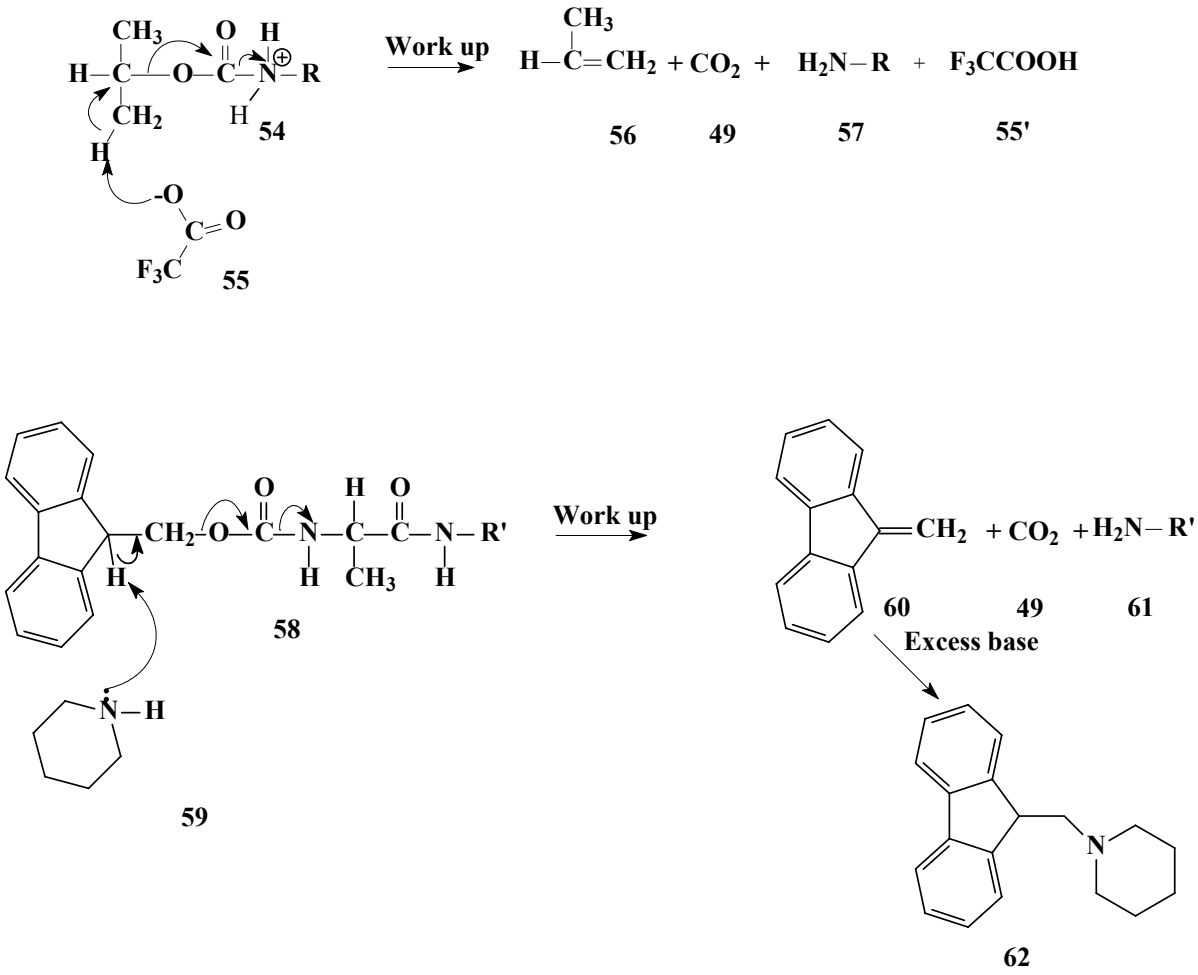


53

Figure 12: Rotation of the peptide backbone defining the ω, ψ and ϕ torsional angles

Peptide synthesis is sub-divided into two types: the classical (solution) method and the solid phase peptide synthesis (SPPS) method. The solution method essentially disappeared due to the difficulties associated with solubility and purification of the peptide as the number of amino acids residues increases.[90] To overcome these difficulties a new SPPS approach to peptide synthesis was initiated by R.B. Merrifield in 1963. Solid phase synthesis is based on the attachment of the protected amino acid of the chain to a linker, which is covalently bonded to a solid polymer (resin). Successive protected amino acids are added one at a time in a stepwise manner until the desired sequence is accomplished. In many textbooks N^{α}-protection (temporary protection) is achieved by the t-Boc-group; however in this dissertation the Fmoc-group was utilized for N^{α}-protection allowing acid activation through conversion of the acid function to an acid fluoride if required.

A t-Boc chemistry uses trifloroacetic acid (TFA) for primary deprotection of the t-Boc-group and TFA can also potentially alter sensitive peptide bonds. In Fmoc chemistry primary deprotection is achieved by a mild base such as piperidine. The advantage with Fmoc synthesis is that TFA is only used for the final cleavage of the crude peptide and deprotection of side chains such as the tert-butyl protecting group in the case of side chain amino acids (e.g. aspartic acid).[91] Fmoc based SPPS can be summarised into the following sections: solid support (resin), linker, attachment of the first residue, protecting groups, Fmoc deprotection, coupling reagents, monitoring, cleavage and removal of the protecting groups, peptide evaluation and peptide modification. The mechanisms for the deprotection of the t-Boc-group and the Fmoc-group are highlighted below in **Scheme 14**.



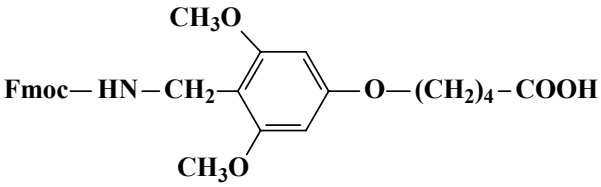
Scheme 14: Deprotection of t-Boc-amino acid with trifluoroacetic acid (TFA) and Fmoc-amino acid with piperidine

Essential to SPPS is that the polymer support (resin) must not be soluble in any of the solvents used, and must swell so as to facilitate the reaction. The advantage of this approach over the classical method is that, when the growing peptide chain is firmly attached to a completely insoluble support,

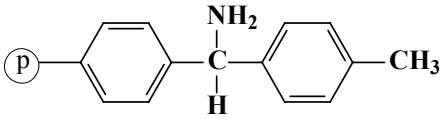
it is easy to filter off the solid resin and to wash away[‡‡] free reagents and by-products.[90,99] Purification of the peptide is not done by recrystallisation but by ensuring 100% coupling (negative ninhydrin test) and removal of the impurities, which simplifies the manipulations and shortens the time required for the synthesis of the peptides.

The development of the handles[§§] (linker) for the SPPS method has resulted in the successful synthesis of many useful peptides like Human Gastrin-I.[92] The linker must be soluble but not cleaved from the resin by any solvent used [i.e. N,N-dimethylformamide (DMF) and isopropanol] in order to allow added residues to couple to it. However it should be easily cleaved at the final deprotection of the crude peptide without destroying the desired product.[69,91,103] Alberico *et al.* have developed a linker (5-(4-(9-fluorenylmethyloxycarbonyl)aminomethyl-3,5-dimethoxy-phenoxy)-valeric acid, PAL) **63** which works well with the p-methylbenzhydrylamine resin (MBHA) **64** to produce peptide amides. Progress since Fischer's first dipeptide in 1901, especially the development of SPPS in the 1970s enables us to synthesise numerous complex peptides in the laboratory.

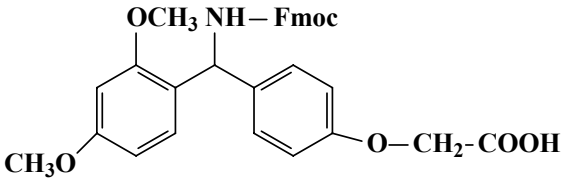
Most studies in Kruger's group at UKZN make use of p-[(R,S)-α–[1–(9H–fluoren-9-yl)methoxy-formamido]-2,4-dimethyl-benzyl]-phenoxyacetic acid (Fmoc-AM linker) **65** and the MBHA resin as it is stable in basic conditions. The linker **65** also produces peptide amides.[69,91,103]



63 PAL (Peptide Amide linker)



64 MBHA



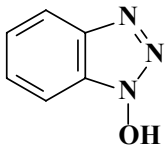
65 Fmoc-AM linker

Figure 13: Linkers and resin (PAL, MBHA and Fmoc-AM linker)

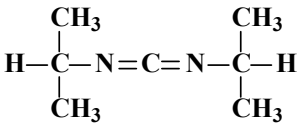
[‡‡] A typical washing procedure is: 3x DMF, 3 x Isopropanol and 3x DMF

§§ "Handles (linkers) are bifunctional spacers which serve to attach the initial residue to the polymeric support in two discrete steps. One end of the linker incorporates features of a smoothly cleavable protecting group, and the other end allows facile coupling to a previously functionalized support."

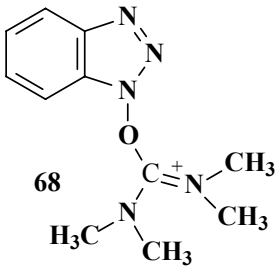
The coupling of the FmocAM linker **65** to the resin was achieved with the help of *in situ* coupling reagents like 1-Hydroxybenzatriazole (HOBt) **66**, N,N-Diisopropylcarbodiimide (DIPCDI) **67** and 2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) **68** (see Figure 14).[69,93]



66



67

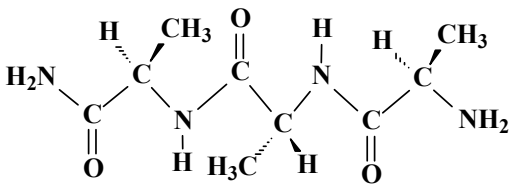


68

Figure 14: *In situ* coupling reagents (HOBt, DIPCDI, and HBTU)

HOBt was used as a catalyst in the reaction while N,N-Diisopropyldicarbodiimide (DIPCDI) acts as a dehydration agent, producing a soluble urea which is washed away with N,N-Dimethylformamide (DMF) and isopropanol. HOBt inhibits the formation of oxazolone (cyclic lactone) by forming an ester with the carboxylic acid group and also suppresses the racemization of amino acids. Other *in situ* coupling reagents like O-(7-Azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium-hexafluoro-phosphate (HATU) can also be used.

N,N-Dicyclohexylcarbodiimide (DCC) is not used in SPPS since it forms an insoluble urea which can block the polystyrene filter. The successful attachment of the linker is monitored with the ninhydrin test. The ninhydrin test is negative when the solution is red-brown and positive if the solution turns dark blue or purple. The TNBS (2,4,6-Trinitrobenzenesulfonic acid) test[94] can be used as an alternative, since some deprotected amino groups do not show the required violet-blue colour with the ninhydrin test.[95] The TBNS test[96] is positive when resin beads turn yellow or red within one minute and negative when the beads remain colorless.[94,95] As the synthesis of the peptides with SPPS required many steps, it was decided to first synthesize short peptides to gain experience with this technique. The first attempted peptide was a tri-peptide (Ala-Ala-Ala, **69**).

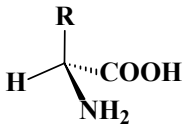


69

Figure 15: Ala-Ala-Ala (tri-peptide)

The resin **64** and a linker **65** were coupled in the presence of *in situ* coupling reagents **66** and **67** in DMF. The successful attachments were verified by a negative ninhydrin test. The Fmoc group was removed by 20% piperidine in DMF and a positive ninhydrin test (**Scheme 13**) was observed. This was followed by the successive addition of the three alanines in the presence of **66** and **67**. The peptide **69** was washed*** with dichloromethane, methanol and diethyl ether, followed by drying in a vacuum dessicator. Final cleavage was achieved using 95% trifluoroacetic acid, 2.5% triisopropylsilane and 2.5% water.[31,69] Triisopropylsilane was added as a carbocation scavenger. The peptide was filtered and cold diethyl ether was added until the precipitate of the crude peptide occurred. The peptide was centrifuged and washed several times with diethyl ether under nitrogen to remove the linker. The characterization of the tripeptide is discussed below in the results and discussion section.

Attempts to incorporate sterically hindered amino acids into peptides using classical (solution) methods have been made and have led to highly useful analogues of many bioactive peptides.[97] The effect of hydrophobic side-chains upon helix stability have been tested by the incorporation of non-natural hydrophobic amino acids such as α-aminobutyric acid (Abu, **70**), norvaline (Nva, **71**), and norleucine (Nle ,**72**) into peptides.[98]



70 R= CH_2CH_3
71 R= $CH_2CH_2CH_3$
72 R= $CH_2CH_2CH_2CH_3$

Figure 16: Non-natural hydrophobic amino acids (Abu, Nva and Nle)

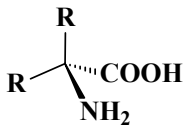
The most frequently used hydrophobic helix-stabilising amino acid is α-aminoisobutyric acid (Aib, **73)** which occurs in microbial peptides, such as alamethicin. Using crystallographic studies, Aib based peptides revealed an overwhelming preference for right or left-handed helices (3_{10} and α).

The incorporation of non-natural amino acids such as Aib **73**, adamantinine **24**, dipropylglycine (Dpg, **74**) and dibutylgylcine (Dbg, **75**) into peptides enhanced the scope of peptide design and protein engineering by imposing local restrictions on polypeptide chain stereochemistry, thus enhancing the stability of specific regions of the secondary structure like α-helix/β-sheets.[99] Further advantages may include:[57,99]

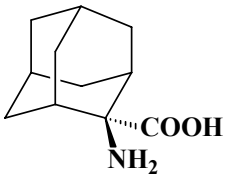
- enhanced binding to a molecular target,

*** A typical washing procedure is provided in the experimental section.

- improved oral bioavailability[100] and longer persistence in circulation in the case of biologically active peptides.



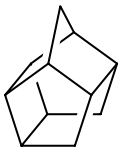
73 R = CH_3
74 R = $CH_2CH_2CH_3$
75 R = $CH_2CH_2CH_2CH_3$



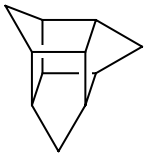
24

Figure 17: Non-natural amino acids (Aib, Dpg, Dbg and adamantinine)

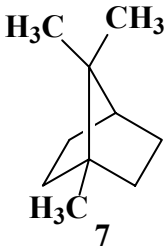
Computational studies predicted promising behaviour of cage amino acids when incorporated into peptides and is leading our synthetic approach at this stage (see the introduction in **Chapter 2**). The challenge is to verify the computational prediction[35,36,37,38,40] through a synthetic approach. This prompted an intensive synthetic effort at UKZN to incorporate the cage moieties such as **4**, **5** and **7** into peptides. It is also hoped that the cage amino acids would also improve the transportation of effective drugs to the central nervous systems (CNS).



4



5



7

Figure 18: Cage[†††] compounds (PCU, 4, trishomocubane, 5 and camphor, 7)

There have been various attempts in the research group at UKZN to incorporate **4** and **5** into oligopeptides. Govender attempted to incorporate the trishomocubane amino acid **14** into a dipeptide using an achiral glycine, but failed to purify the product using high performance liquid chromatography (HPLC). Raasch has attempted to make a tri-peptide with **14** using achiral glycine in the second (*i+1*) position and alanine in the third (*i+2*) position but could not purify it using HPLC. The use of glycine in the second position is suspected of assisting the formation of a six membered ring Schiff base (folding of peptide-**Scheme 15**) due to the lack of steric hindrance in glycine. It was subsequently decided to use alanine rather than glycine.

Bissety *et al*.[35,36,37,38,39,40] used force field and *ab initio* calculations to show that the incorporation of a cage **4** moiety into peptides induces an α-helix (*i, i+4*) mostly 3_{10}-helix (*i, i+3*) due to the steric interaction of the cage. It was also demonstrated[35,36,37,38] that the PCU amino acid is a better β-turn

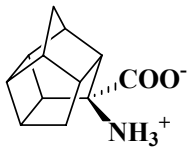
[†††] Camphor is strickly speaking not a classical cage compound, but is structurally related to the PCU skeleton.

inducer than Aib and could be compared to proline. Theoretical [Amber (parm94) and *ab initio,* rhf/6-31G(d)] results have predicted that the trishomocubane amino acid **14** containing peptides would potentially favour four low energy conformers, namely the C_{7ax}, C_{7eq}, 3_{10} and α_L helical conformations.[37,39]

There are several reasons for using alanine (Ala) instead of glycine (Gly) and other amino acids:

- Ala is a strongly helix stabilising residue. It is believed that the methyl side-chain reduces the conformational freedom of Ala, forcing it to adopt a helical conformation.
- Based on electron spin resonance (ESR) studies, Ala based peptides form a 3_{10}-helix (*i*, *i*+3) which is similar to β-turns (*i*, *i*+3) rather than an α-helix.[89,101,102]
- The methyl side-chain of Ala can inhibit the folding of the peptide to form a Schiff base (**Scheme 15**) due to the steric hindrance imposed by the methyl side chain.
- In contrast to other amino acids, the methyl group is small and steric clashes with other side chains are not too large to prevent effective coupling. In addition there is no loss of side-chain entropy on helix formation.

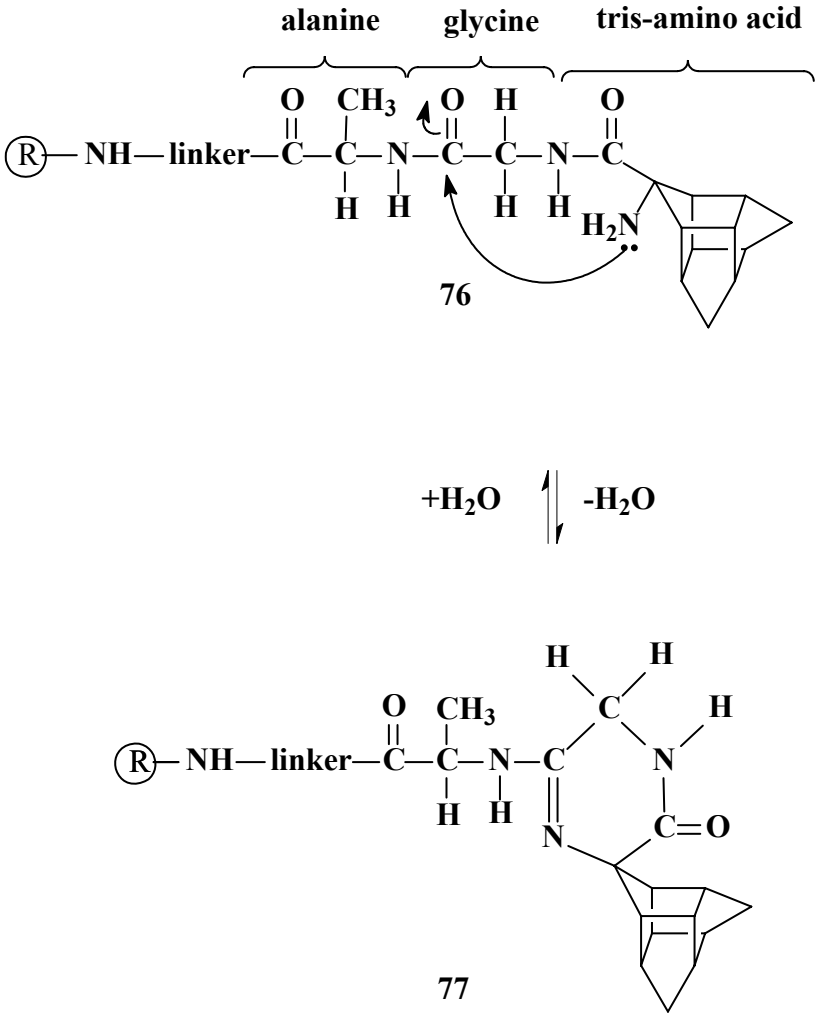
Mdluli[103] has attempted to use alanine (*i+1*) in the Ala-Ala-Ala-PCU, where PCU is the pentacyclo-undecane amino acid **11**, to prove that the methyl side chain of Ala prevents the peptide from excessive folding and subsequent Schiff base formation.



11

Figure 19: PCU amino acid

The fact that a positive ninhydrin test was observed by Mdluli after the incorporation of the cage amino acid **11**, implies that alanine in the second (*i+1*) position had prevented the formation of a Schiff base.



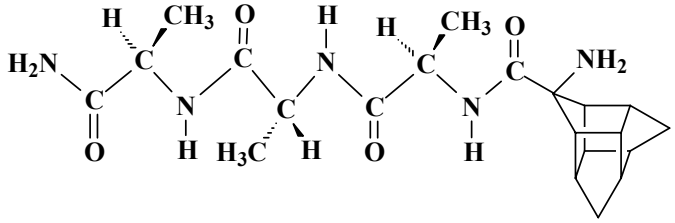
Scheme 15: Formation of a Schiff base

Purification of the peptide using preparative HPLC was found to be difficult due to the small amount of peptide obtained. Further characterisation of the peptide by crystallographic studies, to prove the helical (3_{10} and α) nature of alanine rich peptides and the effect of the hydrophobic cage (PCU) in α-helix to 3_{10}-helix and β-turns were not achieved.

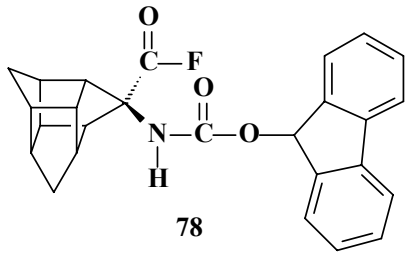
Results obtained from using the SPPS method are discussed below. The successful synthesis of the tri-peptide **69** (Ala-Ala-Ala) was verified by the presence of three carbonyl peaks in the IR spectrum at 1554, 1627 and 1677 cm^{-1}. The carbonyl peak at 1553 cm^{-1} is at low frequency since it is adjacent to an $-NH_2$ group. The appearance of two sharp peaks at 3406 and 3298 cm^{-1} (-NH stretching) also assisted in further verifying the successful synthesis of **69**. The ^{1}H NMR was difficult to interpret since the peptide was not dry enough. The peptide was recovered and dried again under high vacuum, but somehow it appears as if perhaps water of crystallization was still present. In the ^{13}C NMR spectrum three carbonyl peaks at 168, 172 and 175 ppm confirmed the successful synthesis of **69**. Three methyl peaks at 15.08, 15.41, and 15.98 ppm were also observed.

The purity (40%) of the peptide **69** was confirmed by HPLC analysis.‡‡‡ A full discussion of the HPLC conditions is given in the experimental section (**Chapter 6**).

At this stage the addition of tris-amino acid fluoride **78** to the tri-peptide **69** was attempted due to previously unsuccessful attempts to couple the Fmoc-tris-amino acid **50**. The coupling with halogenated Fmoc amino acid **78** was considered to be essential as acid halides are more reactive than carboxylic acids and to ensure successful coupling of the sterically hindered α,α-disubstituted tris-amino acid.[69,104,105] It was therefore essential to first synthesize the acid halide **78** from Fmoc-tris–amino acid **50**. The conversion is obtained through treatment of **50** with cyanuric fluoride and dry pyridine in dry methylene chloride.[30,45,69] Note that the detailed synthesis of **78** is discussed in the experimental section. Coupling of the Fmoc-tris-amino acid fluoride **78** was attempted to the tri-peptide **69** (still attached to the resin) along with pyridine, as an HF scavenger, to produce the tetra-peptide (Ala-Ala-Ala-tris, **79**). Deprotection of the N^{α}-protecting groups (Fmoc group) was achieved by treatment of the peptide with a 20% piperidine in DMF for 30 minutes (**Scheme 14**). A positive ninhydrin test (**Scheme 13**) was obtained after deprotection confirming the successful coupling of **78** to the peptide. The peptide **79** was stripped from the resin and the ^{13}C NMR analysis further proved the successful deprotection of the Fmoc group. This was confirmed by the disappearance of the Fmoc multiplets between 120 - 140 ppm. The peptide was then washed with DMF and isopropanol. It was prepared for cleavage by washing with dichloromethane, methanol and diethyl ether and dried under vacuum overnight. Final cleavage was achieved using a solution of 95% trifluoroacetic acid, 2.5% triisopropylsilane and 2.5% water.[31,69] The peptide was filtered and cold diethyl ether was added until the precipitate of the crude peptide occurred. The peptide was centrifuged and washed several times with diethyl ether under nitrogen as to ensure that the linker was removed from the peptide.



79



78

Figure 20: Ala-Ala-Ala-tris (79) and Fmoc trishomocubane (tris) amino acid fluoride (78)

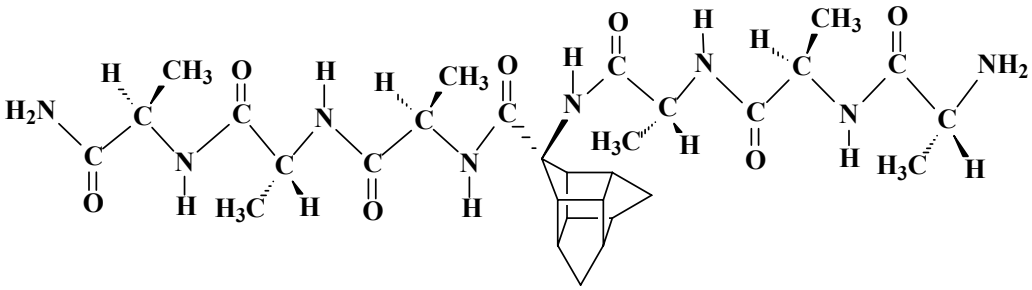
The peptide had some impurities which were confirmed by HPLC. The low yield of peptide **79** implies that the attachment of the cage amino acid to a peptide is problematic and the extension of reaction times as well as excess cage amino acid are required. The purity of peptide **79** was

‡‡‡ C18 Reverse phase column, 80% acetonitrile in water, a flow rate of 1ml/min and ran for 60 minutes in isocratic solution.

measured by HPLC analysis, which gave four peaks at 3.53, 4.69, 5.73 and 14.71 minutes. The single peak at 3.53 min corresponded to the unreacted tri-peptide **69** while the peak at 14.71 min is suspected to correspond to the hydrophobic/ less polar tetra-peptide **79**. It is expected that hydrophobic compounds elute later than hydrophilic compounds since they have a higher affinity for a C18 reverse phase column. The peptide **79** is supposed to be a diasteriomeric mixture as the trishomocubane amino acid **14** is synthesised as a racemate. At this stage it is not yet possible to say if one of the other HPLC peaks is due to the other diastereomer.

The successful synthesis of the tetra-peptide **79** was verified by the presence of four carbonyl peaks in the IR spectrum at 1548, 1624, 1677 and 1756 cm^{-1}. The carbonyl peak at 1548 cm^{-1} is at low frequency since it is adjacent to an $-NH_2$ group. The appearance of two sharp peaks at 3394 and 3298 cm^{-1} due to -NH stretching also assisted in confirming the synthesis of **79**. ^{1}H NMR was difficult to interpret since it again appeared as if water was trapped in the crystals. The appearance of five -NH peaks between 6.90-9.00 ppm confirmed the successful synthesis of **79**. In the ^{13}C NMR spectrum four quaternary peaks at 159, 171, 172 and 174 ppm were also observed.

Further work was undertaken, attempting to incorporate **78** in the centre of the hepta-peptide **80**. The same procedure used for the synthesis of **69** and **79** was followed.



80

Figure 21: Ala-Ala-Ala-14-Ala-Ala-Ala (hepta-peptide)

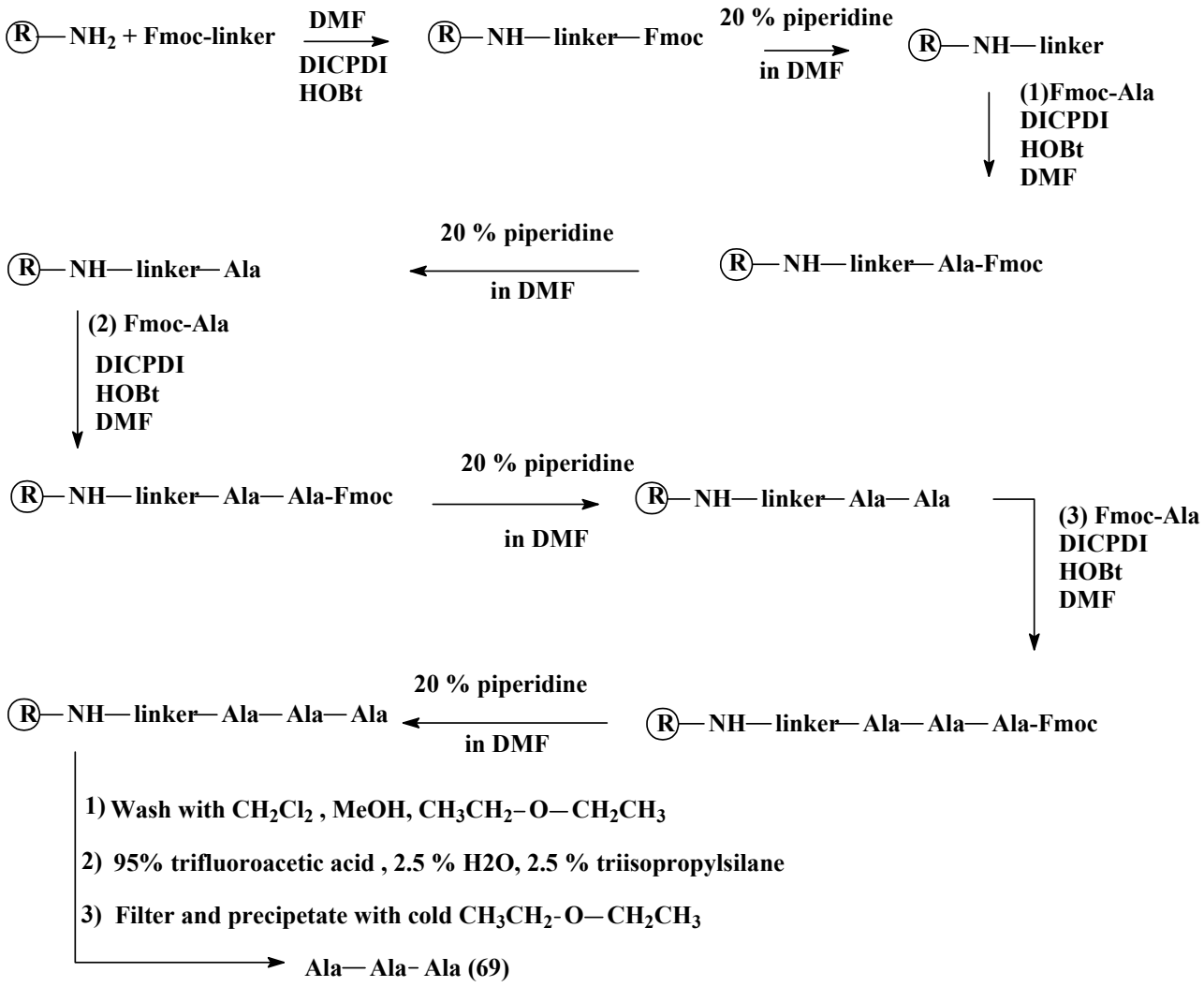
The synthesis of peptide **80** was performed in one attempt. Excess cage amino acid fluoride was used for the coupling of the trishomocubane amino acid, as well as extended reaction times.

According to the HPLC results the purity of **80** was low in the crude peptide mixture. Two peaks eluting at 2.76 min and 32.59 minutes were observed in the HPLC chromatogram. The peak at 2.76 min is suspected to correspond to the unreacted tri-peptide **69** and the peak at 32.59 min was due to the desired product **80**. The time required to elute **80** is different from the time to elute **79**, this may due to the different size of the two peptides. Therefore **80** has a higher affinity for the C18 reverse

phase column.[§§§] It was clear that the polarity of the solvent could be changed from 80% acetonitrile in water so as to minimize the time required to elute **80**, however due to time constraints only qualitative analysis of the peptide was undertaken.

The successful synthesis of **80** was also verified by IR spectroscopy which showed seven carbonyl peaks at 1544, 1558, 1622, 1634 (as a shoulder), 1653 (as a shoulder), 1678 and 1731cm^{-1}. The pure product may in practice only exhibit five carbonyl peaks in the IR spectrum, since the carbonyl groups in the amino acid segments *i+2* and *i+5* may overlap. The presence of two broad peaks at 3295 and 3400 cm^{-1} for –NH stretching also confirmed successful synthesis of **80**. The compound **80** had some impurities which were observed in the ^{1}H NMR spectrum. The appearance of seven peaks between 6.90-9.00 ppm (-NH peaks) also supported the evidence for the successful synthesis of **80**. The ^{13}C NMR (poor quality) showed seven quaternary peaks at 170, 171, 171, 171, 174, 174 ppm and the peak shifted upfield to 155.36 ppm. The peak at 155.36 could be an amide peak of the other terminal. The molecular ion peak of $[M+H]^+$ = 631 m/z in the mass spectrum further confirmed successful synthesis of **80**. Fragmentation in the mass spectrum corresponded well to the breaking of the peptide chain. For example a peak at m/z = 560 correspond to the $[M\text{-Ala}]^+$ and a peak at m/z = 489 correspond to the $[M\text{-(Ala-Ala)}]^+$. However the other peaks become more complex in the spectrum. A schematic summary of the peptide synthesis is depicted below (**Scheme 16**).

§§§ "In a reverse phase column the more hydrophilic compounds elute more quickly than do hydrophobic compounds"



Scheme 16: Summary for the synthesis of Ala-Ala-Ala (69)

A similar procedure was used for the synthesis of **79** and **80**. Schiff base formation (see **Scheme 15**) was prevented by using alanine in the second (*i+1*) position of both reported peptides. It was found difficult to do structural elucidation by NMR due to the impurities observed in ^{1}H NMR spectra and the wetness of the peptides. It is necessary to dry the peptide using a lypholizer for NMR studies since a broad water peak appeared in all NMR spectra. Purification of the peptides by Preparative HPLC would require larger quantities of the peptides.

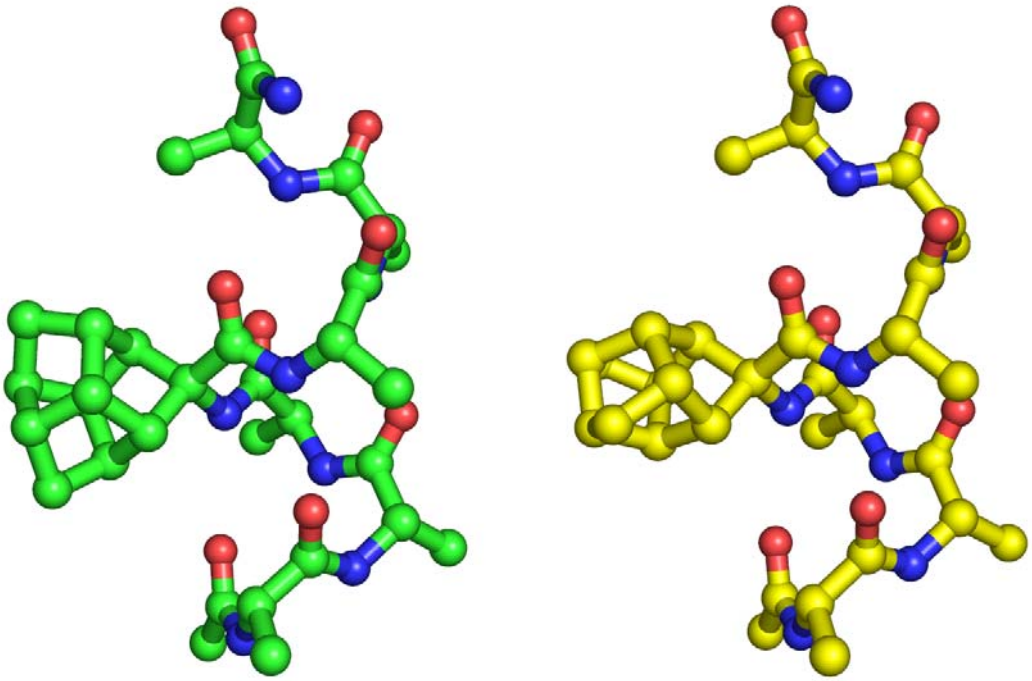
As part of collaboration between the research group at UKZN and the Uppsala University[106] in Sweden, a sample of the amino acid **14** was sent to them. The results obtained in this investigation (UKZN) was communicated to the Uppsala group. The hepta-peptide was successfully synthesized and separated by a former student from the UKZN research group, Dr T. Govender, now working for the Uppsala group. Since the trishomocubane amino acid **14** has D_3 symmetry and is synthesized as a racemate, two products were separated as Ala-Ala-Ala-(R/S)-**14**-Ala-Ala-Ala peptides. The unique symmetry of the cage structures ensures that only two enantiomers are

formed due to the substitution at quaternary carbon see Figure 7, instead of four diastereomers. The peptide was synthesized using the normal SPPS method[****] with a different resin than the one used at UKZN. The two diastereomers were separated by HPLC.[††††] Crystals of both peptides were grown and the diastereomers were analysed using a single crystal X-ray crystallographic analysis.[107] The X-ray results obtained for both enantiomers are presented below. It was some what surprising that both diasteromers produce essentially the same peptide backbone (see Figure 23). Only one of the enantiomers was investigated computationally so far and one would have expected a bigger difference between the peptide backbones of the two diastereomeric peptides.

It is important to note the two diastereomers were separated by approximately 4 minutes in the HPLC spectrum.

[****] Sieber amide resin, Fmoc amino acid (1.5 eq), HATU (3 eq), DIPEA (6 eq), 90 min, TNBS test. Cleavage 95:2.5:2.5 TFA:TIS:water for 1 hour.

[††††] Preparative HPLC on a C8 column, 5% CH_3CN to 95% CH_3CN in 30 minutes and run for a further 15 minutes. Aprox. 4 minutes.



Cage1, Orthorhombic, $P2_12_12_1$
a = 10.7690(5), b = 12.4820(4), c = 26.6490(10) Å
R-factor = 0.0522

Cage2, Orthorhombic, $P2_12_12_1$
a = 11.0100(3), b = 12.3470 (3), c = 26.7020(4) Å
R-factor = 0.0435

Figure 22: X-ray structure of Cage 1& 2 (Ala-Ala-Ala-tris-Ala-Ala-ala)

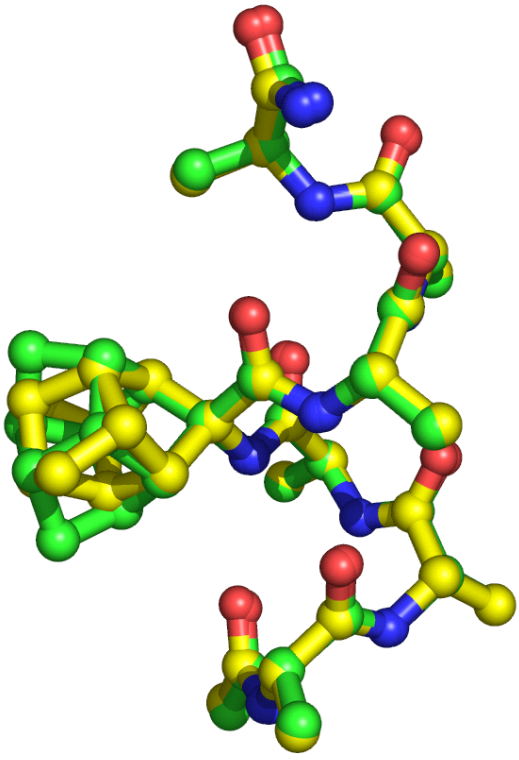


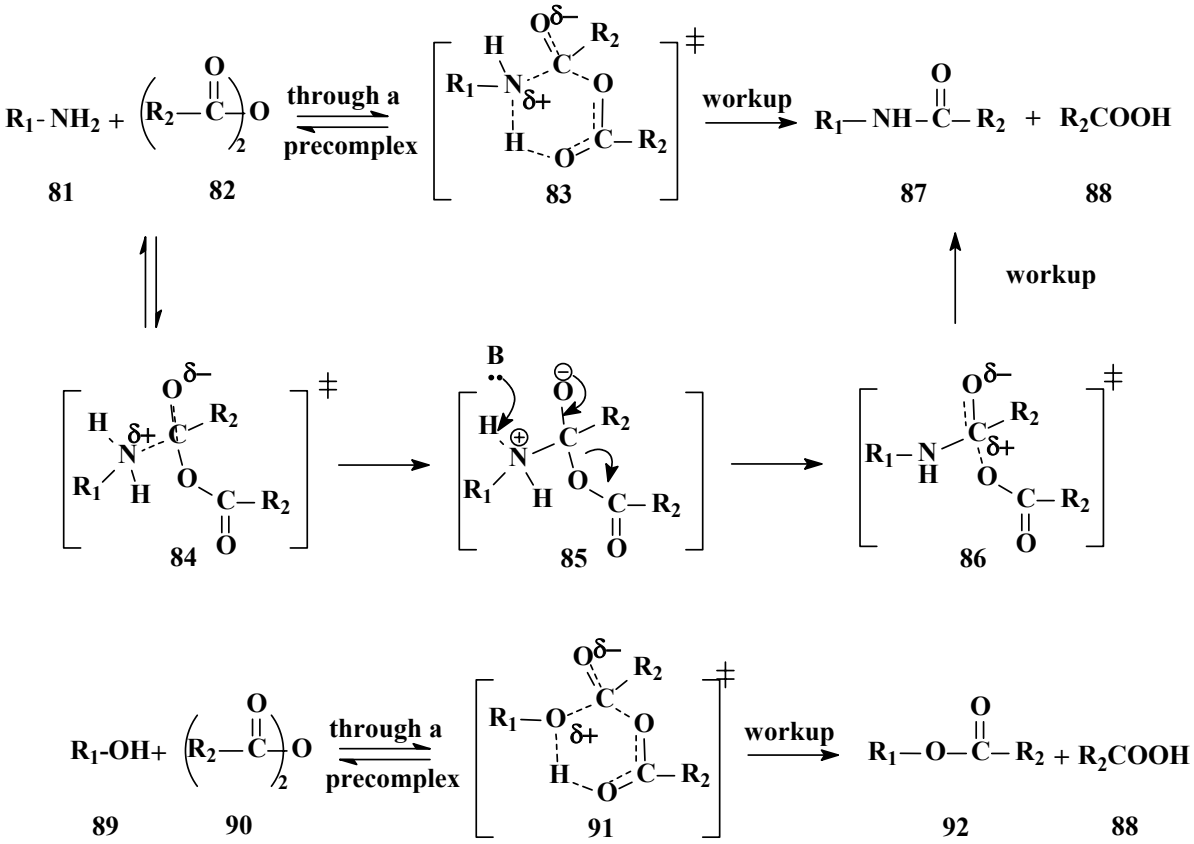
Figure 23: Overlay of the two X-ray structures

Chapter 4

Computational chemistry studies and theory

As indicated in the introduction (**Chapter 1**), a simplified computational model for the esterification of acids was required to enable a better understanding of the mechanism of esterification. Once that is achieved, the model could be extended to investigate more complex systems such as the cage amino acids **11** and **14**. Kruger[108] proposed a unique cyclic transition state for the acetylation of amines and alcohols with acetic anhydride. The purpose of this computational study was to extend the model to include acetylation of alcohols with acid chlorides and carboxylic acids.

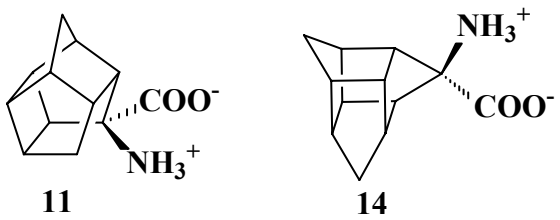
The model developed by Kruger was based on *ab initio* calculations. A mechanism via a six membered ring cyclic transition state (**83**) between alcohols, amines and the incoming acetic anhydride was postulated (**Scheme 17**). This project was then extended to include the regioselective acetylation of a pentacyclo-[5.4.0.0^{2,6}.0^{3,10}.0^{5,9}]-undecane lactam (δ-lactam, **96**) and various hydantoins (5-methylhydantoin **100**, 5,5-dimethylhydantoin **101**, and pentacyclo-[5.4.0.0^{2,6}.0^{3,10}.0^{5,9}]-undecane hydantoin, **102**) with acetic anhydride.[109,110]



Scheme 17: Cyclic mechanism vs acyclic mechanism

Esterification reactions are performed daily in many research laboratories throughout the world. In C→N peptide synthesis methods, esterification (protection of acid side chains) has been used for the protection of carboxylic acid groups of amino acids, since none of the amino acid forms (i.e. cationic, zwitterionic, and anionic, **Scheme 3**) which are present in solution at any pH are viable candidates for peptide bond formation.[111,112] This is due to the presence of the protonated amino or the anionic carboxylate group, and it is only by esterification that the neutral form of the amino function dominates in solution at neutral pH.[113] Esterification enhances peptide bond formation most likely by transforming amino acids to neutral species, in which nucleophilic attack of the amino nitrogen on an electrophilic carbonyl carbon is more feasible than for zwitterionic or anionic amino acids.[114] The N→C peptide synthesis, which required protection of the acid function via esterification, has disappeared since esters are good leaving groups which results in the formation of diketopiperazines (cyclic dipeptides).

Side chain carboxylic acid group protections are also performed in SPPS by alcohols such as adamantan-1-ol, adamantan-2-ol and by isobutene (2-methyl-1-propene). Amino acids with side chains such as aspartic acid (Asp) and glutamic acid (Glu) are protected by alcohols for C→N synthesis.[115] The common side reactions that occurs when coupling unprotected Asp is intramolecular cyclisation to form succinimide and β-aspartyl peptides.[116,117] Glutamic acid also undergoes cyclization to form γ-glutamyl peptides. In these amino acids (Asp, Glu) side chains need to be protected since the desired peptide may be not achieved if they are unprotected. In Fmoc chemistry the side chain protecting groups which are used must be stable in basic conditions for example adamantan-1-ol and isobutene.[115,118,119] So, apart from an interest in understanding the mechanism of esterification to ultimately explain the difference in reactivity between the acid functions of **11** and **14**, peptide chemistry as a field could also benefit by a better understanding.



11 **14**

Figure 24: PCU(11) and trishomocubane amino acids (14)

There are various important questions that need to be answered with regards to the understanding of the esterification mechanism of alcohols with acids and acid chlorides:

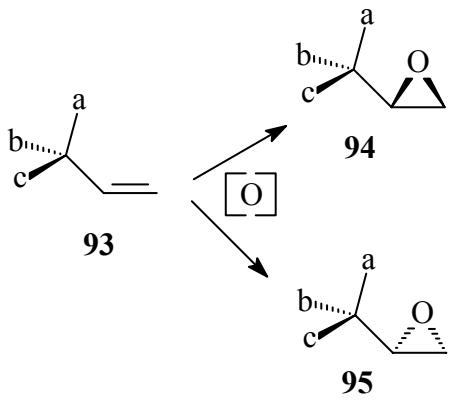
- Can a six membered ring transition state be used as computational model for these reactions?
- A four membered ring transition state should be considered. How will the activation energy of a four membered ring compared with a six membered ring transition state (if possible)?
- Does bond formation occur in a concerted or stepwise mechanism?
- What is the molecular and electronic structure of each transition state?
- What are the associated activation energies?
- What are the effects of entropy on the mechanism?
- What is the nature of solvent effects on the mechanism?

Computational simulations were undertaken with acetic acid, acetyl chloride, and methanol so as to understand the mechanisms and to investigate a similar six membered ring transition state. The field of computational chemistry is a field that can be said to be both old and young.[120] The fact that its foundation was laid with the development of quantum mechanics in the twentieth century makes it old. It is young since there is no technology in human history that develops at a pace like digital computers. Computational chemistry does not only cover the quantum mechanics but it also covers molecular mechanics, minimizations, simulations, conformational analysis, and other computer-based methods for the understanding and the prediction of the behaviour of molecular systems.[121] Theoretical chemistry and computational simulations have gained much attention as powerful computers and user-friendly software packages have become more widely available.[122]

Computational chemistry provides solutions that relate the laboratory data (e.g., IR, UV/vis, NMR spectra, heat capacities, reaction cross-sections) to molecular properties (e.g., geometries, bond energies, activation energies) under investigation. Another important application is that it can calculate the shapes of various orbitals (e.g., HOMO and LUMO) of the molecule. All the simulations of chemical structures carried out in computational chemistry use mathematical methods and the fundamental laws of physics.

Computational chemistry has become a powerful tool in many research fields including the pharmaceutical industry and the academic chemistry environment. It can provide important information about the kinetic and thermodynamic control of a reaction, like the energy barriers relative to reactants between intermediates and products.[123] Thermodynamic and kinetic data, such as the energies of starting materials, intermediates, transition states and products, are difficult to

obtain by experimental techniques.[114,124] Computational chemistry offers a unique opportunity to enhance our understanding of basic chemistry. For example when there are two competing reaction pathways from the same starting material, one can use computational modelling to picture the kinetics (activation energy) of the respective transition states‡‡‡‡ leading to the different products. The products can be optimized and the relative stabilities provide a measure of the thermodynamic stability of the products. For example, a reaction was modelled where a single enantiomer of an alkene (having a stereogenic centre which was directly attached to the prochiral double bond) was experimentally observed to produce two competing products when it was treated with an achiral epoxidising agent (**Scheme 18**).



Scheme 18: Epoxidation of a prochiral (93) alkene

The rate of reaction was dependant upon the energy barriers (activation energies) of the transition states between these two competing reactions. The energy profile of the computational model showed that the epoxidation to **95** was both kinetically and thermodynamically favoured. This may seem easier than it really is, since it requires a lot of organic chemistry knowledge and experience to visualize the nature/geometry of the transition state connecting the alkene with the epoxide. Fortunately there are strategies available to assist the novice computational chemist to find difficult transition states. Some of these strategies will be discussed below.

Modelling of chemical structures is based on two categories which are the backbone of computational chemistry viz: Molecular Mechanics (MM) or Force Field methods and Electronic Structure methods. They both perform the same basic types of calculations among others, the following are basic examples:

- Computing the energy of a particular molecular structure.
- Performing geometry optimization, which locates the lowest energy molecular structure.

‡‡‡‡ A transition state is a stage that link reactants to products in a reaction. A transition state is a maximum on the energy surface of the system.

- Computing the vibrational frequencies of molecules resulting from interatomic motion within the molecule.

The two methods can differ in that the Molecular Mechanics methods perform computations based on the interaction among the nuclei, while the Electronic Structure methods consider electronic effects (electron density) for its calculations.

Molecular mechanics performs simulations using laws of classical physics e.g. Newton's mechanics (eqn: 1) to predict the structures and properties of molecules.[123,125,126]

$$F = m\mathbf{a} \qquad (\mathbf{1})$$

This equation forms the basis of Hook's law, which is again the closer basis of MM. The molecule is seen as a collection of hard spheres (the nuclei) connected by flexible springs (the chemical bonds). Electrons are not taken into account explicitly and, as a result, thermochemical properties such as the heat of formation, ΔH_f^o, has not been easly calculated until the period of 1976-1977, when standard powerful methods for studying molecular structures and related properties were developed.[127,128] Molecular mechanics is a fast method that uses minimal computer resources. They are useful tools in the study of very large molecules like proteins and nucleic acids. The foundation of molecular mechanics methods is due to the observation that molecules tend to be composed of units which have structural similarities in various molecules. For example, all C-H bond lengths are roughly constant in all molecules, between 1.06 and 1.10 Å. The C-H stretch vibrations are also similar, between 2900 and 3330cm^{-1}, which implies that the C-H force constants are also similar.

Molecular mechanics methods are available in many computer programs including *MM3*,[128,129] *HyperChem*,[130] and *Alchemy*.[131] Each of these programmes is characterised by its particular force fields. The common element of all MM programmes is that the calculation of energy is related to atomic characteristics by force constants via an equation.

Electronic structure methods use the laws of quantum mechanics rather than classical mechanics since electrons are very light particles (compared to nuclei) that cannot be described quantitatively by classical mechanics. Quantum mechanics states that the energy and related properties of a molecule may be obtained by solving the Schrödinger equation (eqn: 2)

$$H\psi = E\psi \qquad (\mathbf{2})$$

Where, H is the Hamiltonian operator:

E is the energy.

ψ is the wave function.

There are two major classes of electronic structure methods that are used for modelling viz:

- *Semi-empirical methods* such as AM1, MINDO/3 and PM3 which are implemented in programmes like *MOPAC*[132], *AMPAC*, *HyperChem* and *Gaussian*[133]. These methods use parameters derived from experimental data or high level *ab initio* results to approximate the solution to the Schrodinger equation.
- *Ab initio*[§§§§] methods use no experimental parameters in their computations and are also implemented in programmes like *GAMESS*,[134] *Gaussian* and *Jaguar*.[135] Its mathematical algorithms are based solely on the laws of quantum mechanics and on the values of small physical constants such as the *speed of light*, the *mass & charge of electrons* and *Planck's constant*.

Examples of *ab initio* methods are Hartree-Fock (HF) theory, second order Møller-Plesset (MP2) [136] and Coupled Cluster theory[137,138]. *Ab initio* methods compute solutions to the Schrödinger equation using a series of rigorous mathematical approximations. The Born-Oppenheimer approximation is an important approximation which enables electronic and nuclear motions to be separated. In the Born-Oppenheimer approximation the change in electronic motion is a function of the nuclear shift, so that a small change in nuclear position can result in a change in electronic position and also a change in the energy of the molecule. *Ab initio* methods do not solve the Schrödinger equation exactly but an algorithm is used that can lead to a reasonable approximation to the solution of the Schrödinger equation. An exact solution to the Schrödinger equation can be obtained for small molecules like H_2.

The algorithm makes use of a basis set, which is a mathematical description of the orbitals involved. A more elaborate discussion on the nature of basis sets will be provided below. One can use primitive basis sets, which use severe approximations, or one can use better basis sets that represents a more realistic description of the orbitals. It is important to note that a method (*semi-empirical* or an *ab initio* method) and a basis set which are adequate for one application may be inadequate for another. The disadvantages of *ab initio* methods relative to other methods are with respect to computational resources and speed.[139]

Density functional theory (DFT) method is another electronic structure method which has gained wide application in the field of computational chemistry. It remains controversial as to whether DFT is an *ab initio* method or not, but it uses the same CPU resources as Hartree Fock (HF) theory, the least expensive *ab initio* method according to computer time. DFT methods are based on the theory by Hohenberg and Kohn[140,141] in which they demonstrated that the ground state energy of

any molecule can be described in terms of the total electron density, which means that each molecule has a unique functional form which exactly determines the ground state energy and electron density of the molecule.

DFT approximates the effect of electron correlation***** as it calculates, as accurate as possible, the total electron density of the molecule. HF does not explicitly calculate this effect but considers it in an average sense.[142] The effect is that HF methods result in molecular orbitals where the electrons are slightly closer together on average. This again results in an increase in the potential energy of the system due to an increase in electron-electron repulsion.

The first *ab initio* level of theory that specifically calculates the effect of electron correlation is second order Møller-Plesset (MP2) perturbation theory. In the DFT method the energy is expressed as a function of the density of a uniform electron gas E (ρ) rather than as wave functions. This was then modified to express the electron density around the molecules.[123,143] Some examples of density functional theory include BLYP and B3LYP.[144,145] All the above methods are defined as the level of theory. A "level of theory" cannot work on its own to perform calculations since the molecular orbitals (MOs) have not yet been described. All electronic methods make use of basis functions. Basis functions are sets of mathematical functions which are designed to describe the MOs and to provide maximum flexibility (subject to the cost of doing the calculation) to the molecular orbitals. The basis functions are collectively called basis sets. They are derived from the following equation (eqn: 3):

$$\psi_i = \sum^{n}_{\mu=1} C_{\mu i}\phi_{\mu}. \qquad (3)$$

Where, $C_{\mu i}$ is a molecular orbital expansion which is chosen to optimize the shape of the basis function sum and to ensure normalization. Lastly, ϕ_{μ} is a basis function of atomic orbitals.

Increasing the flexibility of the basis set allows the mathematical algorithm, which solves the wave function of the molecule, to obtain lower energies as restrictions to the space occupied by the MO are removed.

4.1 Choosing a basis set

In early *ab initio* programs, it was a very tedious and error-prone task to specify a basis set for computations. The user was required to punch several dozen floating point numbers into specific columns of the input deck without a single typing error. With current programs like *Gaussian*,

---

§§§§ The term *ab initio* is a latin word which means "from the beginning"

***** Electron correlation is the phenomenon where the electrons in a molecular system avoid each other in the orbital and attempt to stay as far as possible apart.

HONDO,[146] *GAMESS*, and *Jaguar* the user may select from among a wide variety of internally defined basis set by specifying the appropriate keyword (usually the basis set name).

4.2 Types of basis sets

The most widely calibrated basis sets for general purpose usage are those of Pople and co-workers, which range from the small STO-3G minimal basis set all the way up to basis sets with extra diffuse functions. Besides their wide usage, these basis sets offer the computational advantage that, when used with computer programs that exploit their features, the calculation of the second and first derivative integrals is faster than for most other basis sets, because exponents are shared between s and p functions.[123,139,147]

There are mainly five types of basis set, *minimal basis sets* (contracted basis sets), *split valence basis sets*, *polarised basis sets*, *high angular momentum basis sets* and *basis sets for post-third-row atoms*.

Minimal basis sets contain the minimum number of basis functions needed for each atom. They use fixed-sized (contracted) atomic-type orbitals e.g. Slater type orbitals (STOs) and Gaussian type orbitals (GTOs). Examples are STO-3G, MINI-1 and MINI-2. STO-3G uses three Gaussian (3G) primitives per basis functions after Pople *et al.*[148] discovered that the optimum speed and accuracy (compared to calculations using STOs) of calculations was achieved for M=3, where M is the number of Gaussians used in the linear combination of atomic orbtals. The STO-3G basis set uses Slater type orbitals (STO). This basis set was initially developed for the elements H-Ne, but as time went on it was expanded to cover Na-Ar, P-Kr, and Rb-Xe.[149,150,151,152]

In *Split valence basis sets* the number of basis functions per atom is increased which makes the basis set larger and increases its flexibility. Examples are 3-21G, 6-21G, 4-31G, 6-31G, 6-311G and D95.[153] The nomenclature is a guide to the contraction scheme. The first number indicates the number of primitives used in the contracted core functions. The numbers after the hyphen indicate the numbers of primitives used in the valence functions. If there are two such numbers, it is a valence-double-ζ basis. Split valence basis sets have two sizes of basis functions for each valence orbital, e.g. hydrogen and carbon can be represented as:

H: 1s, 1s'.

C: 1s, 1s', 2s, 2s', $2p_x$, $2p_y$, $2p_z$, $2p_x$', $2p_y$', $2p_z$'.

It is important to note that primed and unprimed orbitals differ in size.

Polarised basis sets add polarisation functions into split valence basis sets. Split valence basis sets only expand (change the size) the orbitals but do not change the orbitals shapes. Polarized basis sets remove this limitation by adding orbitals corresponding to one quantum number of higher angular momentum than the valence orbitals. They add *d* functions for atoms like C, N and O, while some add *p* functions to hydrogen atoms and *f* functions to transition metals. It is important to note that the *d* functions that are added to C, N and O atoms have similar sizes to 2*p* orbitals and *p* functions in hydrogen atoms are also similar in size to the 1*s* orbital and so on. The primary purpose of the polarization functions (d and p functions) is to give additional flexibility to the linear combination of atomic orbitals (LCAO) process in forming bonding orbitals between pairs of valence atomic orbitals (1s, 2s, 2p, etc). Polarization functions are essential in strained ring compounds because they provide the flexibility needed to direct the electron density into regions between bonded atoms. They can also be used for accuracy in unstrained compounds. An important point is that most basis sets are defined to use five spherical d functions, except 6-31G(d) which uses six Cartesian d functions. The six Cartesian d functions are d_x^2, d_y^2, d_z^2, d_{xy}, d_{xz}, d_{yz} and pure spherical d functions are $d_{z^2-r^2}$, $d_{x^2-y^2}$, d_{xy}, d_{xz}, d_{yz}. Examples of polarized basis sets include 6-31G(d) and 6-31G(d,p).

Since highest energy molecular orbitals of anions, molecules with lone pairs, highly excited electronic states, and lose supermolecular complexes tend to be much more spatially diffuse than normal molecular orbitals, diffuse functions are needed to minimise errors in energy calculations of such systems. To minimise this error in energy, standard basis sets like polarized basis sets [3-21G(d), 6-31G(d), etc] are often augmented with *diffuse basis functions*, which are indicated by adding '+' in the basis set name. Diffuse functions are typically a better treatment for cases where negatively charged species are being studied because the ionic radius of an anion increases significantly as compared with the corresponding neutral species. Diffuse functions are large versions of s- and p-type functions. Thus, 6-31+G(d) indicates that heavy atoms have been augmented with an additional s and an additional set of p functions. A second '+' in 6-31++G (d) indicates the presence of s functions on H as well.

High angular momentum basis sets add multiple polarization functions per atom to the triple zeta basis set. For example, the 6-31(2d) basis set adds two d functions per heavy atom instead of one *d* function. The 6-31++G(3df, 3pd) basis set contains three sets of valence region functions, diffuse functions on both heavy atoms and hydrogens and multiple polarization functions (3p functions and 1d functions) on hydrogen atoms. These basis sets are useful for describing the interactions between electrons in electron correlation methods. They are not generally needed for HF calculations.

Basis sets for post-third-row atoms are treated in a different way from normal basis sets. In these basis sets for very large nuclei, electrons near the nuclei are treated in an approximate way, via effective core potentials, a method developed by Hellmann in 1935.[154] Examples of these basis sets are LANL2DZ, LACVD, LAV2D and LACV3P, which are available in programmes like *Jaguar*.

4.2.1 The Schrodinger equation and the energy surface of the molecule

The aim of solving the Schrodinger equation is to obtain information about the structure and energy of the molecule being investigated. These theoretical calculations mimic more closely the path of the reaction in experiments. In order to get a better understanding of the reaction, all possible structures linking reagent to product should be considered. The exploration of all structures corresponds to the characterisation of a full potential energy surface (PES) of the reaction path. A potential energy surface is the energy obtained as a function of reaction coordinates in a chemical structure (i.e. rotation of single bonds and dihedral angles). PES is normal taken as the potential energy in a two-dimensional energy profile.

The classical example is the PES for butane found in text books.[155] Below is an example of the PES for dibromoethane.

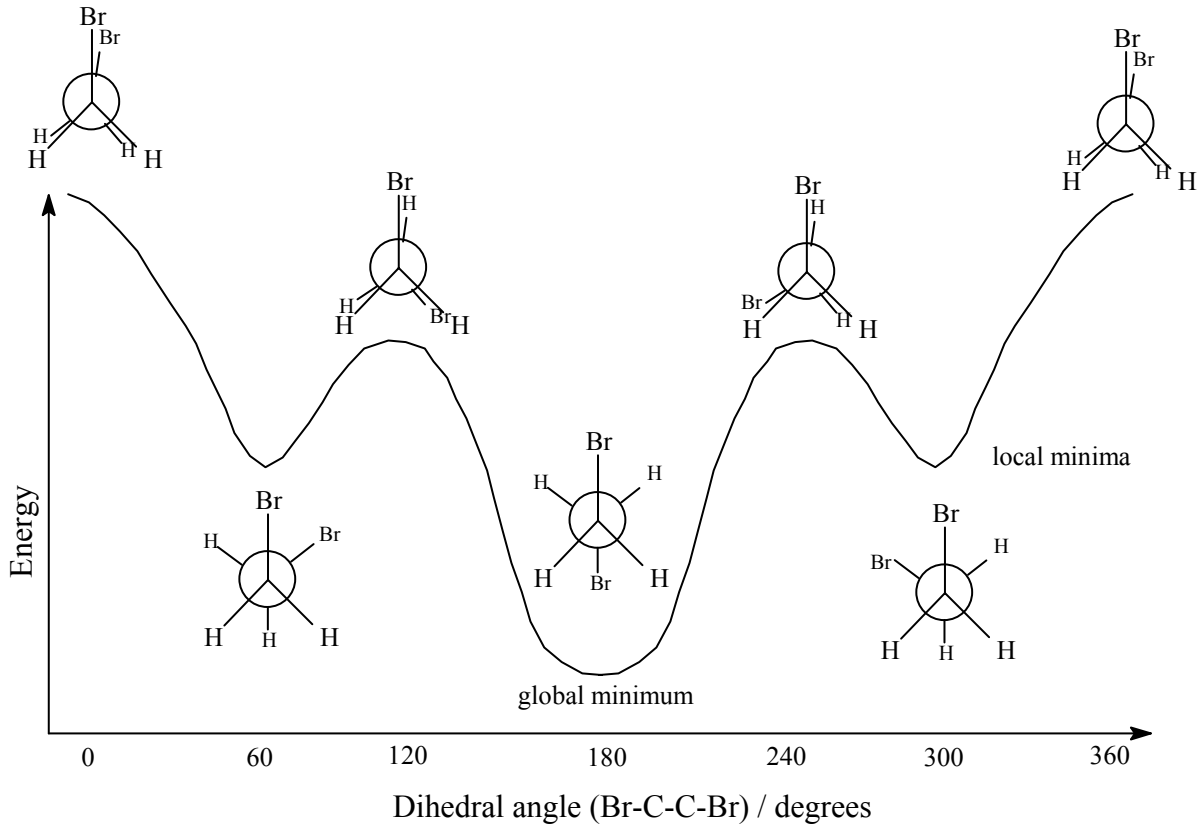


Figure 25: Energy profile of dibromoethane

The staggered conformation (global minimum) has the lowest energy since the two bromo atoms experience the least steric hinderance and is sometimes called an *anti* comformer. A *gauche*

conformer is higher in energy and is called a local maximum. In cases of peptide modelling where there are many minimum structures, the structure that is native is referred to as the global minimum. The eclipsed conformers are maxima on the energy profile and are also transition states.

4.3 The optimization algorithm

The optimization algorithm makes generally use of the energy surface of the specific molecule under investigation. The algorithm calculates the slope of the energy surface and in the case of the search for a minimum, it will take a step towards the minimum. The algorithm will then re-optimise the structure by moving the atoms so that a structure with lower energy is obtained. The slope of the energy surface is again calculated and the next "down hill" step is taken. This is repeated until the slope eventually is reversed due to an overstepping of the minimum on the energy surface. The algorithm will then take a step backwards and the process is repeated until the lowest energy structure is obtained.

In the case of searching for a transition state, the same principle is used, except that the step is taken "up hill" as the transition state is represented by a maximum on the energy surface.

Various modifications to the basic optimization algorithm exist which are designed to increase the efficiency of the optimization in terms of speed and quality. It is important to note that most optimization algorithms are sensitive enough to be trapped in local minima. If one starts with a transition state, a normal optimization searching for minima may even get trapped at the maxima and find exactly the same transition state structure. Optimization algorithms such as Monte Carlo are designed not only to take a step down the energy surface, but also to randomly take a step in any direction. This will enable the algorithm to find minima lower than a local minimum.

4.4. Searching for a transition state

A transition state (TS) is a molecular species that is represented by the maximum on a potential energy curve in a two-dimensional reaction co-ordinate diagram. In experimental cases it is difficult to predict the geometry of a transition state, however nowadays there are spectroscopic methods which can be used to determine the structure of the transition state, such as femtosecond pulse laser spectroscopy, although this is limited to specific systems.[69,121] Therefore the whole process of transition state modelling still depends heavily on computational methods. In order to locate maxima (transition states), it is necessary to have fairly accurate initial geometries and matrices of the second derivatives of the energy (e.g. Hessian matrix). The reason for this is because a typical optimization algorithm performs a crude guess at the initial solution to the wave function of the molecule under investigation. The optimization algorithm then starts a series of "up hill" iterations to solve the wave function. If the initial structure is too far away from the correct

transition state, then the initial guess would not locate the correct wave function of the specific transition, leading to an incorrect solution.

In order to obtain a transition state, it is important to know which reaction coordinates will need to be controlled in order to get the maximum. Typical reaction coordinates may be bond breaking/formation, dihedral angle, angle or a combination of these. A program for modelling structures between reactant and product could be used to perform a linear search for the lowest maximum on the energy surface. A SCAN[†††††] calculation is used as the most convenient method to obtain a transition state, where an accurate starting structure for a transition state is found at the maximum of the potential curve. The SCAN jobs increase/decrease bond length, angle or dihedral angle in a stepwise manner. In each step of a SCAN calculation the molecule is optimised while only constraining the specified parameters. Needless to say the method is quite expensive in terms of time. A practical hint is to first use a crude scan where the step size is quite large, and then to "zoom" in closer to the maximum by using smaller step sizes in a subsequent scan job.

The structure closest to the maximum on the energy surface is manually extracted and used as a starting structure for a non constrained transition state optimization.

A frequency calculation is performed so as to confirm that the maximum structure is really a transition state by confirming that the movement of atoms associated with the transition (the negative eigenvalue/frequency) corresponds to the specific transition. Imaginary frequencies are found in the output file as negative eigenvalues.[‡‡‡‡‡] For a transition state only one imaginary frequency/negative eigenvalue is accepted, since it is a first order saddle point. It may sometimes be difficult to say without doubt if the movement of atoms associated with the transition really corresponds to the correct transition. When this happens the correctness of a transition state is determined by using the intrinsic reaction coordinate (IRC) method to find the minimum energy path (MEP) from the TS to the local minima on either side of the maximum.

More experienced computational chemists prefer a shorter method to determine transition states. This is achieved by manually constructing a transition state input structure and performing a full unrestricted transition optimisation on it. This method can be slightly modified by fixing the reaction coordinate at an approximate value, which is known from experience. The rest of the molecule is optimised and then all constraints are released to find the real transition state.

---

††††† A SCAN job varies the presetted reaction coordinate moving the structure from reactant to product or *vice versa*.
‡‡‡‡‡ "eigenvalues are set of scalars associated with a linear system of equation e.g. a matrix equation".

4.5 The solvent effects

A"Gas phase"[§§§§§] calculation excludes solvent effects and is a simplified technique of obtaining the information about the molecule under investigation. Once a clear understanding of the behaviour of the molecule is obtained through gas phase calculations, one can include computational methods to mimic the effect of surrounding condensed phase. The theory to approximate the effect of solvents were developed in the late 1980s and are still currently used. If the theory was only restricted to gas phase calculations, it would have been inapplicable to vast tracts of chemistry and biochemistry.

Since most chemical reactions are performed in solution, it is important to consider the effect of the solvent on the behaviour of the system. In some cases solvents are so tightly bound that they can be a fundamental part of the solute. Those systems are modelled explicitly, but there are cases where solvent does not interact with the solute, but it acts as a "bulk medium" which also affect the behaviour of the solute.

There are different methods that are used to introduce solvent effects in the quantum mechanical optimization of molecules in solution viz: The polarisable continuum model (PCM) developed by Tomasi *et al.*[156] and the Onsager self-consistent reaction field (SCRF) model.[157] PCM represent a cavity in which the solute is embedded by a solvent and it also represent a tool to accurately describe the effects of solution on the molecle under investigation.[158] The aim of the continuum solvent models is to acts as a perturbation on the gas-phase behaviour of the system. In the SCRF model the solute-solvent interaction is treated as a perturbation of the Hamiltonian of the isolated molecule. The solute is assumed to occupy a spherical cavity of radius a_0.[159] In this study *Jaguar* at the B3LYP/6-31+G(d) level was first used for optimizations in the gas phase. The calculation was then repeated starting with the optimised "gas phase"structure and the solvent is selected from the program. The radius of the spherical cavity a_0 is automatically included in the optimization. Solvent effects often play an important role in determining equilibrium constants and reaction rates (i.e. activation energies). In an S_N2 mechanism the rate of the reaction can be determined by the type of solvent used, e.g. if a reaction involves an alkyl halide and a neutral nucleophile, increasing the polarity of the solvent will increase the stability of the transition state, because the charge on the transition state will be higher than that of the neutral reactants. While using a non polar aprotic solvent in a charged transition state will destabilise the transition state by increasing its activation energy.

[§§§§§] The term "gas phase" means in *vacuo*.

4.6 Previously reported results and considerations

Previous "gas-phase" *ab initio* studies by the organic research group at UKZN were undertaken for the acetylation of amines and alcohols with anhydrides, using density functional theory (DFT), Restricted Hartree Fock (RHF) and second order Møller-Plesset (MP2) levels of computation. These studies suggested the formation of a precomplex that preorganizes the molecules in a favourable orientation that facilitates the formation of a six-membered ring transition state of acetic anhydride with the reagents. In all these investigations the model neglected the effect of solvent and catalyst for simplicity. The six-membered cyclic transition state was found to be independent of the level of theory and basis set used.[108,109,110] The model was also applied successfully to the regioselective acetylation of the δ-lactam **96** where five isomers **98a-98e** could potentially form.

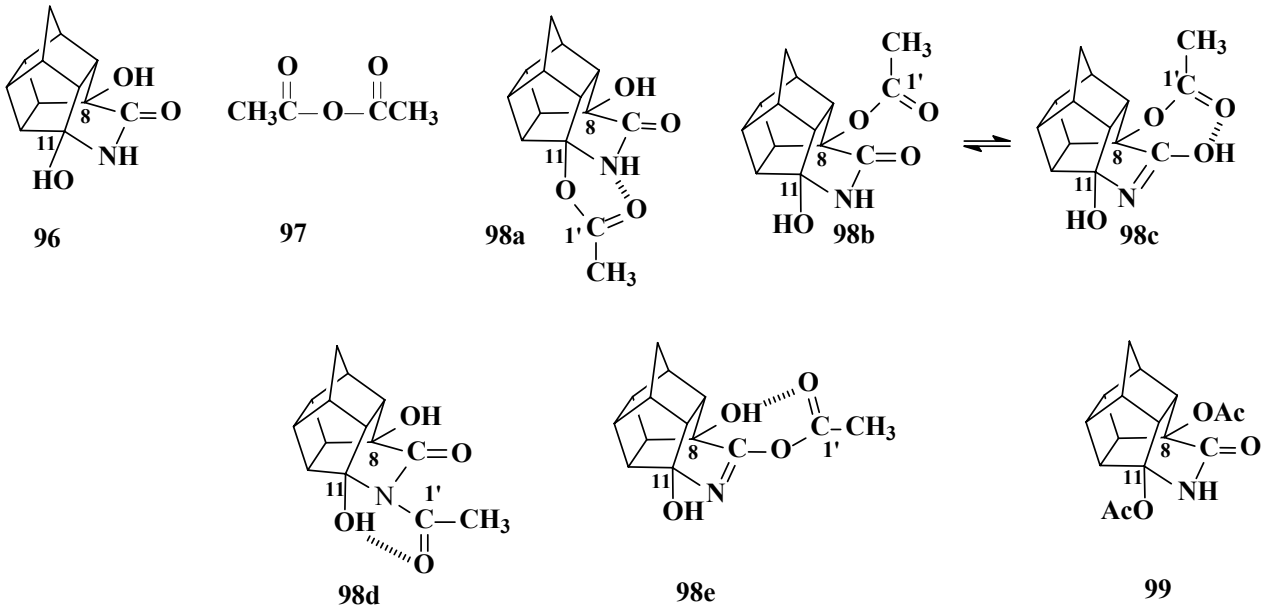
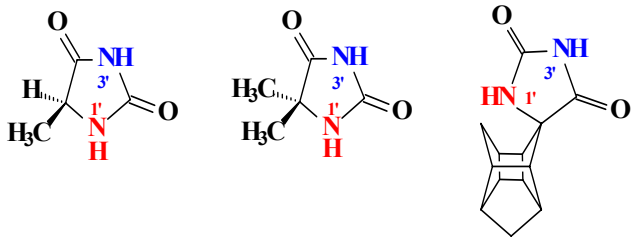


Figure 26: The δ-lactam with possible acetylation products

The computational model has revealed that the transition state reading to **98a** is both kinetically (by 2.5 kcal. mol^{-1}) and thermodynamically preferred product (by 2.8 kcal. mol^{-1}) over the nearest possible isomer. The theoretical predictions correlated with the experimental results since it was experimentally observed that **98a** is obtained between 25-30°C, while a mixture of the monoacetate **98a** and the diacetate **99** is obtained at 38°C.[29,110]

The model was extended to include various hydantoin systems. The same cyclic transition state was obtained for all three hydantoin compounds **100**, **101**, and **102**, both at the N-1' and N-3' positions. Some precomplexes for the PCU lactam systems between hydantoin and acetic anhydride **97**, could not be obtained. Note that the three hydantoin systems **100**, **101**, and **102** exhibit increasing steric hindrance. It was expected that protection would preferentially occur at the

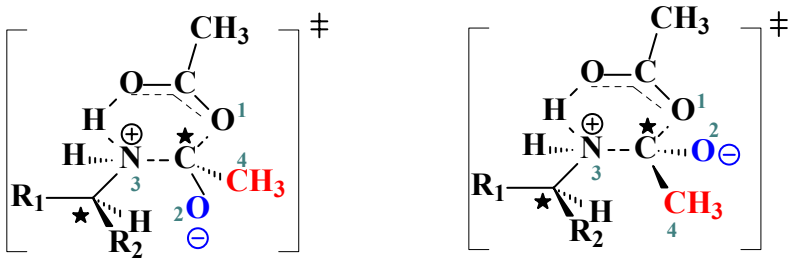
more nucleophilic but more sterically hindered nitrogen (N-1'), since N-3' is less nucleophilic (electron deficient) as it is situated between two carbonyl groups.



100 101 (±)-102

Figure 27: Protection of hydantoins, N-1' versus N-3'

Diastereomeric effects were considered in these calculations as it was shown by Govender *et al.* that chiral alcohols and amines form diastereomeric transition states with anhydrides (Figure 27). It is important to note that both transition states lead to the same product.



[(*R*)-103]-TS1 **[(*S*)-103]-TS2**

Figure 28: Chiral diastereomeric transition states for a chiral amine

According to the computational results protection of 5-methylhydantoin **100** at N-1' of the hydantoin ring was found to be both kinetically (-1.4 kcal.mol^{-1}) and thermodynamically (-0.1 kcal.mol^{-1}) favoured over the closest competing alternative route. The same preference was observed in 5,5-dimethylhydantoin **101**, where protection at N-1' was both kinetically (0.3 kcal.mol^{-1}) and thermodynamically (-23.3 kcal.mol^{-1}) favoured. The justification for these preferences seems logical as it was mentioned above that although N-1' experiences more steric hindrance, it is more nucleophilic than N-3' which favoured protection over N-3'.

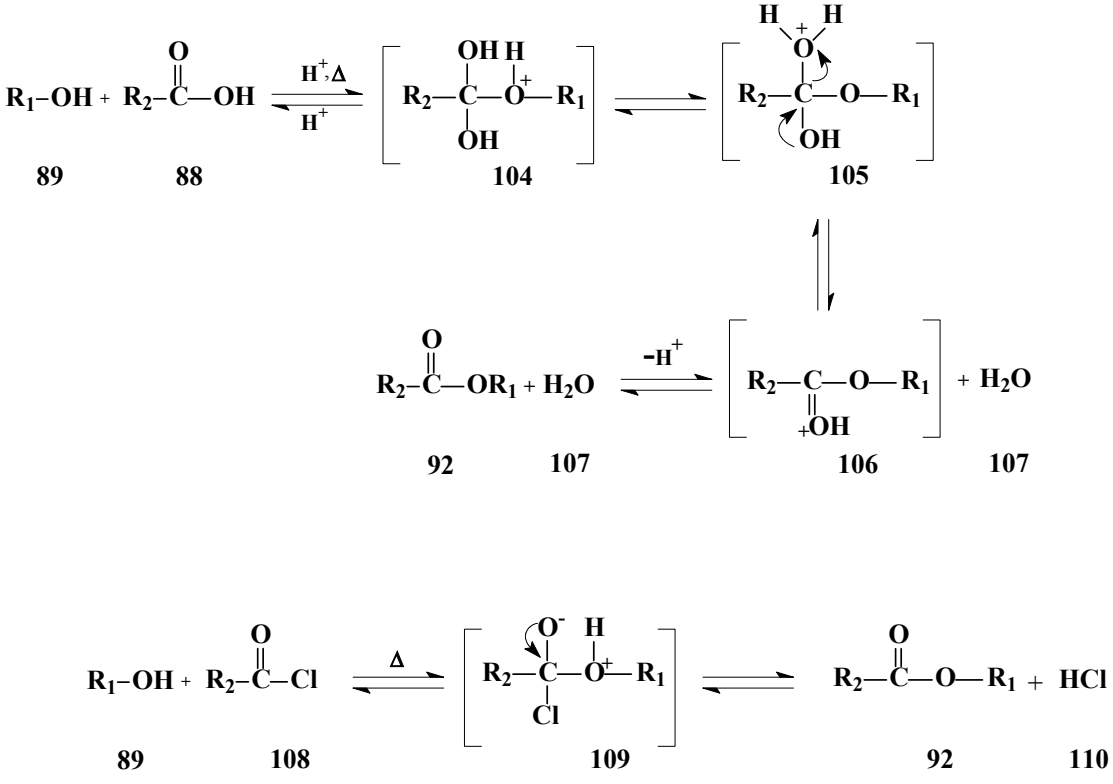
In contrast it was found that protection of the PCU hydantoin **102** at N-3' is both kinetically (2.5 kcal.mol^{-1}) and thermodynamically (2.0 kcal.mol^{-1}) favoured. It is clear that the larger steric hindrance imposed by the rigid PCU skeleton forced protection away from the stronger nucleophilic nitrogen (N-1') towards the less hindered but also much less nucleophilic nitrogen (N-3').

The purpose of this investigation was to investigate the possibility of a similar six membered ring transition state for the esterification reaction of alcohols with acetic acid and acetyl chloride.

4.7 Computational methodology and discussions

Gas-phase studies were undertaken using the *Gaussian* 03 and *Jaguar* programmes at the DFT (B3LYP)[144,145] level of theory with the 6-31+G(d) basis set. The effects of solvent and entropy in the reactions were considered but catalytic effects were ignored for simplicity. The model mimics a normal Fischer esterification[49,160] reaction. The model was further simplified by using methanol, acetic acid and acetyl chloride to reduce the use of computational resources and time.

Modern textbooks[160,161] are teaching slight variations of a linear stepwise mechanism for the esterification reaction of an acid with an alcohol (**Scheme 19**). The carbonyl oxygen of **88** is first protonated by an acid at higher temperatures. The reaction proceeds through the carbocation intermediates **104**, **105** and **106** and the ester **92** is formed after a series of steps. A similar linear step-wise mechanism is accepted for the esterification reaction involving acetyl chloride (**Scheme 19**).[160,161]



Scheme 19: Mechanism of esterification[160,161]

It was previously reported,[108,110] that linear transition states for the protection of alcohols (and amines) with acetic anhydride are not stationary points on the energy profile of a "gas phase" calculation excluding solvent and catalytic effects, therefore this claim had to be verified for the

reaction presented in **Scheme 19**. Intermediate cation structures such as **104**, **105**, **106** and **109** (see **Scheme 19**) are supposed to be minima on the energy surface of the reaction when presented as intermediates in the reaction.

All the carbocations fall apart during an optimization, to its corresponding precursors when optimized in both the "gas phase" as well as when solvent effects are considered (computational results optained in this investigation). This indicates that the conventional mechanism above is problematic as it is accepted that the carbocations are intermediates[******] in the reaction. The computational result is in fact not contrary to what we are taught in organic lectures. It is generally accepted that cations such as **104**, **106** and **109** will automatically lose a proton to the solvent, while **105** will lose water. However, the mechanism as taught in **Scheme 19** implies that those cations should be stationary points in terms of computational chemistry (even a transition state is a stationary point!). This is clearly the case for the gas phase and we know from theory that it is less likely that it will be stationary points in solution, as the solvent will automatically neutralise the excess charge on the substrate by either removing a proton or providing a proton. It is also apparent that these carbocations will have much higher energies than neutral species, that is after all the reason why they would automatically lose charged atoms/fragments to the solution, as it is a thermodynamically driven process. One can therefore rightly state that the mechanism of esterification deserves a fresh theoretical investigation. In order to postulate a six membered ring transition state for this Fischer esterification reaction (**Scheme 19**), the selective use of one solvent molecule (methanol) is required. This is a reasonable assumption as the reactants will be completely surrounded by methanol

There are two types of mechanism proposed in this investigation for the reaction of methanol with acetic acid and acetyl chloride:

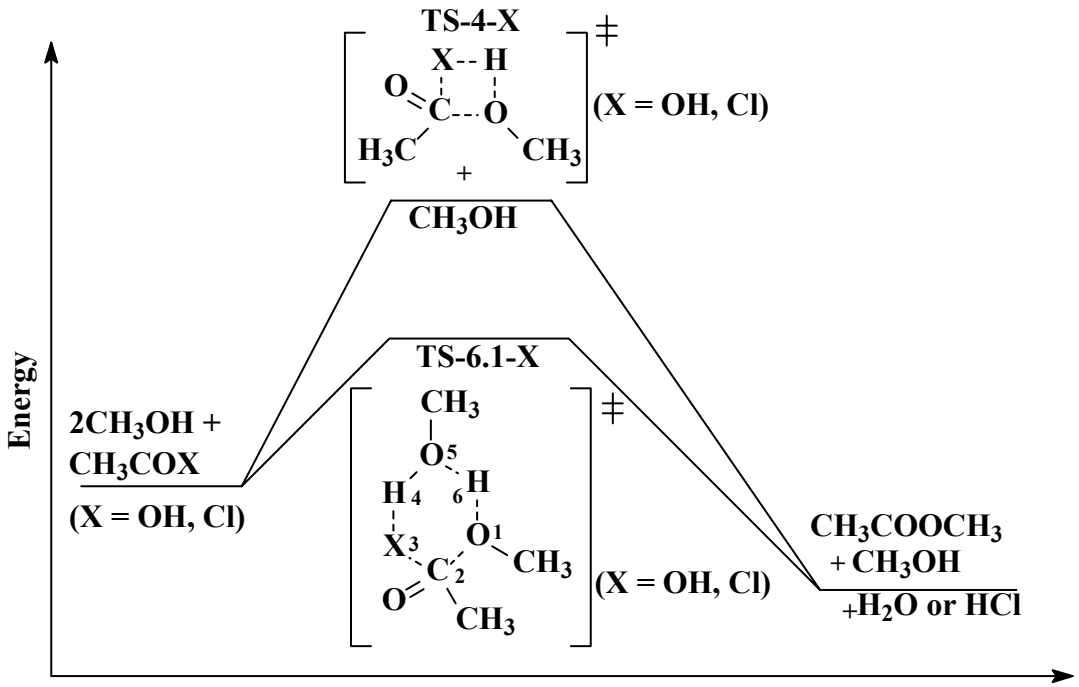
(i) two one step concerted mechanisms which go straight to products i.e. a four-membered ring and six-membered ring transition states (**Scheme 20**)

(ii) A step-wise mechanism which takes place through a stable intermediate (**Scheme 21**).

The discussion below will focus on the acetylation reaction using methanol and acetic acid. The same mechanism can be extrapolated to the esterification of methanol with acetyl chloride. For both the concerted and step-wise mechanisms of ester formation, the first transition state involves simultaneous nucleophilic attack by the alcohol oxygen on the carbonyl carbon (i.e. partial O1----C2 bond formation) and partial transfer of an alcohol hydrogen (H4) to the

[******] An intermediate is a structure that can not necessarily be isolated, but it is stable enough to exist for a short time. An intermediate in terms of computational chemistry should by definition therefore be a stationary point (maxima or minima) on the energy profile.

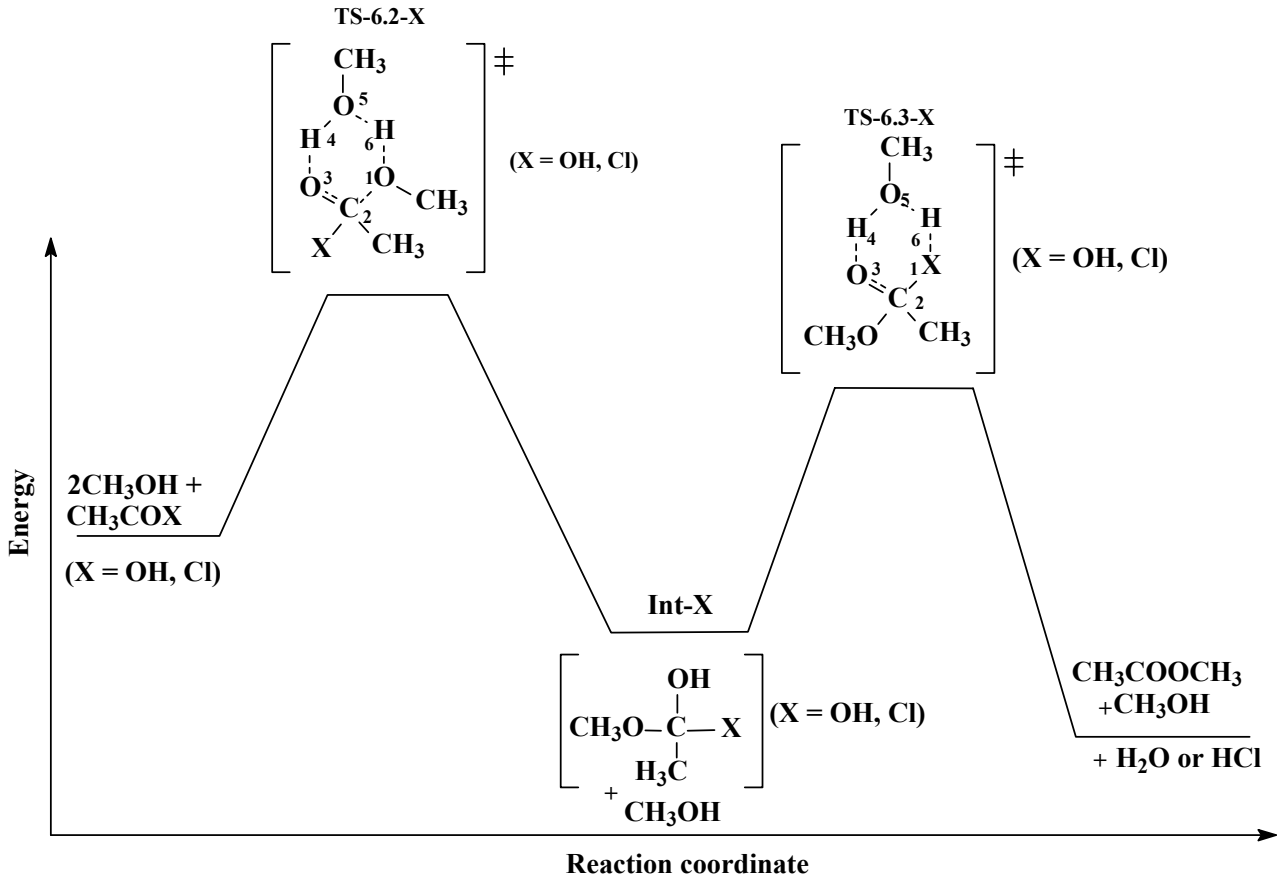
carboxylic acid oxygen atom (i.e. partial X3----H4 bond formation). Simultaneous transfer of H6 from O1 to O5 also occurs.



Scheme 20: Proposed one step concerted mechanism

In the six membered cyclic transition state (**Scheme 20**, **TS-6.1-OH**) both O1----C2 bond formation and the release of water and a methanol molecule will take place simultaneously. A four membered ring transition state (**Scheme 20**, **TS-4-OH**) is also apparent. It is expected that the relative activation energy of the four membered ring transition state (**TS-4-OH**) will be higher that of the six membered ring (**TS-6.1-OH**) due to higher ring strain.

In contrast to the concerted mechanism (**Scheme 20**), the step-wise mechanism (**Scheme 21**) is assumed to take place through a stable intermediate (**Scheme 20**, **Int-OH**). In the first step the transition state (**TS-6.2-OH**), represents the O1----C2 bond formation through the nucleophilic attack of the methanol oxygen atom on the carbonyl carbon atom. Similar hydrogen transfer as above occurs to form the intermediate (**Int-OH**).



Scheme 21: Proposed step-wise mechanism

In the second step, the intermediate (**Int-OH**) is converted to the product through a similar six membered cyclic transition state (**TS-6.3-X**), in which water and methanol are released to form the products. Analogous mechanisms are possible for the acteyl chloride reaction.

Geometry optimizations were performed without any constraints using B3LYP/6-31+G(d) so as to find the lowest energy structures for the reactants, intermediates, transition states and products. Cyclic transition states were characterized by single imaginary frequencies which were associated with the movement of atoms consistent with the expected transition. The transition states seem to belong to a class of pseudopericyclic transition states reported before by Birney *et al.*[162] It is beyond the scope of this study to compare the similarities and difference between pseudopericyclic transition states and the cyclic transition states reported herein. As pointed out before, special attention should be paid to these cyclic transition states (**TS-4-X**, **TS-6.1-X**, **TS-6.2-X** and **TS 6.3-X**) as the carbonyl carbon is chiral during the formation of the transition state. At least two possible transition states exist for each of **TS-4-X** and **TS-6.3-X**. At least four transition states exist for each of **TS-6.1-X** and **TS-6.2-X**. See Figure 29 below.

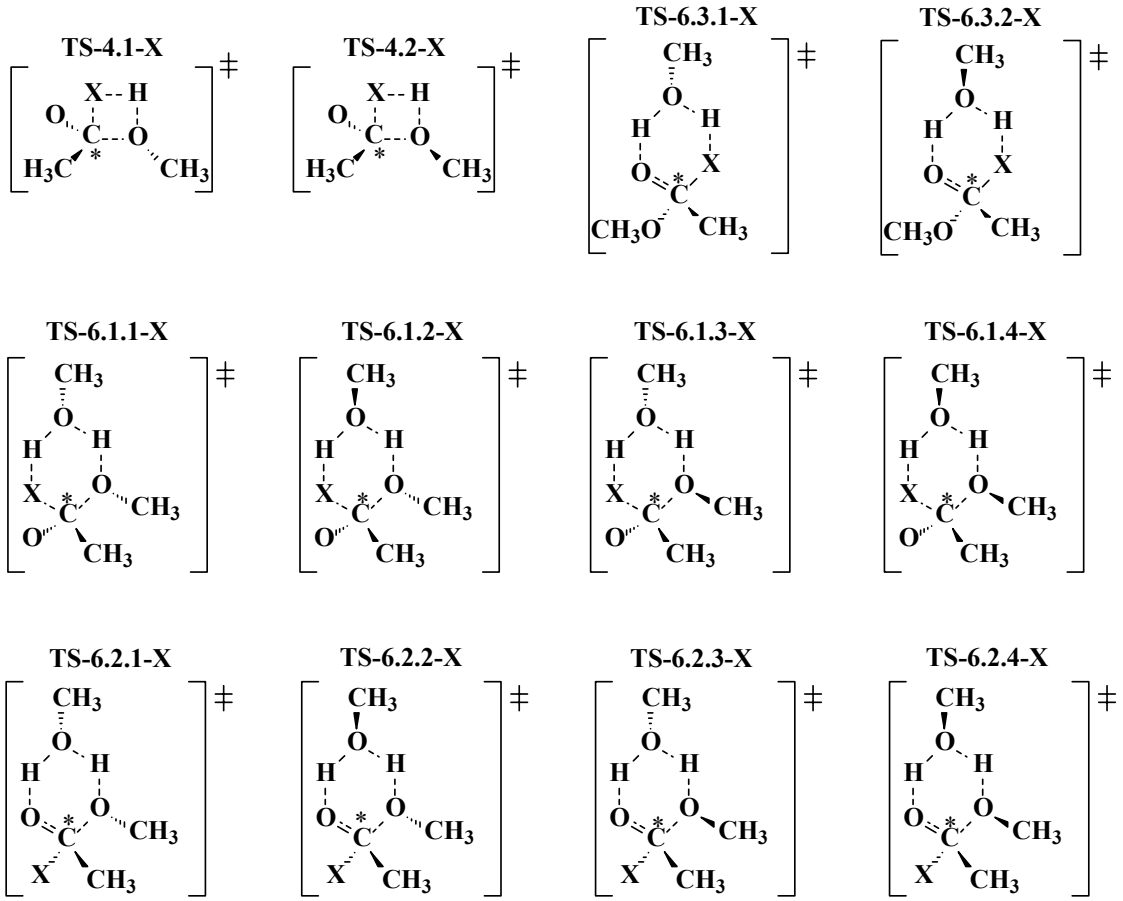


Figure 29: Diastereomeric cyclic transition states (X = OH, Cl)[163]

It is important to note that all the possible diasteriomeric (and enantiomeric) transition states lead to one product (ester) only. The six mechanisms (**Schemes 20**, **21** and Figure 29) were modelled and the results are presented in **Table 1**. The Cartesian coordinates of all optimised structures are available as supplementary material on the attached CD.

Table 1: ΔE vs ΔG in gas phase and in solution

	Relative energies[a,b] using B3LYP/6-31+G(d) / kcal.mol^{-1}							
	X = OH				X = Cl			
	"Gas Phase"		MetOH		"Gas Phase"		MetOH	
	ΔH	ΔG	ΔH	ΔG	ΔH	ΔG	ΔH	ΔG
Reactants[c]	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
TS-4.1-X[d]	41.2	11.7	47.9	11.5	16.0	11.3	12.9	10.6
TS-4.2-X[d]	45.4	11.6	50.4	11.8	21.2	11.2	15.0	10.7
Products[e]	-2.0	1.0	-2.7	2.3	-12.5	1.7	-10.0	2.9
Reactants[c]	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
TS-6.1.1-X[d]	29.3	22.4	36.8	24.2	2.2	22.0	9.2	23.1
TS-6.1.2-X[d]	28.0	22.5	36.9	23.2	2.3	21.9	9.5	23.0
TS-6.1.3-X[d]	-[f]	-[f]	-[f]	-[f]	3.7	22.6	10.5	22.6
TS-6.1.4-X[d]	-[f]	-[f]	-[f]	-[f]	5.8	22.3	10.5	22.6
Products[e]	-2.0	1.0	-2.7	2.3	-12.5	1.7	-9.9	2.9
Reactants[g]	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
TS-6.2.1-X[h]	19.4	23.0	32.4	22.4	-4.7	21.8	0.7	22.4
TS-6.2.2-X[h]	18.8	22.8	32.1	23.2	24.5	22.7	10.5	23.2
TS-6.2.3-X[h]	19.3	23.1	32.2	23.7	-[f]	-[f]	-[f]	-[f]
TS-6.2.4-X[h]	-[f]	-[f]	-[f]	-[f]	16.6	23.3	9.3	22.8
Int-X[h]	5.5	11.9	11.4	12.2	0.3	12.1	4.6	12.1
TS-6.3.1-X[h]	18.7	23.0	32.8	22.4	-13.7	22.5	-14.8	22.0
TS-6.3.2-X[h]	18.3	23.0	32.0	24.3	-10.3	22.6	19.1	23.5
Products[i]	-2.0	1.0	-2.7	2.3	-12.5	1.7	-10.0	2.9

a Energy reported at the same level of theory used for the optimization. The Cartesian coordinates of all optimised structures are available as supplementary material on the attached CD.
b. All energies are reported relative to the sum of energies of the reactants.
c. 2 CH_3OH + CH_3COX (refer to Scheme 20).
d. Refer to Scheme 20.
e. CH_3OH + CH_3COOCH_3 + H_2O or HCl (refer to Scheme 20).
f. One of the methyl groups flipped over during the transition state optimization.
g. 2 CH_3OH + CH_3COX (refer to Scheme 21).
h. Refer to Scheme 21.
i. CH_3OH + CH_3COOCH_3 + H_2O or HCl (refer to Scheme 21).

Four-and six-membered ring cyclic transition states were obtained for methanol, acetic acid and acetyl chloride. The second methanol molecule assisted in facilitating the six membered ring transition states.[108,109,110] According to the results in **Table 1**, values for a change in free energy (ΔG) shows clearly that the four-membered ring transition states (**TS-4-X**) are energetically preferred when compared to the six-membered ring transition states (**TS-6.1-X**), while the opposite is true when entropy effects (ΔS) are ignored.

The fact that four membered rings are more strained increases the activation energy as ΔH considers bond energies, while six membered rings are expected to be less strained and show

smaller ΔH values. As the six membered ring transition states involve a solvent molecule, the disorder is decreased leading to an expected decrease in entropy. It is somewhat surprising that this entropy effect is so large to reverse a much favourable kinetic preference of the six membered ring transition states when entropy is ignored. Since four membered rings are small compared to six membered rings, the former have much higher randomness in terms of translational, vibrational and rotational energies, which increase the entropy of four membered ring system.

The acetyl chloride mechanism is energetically favoured in all respects as is expected. The Gibbs kinetic energy results in **Table 1** shows that the esterification mechanism with acetyl chloride in the gas phase is marginally favoured over acetic acid for both the four membered ring (0.4 kcal.mol^{-1}) and the concerted six membered ring mechanism (0.4 kcal.mol^{-1}).

The six membered ring step-wise mechanism and concerted mechanism have similar Gibbs kinetic energy values for both gas phase and including solvent effects.

Esterification with acetic acid may need mild heating conditions, as it is known that molecules have enough thermal energy at room temperature to allow processes requiring up to ~15 kcal.mol^{-1}.

The reason for the dominance of acetyl chloride mechanism may be because chlorine is a better leaving group than hydroxide (-OH), and is also a weaker base. The intermediate (**Int-OH**) for acetic acid mechanism is preferred by 0.3 kcal/mol over the acetyl chloride intermediate. Similarly the products for the acetic acid reactions are more stable than for acetyl chloride.

The esterification experiment is normally performed in a polar protic solvent (methanol). Since the concerted mechanism model of the Fisher esterification (see **Scheme 20**) represents a normal S_N2 mechanism, the same stabilization effect by a polar solvent is expected. This stabilization is observed for the acteyl chlorie reaction (see **Table 1**), but not for the acetic acid reaction.

Since in theory, the entropy change in gas phase ($\Delta S >> 0$) is higher than the entropy change in liquid phase (solution), i.e. the randomness of the gas phase system is higher than that in a liquid system, therefore the free energy in a gas phase is expected to be more negative implying the spontaneousity of the reaction in gas phase than in solution. The solvent effect plus entropic effect can cause the reaction performed in solution to have a higher free energy than in the gas phase,[164] since there is a decrease in number of participating species (the molecules are trying to be ordered). This is likely to cause a decrease in entropy ($\Delta S << 0$).[165] This may result in a greater value of

ΔG.[††††††] The effect of a decrease in entropy is observed in the six membered ring reactions only (see **Table 1**).

†††††† If $\Delta S << 0$, then $\Delta G = \Delta H - (-)T\Delta S$ and is greater than zero.

Chapter 5

Conclusion

The first part of the thesis involves the synthesis of cage amino acids and the incorporation into peptides. The synthesis of a trishomocubane amino acid **14** and its Fmoc derivative was successfully performed. The Fmoc-tris amino acid fluoride **78** was also synthesised. Acid fluorides are crucial for the incorporation of sterically hindered α,α-amino acid derivatives into short peptides. An excess of **78** and prolonged reaction times are necessary for the successful coupling of the cage amino acid to peptides. The synthesis of the tri-peptide (Ala-Ala-Ala, **69**) was successful and it has been useful in obtaining experience in the SPPS method. Ala-Ala-Ala-tris **79** was successfully synthesized in 40% yield according to HPLC analysis. The successful synthesis of the tetra-peptide encouraged the incorporation of the cage moiety into the centre of the hepta-peptide (Ala-Ala-Ala-tris-Ala-Ala-Ala, **80**). The mass spectroscopy results confirmed the successful synthesis of the hepta-peptide with a peak at m/z = 631which corresponded to the molecular mass of the peptide **80** (630.74 $g.mol^{-1}$). Seven different carbonyl peaks were observed in the ^{13}C NMR spectrum of **80** at 155.36, 170.93, 171.38, 171.54, 171.59, 174.16 and 174.18 ppm. The experience gained in this study was communicated to our collaborators in Sweden and a sample of the amino acid **14** was also sent to them. They were able to synthesise the peptide **80** in pure form recently and separated both diasteriomeric peptides using preparative HPLC. The X-ray structure of each peptide was also obtained and correspond well with the previously predicted structures.

The second part of the thesis involved a theoretical investigation of esterification reactions. Esterification mechanisms were proposed for acetyl chloride, acetic acid and methanol. A concerted four membered ring and six membered ring mechanism was considered. The use of an additional methanol molecule from the solvent facilitated the formation of the six membered ring transition states. A step-wise reaction mechanism involving six membered ring transition states was also investigated. Calculations were performed in both the "gas" phase and in solvent. As expected, the four membered ring transition states were higher in energy, although the disadvantage is overcome when entropic effects are considered. The six membered ring transition states experience a huge entropy penalty, making the four membered ring mechanism favoured in terms of Gibbs energy. The activation energy of the reaction (~11 $kcal.mol^{-1}$) is just inside the range of a spontaneous reaction at room temperature, provided that the by-products are removed. The proposed mechanism does not include any carbocation intermediates. The computational results suggest that the carbocation intermediates which are still proposed in modern text books are not stationary points on the energy profile of the reaction even when solvent effects are considered.

Chapter 6

Experimental

Melting points were recorded using a Bombay 400 013 instrument from Shital Scientific Industries. All melting points are uncorrected. Infrared spectra (IR) were obtained using a single beam Nicolet Impact 410 spectrometer. The one-dimensional NMR spectra were recorded on a Varian Gemini 300 MHz spectrometer. The fast atom bombardment (FAB) mass spectra were obtained from a Micromass VG70-70E mass spectrometer at North-West University (Potchefstroom campus), equipped with an In-tech FAB gun. FAB mass spectra were obtained by bombardment of samples with xenon atoms (1 mA at 8 keV). *m*-Nitrobenzyl alcohol was used as the matrix. Elemental analysis was done using a Leco CHNS 932 instrument. HPLC analysis was performed using a Waters model 510 solvent delivery system and SP8490 detector. A Nucleosil 100 C18 column (5 μm particle size, 250 mm x 4.6mm i.d.) was used. The details for the HPLC solvent conditions are provided below at the relevant section. All the experimental work was performed by the author in the laboratory.

6.1 Synthesis of 5,8-methano-4a,5,8,8a-tetrahydro-1,4-naphthoquinone (18, adduct)

p-Benzoquinone **16** (200 g, 0.54 mol) was dissolved in dry toluene (4 l) and stirred in an ice/salt bath within a dark fume hood. Cold, freshly cracked cyclopentadiene **17** (132 ml, 1.96 mol) was slowly added over two hours via a dropping funnel. The slow addition and cool temperatures allowed for a successful Diels-Alder reaction, without dimerisation of the cyclopentadiene or diadduct formation. The reaction was stirred overnight. The solution was evaporated in a dark fumehood using an evaporating dish, and greenish-yellow crystals of an adduct were obtained. The product was recrystallized from petroleum ether (40-60°C) to yield **18** as yellow crystals. (230 g, 74%, Mp. 77°C). NMR and IR were identical to those of an authentic sample.

6.2 Synthesis of pentacyclo-[5.4.0.0^{2,6}.0^{3,10}.0^{5,9}]-undecane-8,11-dione (19)[33,69]

The adduct **18** (200 g, 1.15 mol) was added to a volumetric flask and dissolved in a (5 l) solution of 10 % (v/v) acetone in hexane. The volumetric flask was exposed to direct sunlight until a colourless solution was obtained. The photocyclisation normally takes three to five days in Durban. The absence of the starting meterial was confirmed by TLC (4:96) of hexane in ethyl acetate. The solvent was evaporated in *vacuo* to give a white, microcrystalline solid of **19** (195 g, 98%, Mp. 240°C). NMR and IR were identical to those of an authentic sample.

6.3 Synthesis of pentacyclo-[5.4.0.0^{2,6}.0^{3,10}.0^{5,9}]-undecane-8,11-dione-mono-ethylene ketal (28)

A mixture of the diketone **19** (183.00 g, 1.05 mol), ethylene glycol (81.20 ml, 1.45 mol), *p*-toluenesulfonic acid (6.11 g, 3.21x10^{-2} mol) and 4Å freshly regenerated molecular sieves (2 g) in dry toluene (800 ml) was gently refluxed with stirring in a Dean and Stark trap for four days. The reaction mixture was left to cool and poured slowly into ice-cold 10 % (m/v) aqueous sodium carbonate (1 l). The aqueous mixture was extracted with dichloromethane (3 x 500 ml). The combined organic extracts were dried over anhydrous sodium sulphate. The mixture was filtered and the solvent evaporated in *vacuo*. The resulting brown residue was recrystallised from n-hexane to give **28** as a white solid (170 g, 74 %, Mp. 73ºC). NMR and IR were identical to those of an authentic sample.[31,69]

6.4 Synthesis of 11-hydroxypentacyclo-[5.4.0.0^{2,6}.0^{3,10}.0^{5,9}]-undecane-8-one-ethylene ketal (29)[31,64]

The mono-ketal **28** (10.00 g, 4.59x10^{-2} mol) was dissolved in ethanol (100 ml) and placed in an ice-bath to cool. An ethanolic solution of sodium borohydride (3.50 g in 50 ml ethanol) was added, with stirring, over a 10 minute period. The mixture was left to stir for two hours in an ice-bath and an additional two hours at room temperature. A small amount (1 ml) of the mixture was withdrawn from the flask and extracted with dichloromethane to confirm complete reduction of the keto functional group using IR spectroscopy. The solvent was removed in *vacuo*, and the residue extracted using dichloromethane (50 ml). A white solid **29** was obtained in good yield (7.80 g, 78%, Mp. 250°C). NMR and IR were identical to those of an authentic sample.

6.5 Synthesis of 11-hydroxypentacyclo-[5.4.0.0^{2,6}.0^{3,10}.0^{5,9}]-undecane-8-one (30)

The hydroxy-ketal **29** (92.5 g, 0.42 mol, 27) was placed in an ice-bath at which stage 10 % (v/v) hydrochloric acid (200 ml) was cautiously added. The mixture was stirred at room temperature for 18 hours. The product was extracted with dichloromethane (150 ml), the organic layer was dried with anhydrous sodium sulphate and evaporated in vacuo to yield **30** in good yield (89.5 g, 96%, Mp. 255°C). NMR and IR were identical to those of an authentic sample.

6.6 Synthesis of *endo*-pentacyclo-[5.4.0.0^{2,6}.0^{3,10}.0^{5,9}]-undecane-8-ol (31)[64,65]

A mixture of the hydroxyketone **30** (13.00 g, 7.52x10^{-2} mol) and hydrazine hydrate (23.10 ml, 0.74 mol) in 1,2 ethane diol (200 ml) was refluxed, with stirring, at 120°C for two hours. The mixture was allowed to cool to 80°C at which stage excess KOH (12.00 g, 0.21 mol) was added cautiously. Excess hydrazine hydrate and water was distilled from the mixture until the temperature reached 185°C. The mixture was refluxed for a further three hours at 185°C. The mixture was cooled,

diluted with water (300 ml) and extracted with dichloromethane (3x100 ml). The combined organic extracts were evaporated in vacuo. A yellow residue was steam distilled to yield the alcohol **31** as a white solid with a distinct camphor-like smell (11.2 g, 56.3%, Mp. 231°C). NMR and IR data were identical to those of an authentic sample.

6.7 Synthesis of pentacyclo-[6.3.0.0^{2,6}.0^{3,10}.0^{5,9}]-undecane-4-ol (27)[31,64,69]

The alcohol **31** (31.2 g, 0.195 mol) was dissolved in glacial acetic acid (360 ml) containing (9 g) of concentrated sulphuric acid and the mixture was refluxed for 6 hours. (The mixture turned black in colour after few minutes of refluxing). The reaction mixture was cooled, diluted with 1500 ml of water and extracted with dichloromethane. The organic extract was washed successively with water and sodium bicarbonate solution. Decolourization was achieved by heating an organic extract with activated charcoal and subsequent filtration. Evaporation of the solvent in vacuo resulted in pentacyclo-[6.3.0.0^{2,6}.0^{3,10}.0^{5,9}]-undecane-4-acetate (orange oil, **32**). The orange oil **32** was stirred overnight in a mixture of methanol (330 ml) and K_2CO_3 (18 g). The volume of the methanol mixture was reduced in *vacuo*, diluted with water and extracted with dichloromethane. The organic extract was dried and evaporated in *vacuo* to afford a white solid alcohol **27** in good yield (20.6 g, 66%, Mp.167 -168^{o}C). NMR and IR were identical to those of an authentic sample.[31,69]

6.8 Synthesis of pentacyclo-[6.3.0.0^{2,6}.0^{3,10}.0^{5,9}]-undecane-4-one (12)

Chromium trioxide (14.00 g, 0.14 mol) was dissolved in deionised water (30 ml) and added to acetic acid (320 ml). The alcohol **27** (10.00 g, 0.0625 mol) was dissolved in acetic acid (100 ml) and added drop-wise to the prepared mixture. The reaction was refluxed at 90ºC overnight. The successful oxidation (Jones oxidation) of the alcohol was evident through reduction of Cr^{6+} (dark red) to Cr^{3+} (green). The reaction mixture was cooled, diluted with deionised water (1000 ml) and extracted with dichloromethane (800 ml). The organic extract was successively washed with water (2 x 500 ml), a saturated bicarbonate solution (2 x 500 ml) and water (500 ml). The organic solution was dried, using anhydrous sodium sulphate. The mixture was filtered, and the solvent removed in *vacuo* to yield pentacyclo-[6.3.0.0^{2,6}.0^{3,10}.0^{5,9}]-undecanone **12** (7.5 g, 76%, Mp. 195°C). NMR and IR were identical to those of an authentic sample.[30,69]

6.9 Synthesis of Tris-hydantoin (13)

A mixture of NaCN (1.00 g, 2.04x10^{-2} mol), trishomocubanone **12** (1.00 g, 6.25x10^{-3} mol) and $(NH_4)_2CO_3$ (2.00 g, 2.08x10^{-2} mol), ethanol (10 ml) and NH_4OH (15 ml) were first stirred and sealed in a glass medium-pressure vessel. The reaction vessel was sealed in a metal pressure vessel and partially filled with water. The reaction vessel was placed in an oil bath and heated to 60 -100 ºC overnight. The cooled reaction was first diluted with deionised water (100 ml) and

extracted with ethyl acetate (150 ml). The solvent was removed in *vacuo* to afford the crude hydantion (**13**). The product was washed successively with acetone and diethyl ether and thereafter, recrystallised from tetrahydrofuran (THF) to yield pure hydantoin **13** as a white solid (1.4 g, 98% yield, Mp. 230°C). NMR and IR were identical to those of an authentic sample.[30,69]

6.10 Synthesis of bis-t-Boc protected Tris-hydantoin (45)[30,46]

The method was adopted from the literature.[73,74] A solution of tris-hydantoin **13** (0.50 g, 2.17×10^{-3} mol), di-tert-butyl-dicarbonate (1.19 g, 5.43×10^{-3} mol), 4-dimethyl-aminopyridine (DMAP) (1.3 mg, 1.09×10^{-4} mol) and triethylamine (0.35 ml, 2.60×10^{-3} mol) in dry THF (50 ml) was left to stir under nitrogen gas for 24 hours. The solution was concentrated in *vacuo* to yield the crude product. Purification was achieved through silica gel column chromatography (dichloromethane) to yield the product as a white powder **45** (0.77 g, 98%, Mp. 227 °C). NMR and IR data were identical to an authentic sample.[30,69]

6.11 Synthesis of Trishomocubane amino acid (14)

The method for the synthesis of trishomocubane amino acid was adopted from the synthesis of adamantanine (**Scheme 4**). The method required the synthesis of hydantoin **13** from trishomocubanone **12** and hydrolysis of the hydantoin at high temperatures to the amino acid **14**. The poor yield obtained with this method challenged the research group to search for another method that could produce **14** in good yield. The solution was the bis-t-Boc protection of the precursor (hydantoin-**13**) by di-tert-butyl dicarbonate to form bis-t-Boc hydantoin **45**. Hydrolysis of the bis-t-Boc hydantoin was found to produce a good yield of trishomocubane amino acid (> 95 % yield). Both methods are presented below.

6.11.1 Synthesis of Tris-amino acid (14) from the Tris-hydantoin (13)

Trishomocubane hydantoin **13** (0.1 g, 4.88 x 10-4 mol) was added to a 100 ml beaker, 50 ml of 10% NaOH solution was added. The mixture was then transferred to a metal pressure vessel which was heated with stirring at 170°C for 12 hours. The yellow solution was cooled, acidified with concentrated hydrochloric acid to ~ pH 6.5. The white precipitate of **14** formed. The volume of the solution was then reduced to 20 ml (on a steam bath) to precipitate more amino acid. The white product (0.06g, 60 %) of **14** was obtained, filtered and washed with diethyl ether, acetone and dried in a vacuum desiccator.

The ninhydrin test was positive and confirmed the presence of the free $-NH_2$ group. IR (KBr) ν_{max}, 3436 cm^{-1}, 3198 cm^{-1}, 2953 cm^{-1}, 2873 cm^{-1}, 1623 cm^{-1}, 1591 cm^{-1} and 1494 cm^{-1}.

6.11.2 Synthesis of Tris-amino acid (14) from a bis-t-Boc Tris-hydantoin (45)

The procedure was adopted from Rebek *et al*. The bis-t-Boc hydantoin **45** in aqueous LiOH (8 mol equivalents, 1.560M) was stirred in a round bottom flask at room temperature overnight. The resulting solution was adjusted to ~pH 6.5 using concentrated HCl. The trishomocubane amino acid **14** precipitated from solution and was filtered off to remove the water and by-products. The product (**14**, > 95%) was washed with acetone and diethyl ether. The ninhydrin test was positive. IR data was identical to an authentic sample.[30,69]

6.12 Synthesis of Fmoc-Tris amino acid (50)[31,69]

To an ice cooled solution of amino acid **14** (0.10 g, 4.88×10^{-4} mol) in 1,4-dioxane (7.50 ml) and 10% (m/v) Na_2CO_3 (20 ml) was added, with stirring, a solution of 9-fluorenylmethylchloroformate (Fmoc-Cl, 015 g, 5.80×10^{-4} mol) in 1,4-dioxane (5.00 ml). This was left to stir in the ice-bath for four hours and at room temperature for a further eight hours. The solution was diluted with deionised water (50 ml) and extracted with diethyl ether (100 ml). The aqueous layer was cooled to 10°C in an ice-bath and acidified to pH 2 with concentrated HCl. The solution was subsequently extracted with ethyl acetate (150 ml). The organic solvent was evaporated in *vacuo* to give the crude Fmoc amino acid derivative **50**. Purification was achieved using a silica gel column chromatography (dichloromethane) and (0.16g, 75%, Mp. 213°C) of the product was obtained as white solid. . Elemental analysis, Found: %C=75.83, %H=5.95, %N=2.92; $C_{27}H_{25}NO_4$ (427.178), required %C=75.86, %H=5.89, %N=3.28%. NMR data was identical to an authentic sample.[30,31] IR(KBr) ν_{max}, 3417 cm^{-1}, 3358 cm^{-1}, 2954 cm^{-1}, 2880 cm^{-1}, 1744 cm^{-1}, 1711 cm^{-1}, 1419 cm^{-1}, 737 cm^{-1}. M.S. data m/z=428

6.13 Synthesis of Fmoc Tris-amino acid fluoride (78)[45,69]

Cyanuric fluoride (40.50 µl, 4.68×10^{-4} mol) was added to a suspension of Fmoc tris amino acid **50** (0.20 g, 4.68×10^{-4} mol) and dry dichloromethane (50 ml). Dry pyridine was added (37.80 µl, 4.68×10^{-4} mol) and the resulting solution was left to stir under nitrogen for 12 hours. Precipitated water-soluble cyanuric acid was extracted with ice water (40 ml). The organic phase was dried over anhydrous $MgSO_4$, filtered and removed in *vacuo*. The resulting white solid was recrystallised from dichloromethane/hexane to give the pure acid fluoride as a white solid (0.17 g, 87 %). The NMR data was identical to those of an authentic sample. IR (KBr) ν_{max}, 3414 cm^{-1}, 2971 cm^{-1}, 2869 cm^{-1}, 1841 cm^{-1} (-CO-F), 1729 cm^{-1}, 1614 cm^{-1}, 741 cm^{-1}.

6.14 Coupling of linker to MBHA resin

The MBHA resin **64** (0.50 g, 0.70 mmol NH_2/g) was swelled by adding DMF to a 10 ml polypropylene syringe. The FmocAM linker **65** (0.47 g, 8.75x10-4 mol), DIPCDI **67** (0.14 ml, 8.75x10-4 mol) and HOBt **66** (0.12 g, 8.75x10-4 mol) were added at room temperature to the resin in DMF. This was left to stir for 24 hours. Successful attachment was confirmed by a negative ninhydrin test. The Fmoc group was cleaved by stirring the resin and linker in 20 % (v/v) piperidine in DMF for 30 minutes. Cleavage of the Fmoc group was confirmed by a positive ninhydrin test. The Fmoc by-product was removed from the solution by filtration, followed by washing of the resin with isopropanol and DMF.

6.15 Washing procedure for solid phase peptide synthesis

The resin was washed with DMF (3 x 10 ml, 30 sec), isopropanol (3 x 10 ml, 30 sec) and DMF (3 x 10 ml, 30 sec).

6.16 Coupling of Fmoc Alanine to the solid support

Fmoc Alanine, DIPCDI (0.14 ml, 8.75x10-4 mol) and HOBt (0.12 g, 8.75x10-4 mol) were added at room temperature to the polymer support (resin) in DMF. This was left to stir for 24 hours. Successful attachment was confirmed by a negative ninhydrin test. The Fmoc group was cleaved by stirring the mixture in 20% (v/v) piperidine in DMF for 30 minutes. Cleavage of the Fmoc group was confirmed by a positive ninhydrin test. The Fmoc group was removed from the solution by filtration, followed by washing of the resin with isopropanol and DMF.

6.17 Alternative method for coupling of Fmoc Alanines and Fmoc-tris amino acid

The Fmoc Alanine (3 equivalents), DIPCDI (3 equivalents) and HATU (3 equivalents) were added at room temperature to the Seiber amide resin in DMF. This was left to stir for 24 hours. Successful attachment was confirmed by a negative TNBS test. All washings were done with DMF. The Fmoc group was cleaved by stirring the mixture in a 20 % (v/v) piperidine in DMF for 30 minutes. Cleavage of the Fmoc group was confirmed by a positive TNBS test. The Fmoc group was removed from the solution by filtration, followed by washing of the resin with isopropanol and DMF.

6.18 Coupling of Fmoc tris amino acid fluoride (78) to the tri-peptide (69) bound to resin

The Fmoc tris-amino acid fluoride **78** (0.67 g, $1.56x10^{-3}$ mol) and dry pyridine (151.00 µl, $1.87x10^{-3}$ mol) was added to the resin in DMF and left to stir for 24 hours. Successful attachment was confirmed by a negative ninhydrin test. The Fmoc group was cleaved by stirring the mixture in a

20 % (v/v) piperidine in DMF for 30 minutes. The Fmoc group was removed from the solution by filtration, followed by washing of the resin with isopropanol and DMF.

6.19 Cleavage of the tetra-peptide (79) from the resin and linker

The supported peptide was separately washed with dichloromethane (3x10 ml), methanol (3x10 ml) and diethyl ether (3x10 ml) and dried overnight under vacuum. The dry resin was placed in a flask containing the cleavage mixture of 95 % (v/v) TFA, 2.5 % (v/v) water and 2.5 % (v/v) triisopropylsilane. The mixture was left to stir for 12 hours. The resin was removed by filtration and washed several times with TFA. The TFA was removed in *vacuo* at 40°C. Higher temperatures can cleave the peptide bonds.[166] A yellow oil remained. The oil was dissolved in water and frozen with liquid nitrogen. The solid peptide was obtained after freeze drying. IR(KBr) ν_{max} 3394 cm^{-1}, 3298 cm^{-1}, 2972 cm^{-1}, 2880 cm^{-1}, 1756 cm^{-1}, 1677 cm^{-1}, 1623 cm^{-1}, 1548 cm^{-1}, 1203 cm^{-1}. Since the peptide was not completely elucidated by NMR it is not easy to assign peaks, but carbonyl peaks appeared at 174, 172, 172 and 159 ppm.

6.20 General procedure for HPLC analysis

HPLC chromatography of **69**, **79** and **80** was performed using a solvent system which was 80 % (v/v) acetonitrile/20 % (v/v) water/0.1 % (v/v) trifluoroacetic acid (solution A) and 100 % water (solution B). The following gradient elution system was used: 80 % solution A and 20 % solution B was changed linearly to 40 % solution A and 60 % solution B at 1.00 ml per minute over 50 minutes.

References

1 Marchand, A. P. *Chem. Rev*. **1989**, 89, 1011.
2 Griffin, G. W.; Marchand, A. P. *Chem. Rev.* **1989**, 89, 997.
3 Marchand, A. P. *Advances in Theoretically Interesting Molecules*, R. P. Thummel, ed, JAI Press, Greenwhich CT, **1989**, 1, 357.
4 (a) Geldenhuys, W. J.; Malan, S. F.; Bloomquist, J. R.; Marchand, A. P.; Van der Schyf, C. J. *Med. Res. Rev.* **2005**, 25, 21.
5 Marchand, A. P. *Science* **2003**, 229, 52.
6 Marchand, A. P.; Reddy, G. M.; Deshpande, M. N.; Watson, W. H.; Nagl , A.; Lee, O. S.; Osawa, E. *J. Am. Chem. Soc.* **1990**, 112, 3521.
7 Marchand, A. P.; Kruger, H. G.; Power, T. D.; Segal, C. *Kem. Ind*. **2002**, 51, 51.
8 Weinstein, D. I.; Alster, J.; Marchand, A. P. *Thermochimica Acta*. **1986**, 99, 133.
9 Marchand, A. P. *Tetrahedron* **1988**, 44, 2377.
10 Danysz, W.; Parsons, C. G.; Kornnhuber, J.; Schmidt, W. J.; Quack, G. *Neuroscience and Biobehavioral reviews* **1997**, 21(4), 455.
11 Parsons, C. G.; Danysz, W.; Quack, G. *Neuropharmacology* **1999**, 38, 735.
12 Kitagawa, K.; Mizobuchi, N.; Hama, T.; Hibi, T.; Konishi, R.; Futaki, S. *Chem. Pharm. Bull.* **1997**, 45, 1782.
13 Marchand, A. P.; Takhi, M.; Kumar, V. S.; Krishnudu, K.; Ganguly, B. *ARKIVOC, Part 3* **2001**, 13.
14 Govender, T.; Hariprakasha, H. K.; Kruger, H. G.; Marchand, A. P. *Tetrahedron Asymmetry* **2003**, 14, 553.
15 Boyle, G. A.; Govender, T.; Kruger, H. G.; Maguire, G. E. M. *Tetrahedron Asymmetry* **2004**, 15, 2661.
16 Boyle, G. A.; Govender, T.; Kruger, H. G.; Maguire, G. E. M. *Tetrahedron Asymmetry* **2004**, 15, 3775.
17 (a) Oliver, D. W.; Dekker, T. G.; Snyckers, F. O. *Eur. J. Med. Chem*. **1991**, 26, 375.
(b) Oliver, D. W.; Dekker, T. G.; Snyckers, F. O.; Fourie, T. G. *J. Med. Chem*. **1991**, 24, 851.
(c) Oliver, D. W.; Dekker, T. G.; Snyckers, F. O. *Arzneim. Forsch./Drug Res*. **1991**, 41, 549.
(d) Oliver, D. W.; Dekker, T. G.; Snyckers, F. O.; Fourie, T. G. *J. Pharmacol. Exp. Ther*. submitted for publication.
18 Walshe, S. F. *Diseases of the Nervous System*, 11th ed.; E & S Livingstone Edinburgh: London, **1970**.
19 http//: www.ukzn.ac.za/kruger/kruger.htm
20 Bailey, E. V.; Stone, T. W. *Arch. Int. Pharmacodyn*. **1975**, 216, 246.
21 Brookes, K. B.; Hickmott, P. W.; Jutle, K. K.; Schreyer, C. A. *S. Afr. J. Chem*. **1992**, 45, 8.
22 (a) Engler, E. M.; Andose, J. D.; Schleyer, P. V. R.; Osawa, E.; Kent, G. J. *J. Am. Chem. Soc*. **1973**, 95, 8005.
(b) Godleski, S. A.; Schelyer, P. V. R. *J. Chem. Soc., Chem. Comm*. **1974**, 976.
23 Kent, G. J.; Godleski, S. A.; Osawa, E.; Schyler, P. V. R. *J. Org. Chem*. **1977**, 42, 3852.
24 Underwood, G. R.; Ramamoorthy, B. *Tetrahedron Lett.* **1970**, 4125.
25 Aitken, R. A.; Kinlenyi, S. N. *Asymmetric synthesis*, Blackie Academic & Professional, **1992**.
26 Atkinson, R. S. *Stereoselective Synthesis*, John Willey & Son Ltd, **1995**.
27 Sessler, J. L.; Andrievsky, A. *Eur. J. Chem*. **1998**, 4, 159.
28 Fujii, N.; Saito, T. *The Chemical Records*. **2004**, 4, 267.
29 Martins, F. J. C.; Viljoen, A. M.; Kruger, H. G.; Fourie, L.; Roscher, J.; Joubert, A.; Wessels, L. P. *Tetrahedron* **2001**, 57, 1601.

30 Fourie, L.; Govender, T.; Hariprakasha, H. K.; Kruger, H. G.; Raasch, T. *Magn. Reson. Chem*. **2004**, 42, 617.
31 Govender, T. *MSc Dissertation,* University of Natal, **2001**.
32 Marchand, A. P.; Power, T. D.; Kruger, H. G. *Croatica Chemica Acta*. **2001**, 74(2), 265.
33 Cookson, R. C.; Crundwell, E.; Hill, R. R.; Hudee, J. *J. Chem. Soc*. **1964**, 3062.
34 Marchand, A. P.; Allen, R. W. *J. Org. Chem*. **1974**, 39(11), 1596.
35 Bisetty, K.; Gomez-Catalan, J.; Aleman, C.; Giralt, E.; Kruger, H. G.; Perez, J. J. *J. Peptide Sci*. **2004**, 10, 274.
36 Bisetty, K.; Corcho, F.; Canto, J.; Kruger, H. G.; Juanje, J. J. *J. Peptide Sci.* **2006**, 12, 92.
37 Bisetty, K.; Catalan, J. G.; Kruger, H. G.; Juanje, J. J. *J. Mol. Struct. (Theochem)* **2005**, 731, 127.
38 Bisetty, K.; Govender, P.; Kruger, H. G. *Biopolymers* submitted for publication.
39 Govender, P. P. *M. Tech. Dissertation*, Durban Institute of Technology, **2004**.
40 Bisetty, K.; Corcho, F. J.; Canto, J.; Kruger, H. G.; Perez, J. J. *J. Mol. Struct. (Theochem)* submitted for publication.
41 Sibanda, B. L.; Thornton, J. M. *Nature* **1985**, 316, 170.
42 Sewald, N.; Jakubke, H. D. *Peptides: Chemistry and Biology*, Wiley-VCH Verlag GmbH & Co.KgaA, **2002**.
43 http://imtech.res.in/raghava/betapred/intro.html
http://www.sanger.ac.uk/Users/sgj/thesis/html/node39.html
44 Singh, A. *Msc Dissertation,*University of KwaZulu-Natal, **2005**.
45 Govender, T.; Hariprakasha, H. K.; Kruger, H. G.; Raasch, T. *S. Afr. J. Chem*. **2005**, 58, 37.
46 Kruger, H. G. *PhD. Thesis*, Potchefstroom University, **1996**.
47 Edelman, J.; Chapman, J. M. *Basic Biochemistry*, Morrison & Gibb Ltd: London, **1978**.
48 Jakubke, H. D.; Jeschkeit, H. *An introduction to Amino acids, peptides and proteins*., The Macmillan Press Ltd, **1977**.
49 Hart, H.; Craine, L. E.; Hart, D. J. *Organic chemistry,* 10th ed.; Houghton Mifflin Company, **1999**.
50 McMurry, J.; Fay, R. C. *Chemistry,* 2nd ed.; Printince Hall, **1998**.
51 Fiechter, A. *Advances in Biochemical engineering/biotechnology*, Springer-Verlag Berlin heidelberg, **1990**.
52 Lee, K. H. *Current Pharmaceutical Design* **2002**, 8(9), 795.
53 Maitland Jones,Jr. *Organic Chemistry*, W. W. Northon & Company New York, **1997**.
54 Dembitsky, M. V.; Srebnik, M. *Tetrahedron* **2003**, 59, 579.
55 Van Betsbrugge, J.; Van De Nest, W.; Verheyden, P.; Tourwe, D. *Tetrahedron* **1998**, 54, 1753.
56 Adamczk, M.; Akireddy, S. R.; Reddy, R. E. *Tetrahedron* **2002**, 58, 6951.
57 Nagasawa, H. T.; Elberling, J. A.; Shirota, F. N. *J. Med. Chem*. **1973**, 16(7), 823.
58 Miller, B. *Advanced Organic Chemistry*, Prentice Hall, **1998**.
59 Allinger, N. L.; Tribble, M. T.; Miller, M. A.; Wetrz, D. H. *J. Am. Chem. Soc*. **1971**, 93, 1637.
60 Eaton, P. E.; Hudson, R. A.; Giordano, C. *J. Chem. Soc., Chem. Comm*. **1974**, 978.
61 Cahn, R. S.; Ingold, C. K.; Prelong, V. *Angew. Chem., Int. Ed. Engl*. **1966**, 5, 385.
62 Nakazaki, M.; Naemura, K.; Arashiba, N. *J. Org. Chem.* **1978**, 43, 689.
63 Mueller, H.; Melder, J. P.; Fessner, W. D.; Hunkler, D.; Fritz, H.; Prinzbach, H. *Angew. Chem*. **1988**, 100(8), 1140.
64 Dekker, T. G.; Oliver, D. W. *S. Afr. J. Chem*. **1979**, 32, 45.
65 Eaton, P. E.; Cassar, L.; Hudson, R. A.; Hwang, D. R. *J. Org. Chem*. **1976**, 41(8), 1445.
66 Sasaki, T.; Eguchi, S.; Kriyama, T.; Hiroaki, O. *Tetrahedron* **1974**, 30, 2707.
67 Mehta, G.; Srikrishana, A.; Reddy, A. V.; Nair, M. S. *Tetrahedron* **1981**, 37(25), 4543.
68 Smith, E. C.; Barborak, J. C. *J. Org. Chem*. **1976,** 41, 1433.

69 Raasch, T. *MSc Dissertation*, University of Natal, **2003**.
70 http://www.emsl.pnl.gov/forms/basisform.html
71 Bucherer, H. T.; Steiner,W. *J. Prakt. Chem*., **1934**, 140, 291.
72 Munday, L. *J. Chem. Soc.* **1961**, 4372.
73 Kubik, S.; Meissner, R. S.; Rebek Jr, J. *Tetrahedron Lett*. **1994**, 35, 6635.
74 Flynn, D. A.; Zelle, R. E.; Grieco, P. A. *J. Org. Chem.* **1983**, 48, 2424.
75 Kaiser, E. *Anal. Biochem*. **1970**, 34, 595.
76 Rogers, L. J. *Amino acids, Proteins and Nucleic acids*, Academic Press Limited vol.5, **1991**.
77 Dekker, T. G.; Oliver, D. W.; Pachler, K. G. R.; Wessels, P. L.; Woudenberg, M. *Org. Magn. Reson*. **1981**, 15, 188.
78 Dekker, T. G.; Oliver, D. W.;Wessels, P. L. *Magn. Reson. Chem*. **1994**, 32(6), 330.
79 Schwartz, M.; Marchand, A. P.; Wang, K. S.; Reddy, S. P.; Redda, G. M.; Gadgil, V. R.; Watson, W. H.; Kashyap, R. P.; Kraweic, M. *J. Chem. Soc., Perkin Trans.* 2 **1993**, 1829.
80 Offord, R. E. *Semisynthetic proteins*, A Wiley-Interscience publication, **1980**.
81 Jensen, J. H.; Baldridge, K, K.; Gordon, M. S. *J. Phys. Chem*. **1992**, 96, 8340.
82 Remko, M.; Rode, B. M. *Phys. Chem. Chem. Phys*. **2001**, 3, 4667.
83 Remko, M.; Rode, B. M. *Structural Chemistry* **2004**, 15(3), 223.
84 Davidson, J. P.; Lubman, O.; Rose, T.; Waksman, G.; Martin, S. F. *J. Am. Chem. Soc.* **2002**, 124(2), 205.
85 Notes from BioQuadrant Pharmaceutical Intermediate, Inc. The role of peptides, peptidomimetics and amino acids in future drug discovery.
86 du Vigneaud, V.; Ressler, C.; Swan, J. M.; Roberts, C. W.; Kattsoyannis, P. G.; Gordon, S. *J. Am. Chem. Soc*. **1953**, 75, 4879.
87 http:// www.oxytocin.org/oxytoc/, http://en.wikinepia.org/wiki/oxytocin
88 King, J. *Protein and Nucleic Acid Structure and Dynamics*, the Benjamin/ Commings Publishing Company, **1985**.
89 Andrews, M. J. L.; Tabor, A. B. *Tetrahedron* **1999**, 55, 1711.
90 Merrifield, R. B. *J. Am. Chem. Soc.* **1963**, 85, 2149.
91 Novabiochem **1995**, S2.
92 Alberico, F.; Kneib-Cordonier, N.; Biancalana, S.; Gera, L.; Masada, R. I.; Hudson, D.; Barany, G. *J. Org. Chem.* **1990**, 55, 3730.
93 Carpino, L. A.; EI-Faham, A.; Minor, C. A.; Albericio, F. *J. Chem. Soc*., *Chem. Commun*., **1994**, 201.
94 Hancock, W. S.; Battersby, J. E. *Anal. Biochem.* **1976**, 71, 261.
95 Navibiochem **1997/98**, S38.
96 Hancock, W. S.; Battersby, J. E. *Anal. Biochem.* **1976**, 71, 261.
97 Vranesic, B.; Tamasic, J.; Smerdel, S.; Kantoci, D.; Beneditti. F. *Helv. Chim. Acta.* **1993**, 76, 1752.
98 Padmanabham, S.; Baldwin, R. L. *J. Mol. Biol*. **1991**, 219, 135.
99 Karle, I. L.; Rao, R. B.; Prasad, S.; Kaul, R.; Balaram, P. *J. Am. Chem. Soc*. **1994**, 116(23), 10355.
100 Peggion, C.; Formaggio, F.; Crisma, M.; Toniolo, C.; Kaptein, B.; Broxterman, Q. B.; Kamphuis, J. *J. Peptide Sci*., **1999**, 5, 547.
101 Miick, S. M.; Martinez, G. V.; Fiori, W. R.; Todd, A. P.; Millhauser, G. L. *Nature* **1992**, 359, 653.
102 Hanson, P.; Martinez, G.; Millhauser, G.; Formaggio, F.; Crisma, M.; Toniolo, C.; Vita, C. *J. Am. Chem. Soc.* **1996**, 118, 271.
103 Mdluli, P. S. *MSc Dissertation,* University of KwaZulu-Natal, **2005.**
104 Wenschuh, H.; Beyermann, M.; Krause, E.; Brudel, M.; Winter, R.; Schumann, M.; Carpino, L. A.; Bienert, M. *J. Org. Chem*. **1994**, 59, 3275.

105 Wenschuhh, H.; Beyermann, M.; Karuse, M.; Carpino, L. A.; Bienert, M. *Tetrahedron Lett*. **1993**, 34, 3733.
106 Department of Chemistry, Uppsala University, S-751 24 Uppsala, Sweden – see http://www.org.kemi.uu.se/Research/Pia/Piahem.shtm
107 Kongsaeree,P. Department of Chemistry and Center for Protein Structure and Function, Mahidol University, Rama VI Road, Bangkok 10400.
108 Kruger, H. G. *J. Mol. Struct. (Theochem)* **2002**, 577, 281.
109 Kruger, H. G.; Mdluli, P.; Power, T. D.; Raasch, T.; Singh, A. *J. Mol. Struct. (Theochem*) submitted for publication.
110 Gokul, V.; Kruger, H. G.; Govender, T.; Fourie, L.; Power, T. D. *J. Mol. Struct. (Theochem)* **2004**, 672, 119.
111 Oie, T.; Loew, G. H.; Burt, S. K.; Brinkley, J. S.; MacElroy, R. D. *J. Am. Chem. Soc*. **1983**, 105, 2221.
112 (a) Fields, G. B.; Tian, Z.; Barany, G.; Alberico, F.; Peptide chemistry lecture notes: Principles and practice of Solid Peptide Synthesis.
(b) Alberico, F., Barany, G. *Tetrahedron Lett*. **1991**, 32, 1015.
113 Hay, R. W.; Porter, L. J. *J. Chem. Soc*. **1967**, 1261.
114 Oie, T.; Loew, G. H.; Burt, S. K.; MacElroy, R. D. *J. Am. Chem. Soc*. **1984,** 106(26), 8007.
115 Navobiochem **1997/98**, S22.
116 Nicolis, E.; Pedroso, E.; Giralt, E. *Terahedront. Lett.* **1989**, 30, 497.
117 Kenner, G. W.; Seely, J. H. *J. Am. Chem. Soc*. **1972**, 94, 3259.
118 Okada, Y.; Iguchi, S.; Kawasaki, K. *J. Chem. Soc*., *Chem. Commun*. **1987**, 1532.
119 Okada, Y.; Iguchi, S. *J. Chem. Soc*., *Perkin Trans. 1* **1988**, 2129.
120 Cramer, C. J. *Essentials of computational chemistry, theories and models.* 2nd ed.; John Wiley & Son Ltd, **2004**.
121 Singh, T. *M. Tech Dissertatio*, Durban Institute of Technology, **2003.**
122 Simos, J. *An Introduction to theoretical chemistry*, Cambridge University Press, **2003**.
123 Foresman, J. B.; Frisch, Æ. *Exploring Chemistry with electronic structure methods*, 2nd ed.; Gaussian, Inc. 1996.
124 Oie, T.; Loew, G. H.; Burt, S. K.; Brinkley, J. S.; MacElroy, R. D. *J.Am. Chem. Soc*. **1982**, 104(23), 6169.
125 Jensen, F. *Introduction to Computational chemistry*, John Wiley & Son Ltd, **1999**.
126 http://webpages.marshall.edu/%7Epricew/chem307/tutorials/Comp_chem/intro.html
127 Allinger, N. L. *J. Am. Chem. Soc*. **1977**, 99, 3279.
128 Allinger, N. L.; Yuh, Y. H.; Lii, J. H. *J. Am. Chem. Soc*. **1989**, 111, 8551.
129 http://europa.chem.uga.edu/allinger/mm2mm3.html
130 http:// www.hyper.com/
131 Alchemy 2000, version 2.05,Tripos, Inc. **1997-1998**.
132 Stewart, J. J. P. *MOPAC manual*, version 6.00, Frank J. Seiler Research Laboratory, **1990**.
133 Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; MontgomeryJr, J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M. T.; Nakajima, Y.; Honda, O.; Kitao, H.; Nakai, M.; Klene, X.; Li, J. E.; Knox, H. P.; Hratchian, Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P.; Morokuma, Y. K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng,

C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; and Pople, J. A. Gaussian, Inc. Pittsburgh PA, **2003**.
134 Dupuis, M.; Spamgler, D.; Wendoloski, J. J. NRCC Software Catalog, Vol.1 Program QG01(GAMESS), **1980**.
135 Jaguar. version 6.0, Schrodinger, LLC,New York, **2005**.
136 Leach, A. R. *Molecular modelling, principles and applications*, Addison Wesley Longman Ltd, **1996**.
137 Cizek, J. *J. Chem. Phys*. **1966**, 45, 4256.
138 Cizek, J. *Adv. Chem. Phys*. **1969**, 14, 35.
139 Lipkowitz, K. B.; Boyd, D. B. *Reviews in computational chemistry*, New York-VCH: Weinheim, **1990**.
140 Hohenburg, P.; Kohn, W. *Phys. Rev*. B **1964**, 136, 864.
141 Kohn W.; Sham, L. J. *Phys. Rev*. A **1965**, 140, 1133.
142 Hehre, W. J. *A Guide to Molecular Mechanics and Quantum Chemical Calculations*, Wavefunction, Inc. **2003**, 458.
143 http://en.wikipedia.org/computaional_chemistry
144 Becke, A. D. *J. Chem. Phys*. **1993**, 98, 5648.
145 Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, 37, 785.
146 Dupuis, M.; Watts, J. D.; Villar, H. O.; and Hurst,G. J. B.; HONDO, version 7.0, **1987**.
147 Hehre, W. J.; Radom, L.; Schleyer, P. V. R.; Pople, J. A. *Ab initio molecular orbital theory*, Wiley, New York, **1986**.
148 Hehre, W. J.; Stewart, R. F.; Pople, J. A. *J. Chem. Phys*. **1969**, 51, 2657.
149 Hehre, W. J.; Ditchfield, R.; Stewart, R. F.; Pople, J. A. *J. Chem, Phys*. **1970**, 52, 2769.
150 Peitro, W. J.; Levi, B. A.; Henhre, W. J.; Stewart, R. F. *Inorg. Chem*. **1980**,19,2225.
151 Peitro, W. J.; Blurock, E. S.; Hout, R. S.; Henhre, W. J.; DeFrees, D. J.; Stewart, R. F. *Inorg. Chem*. **1981**, 20,3650.
152 Peitro, W. J.; Henhre, W. J. *J. Comput. Chem*. **1983**, 4, 241.
153 Binkely, J. S.; Pople, J. A.; Hehre, W. J. *J. Am. Chem. Soc*. **1980**, 102(3), 939.
154 Hellmann, H. *J. Chem. Phys.* **1935**, 3, 61.
155 Bruice, P.Y. *Organic chemistry*, 3rd ed.; Prentice Hall, **2001**.
156 Miertus, S.; Scrocco, E.; Tomasi, J. *J. Chem. Phys*. **1981**, 55, 117.
157 Xu, S.; Wang, C.; Sha, G.; Xie, J.; Yang, Z. *J. Mol. Struct. (TheoChem)* **1999**, 459, 163.
158 Caricato, M.; Ingrosso, F.; Mennucci, B.; Tomasi, J. *J. Chem. Phys*. **2005**, 122, 154501.
159 Wong, M. W.; Wiberg, K. B., Frisch, M. J. *J. Am. Chem. Soc*. **1992**, 114, 1645.
160 Solomons, T. W. G.; Fryhle, C. B. *Organic Chemistry*, 8th ed.; John Wiley & Son, Inc. **2004**.
161 Clayden, J.; Greeves, N.; Wothers, P. *Organic Chemistry*, 1st ed.; Oxford University Press, Oxford New York, **2001**.
162 Quideau, S.; Looney, M. A.; Pouységu, L.; Ham, S.; Birney, D. M. *Tetrahedron Lett*. **1999**, 40, 615.
163 Jali, S., Kruger, H.G., Mdluli, S.P. *J Org. Chem.* submitted for publication.
164 Maguire, G. E. M.; Paraskevopoulos, J. *University of KwaZulu-Natal*, personal communication **2005**.
165 Skyes, P. *A guide to Mechanism in Organic Chemistry*, 1st ed.; Longman group Ltd: London, **1961**.
166 Govender, T. *PhD, University of KwaZulu-Natal*, personal communication in **2004**.