

**Some pages of this thesis may have been removed for copyright restrictions.**

If you have discovered material in Aston Research Explorer which is unlawful e.g. breaches copyright, (either yours or that of a third party) or any other law, including but not limited to those relating to patent, trademark, confidentiality, data protection, obscenity, defamation, libel, then please read our [Takedown policy](#) and contact the service immediately (openaccess@aston.ac.uk)

DNA METHYLATION AT CYTOSINE POSITION 5

Graeme Martin Currie

Doctor of Philosophy

THE UNIVERSITY OF ASTON IN BIRMINGHAM

May 1992

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with its author and that no quotation from the thesis and no information derived from it may be published without the author's prior, written consent.

# DNA methylation at cytosine position 5

by

Graeme Martin Currie

submitted for the degree of Doctor of Philosophy, 1992.

The University of Aston in Birmingham.

## Summary

DNA methylation appears to be involved in the regulation of gene expression. Transcriptionally inactive (silenced) genes normally contain a high proportion of 5-methyl-2'-deoxycytosine residues whereas transcriptionally active genes show much reduced levels. There appears good reason to believe that chemical agents capable of methylating 2'-deoxycytosine might affect gene expression and as a result of hypermethylating promoter regions of cytosine-guanine rich oncogenic sequences, cancer related genes may be silenced.

This thesis describes the synthesis of a number of 'electrophilic' S-methylsulphonium compounds and assesses their ability to act as molecules capable of methylating cytosine at position 5 and also considers their potential as cytotoxic agents.

DNA is methylated *in vivo* by DNA methyltransferase utilising S-adenosylmethionine as the methyl donor. This thesis addresses the theory that S-adenosylmethionine may be replaced as the methyl donor for DNA methyltransferase by other sulphonium compounds. S-[<sup>3</sup>H-methyl]methionine sulphonium iodide was synthesised and experiments to assess the ability of this compound to transfer methyl groups to cytosine in the presence of DNA methyltransferase were unsuccessful.

A proline residue adjacent to a cysteine residue has been identified to be a highly conserved feature of the active site region of a large number of prokaryotic DNA methyltransferases. The thesis examines the possibility that short peptides containing the Pro-Cys fragment may be able to facilitate the alkylation of cytosine position 5 by sulphonium compounds. Peptides were synthesised up to 9 amino acids in length but none were shown to exhibit significant activity.

Molecular modelling techniques, including Chem-X, Quanta, BIPED and protein structure prediction programs were used to assess any structural similarities that may exist between short peptides containing a Pro-Cys fragment and similar sequences present in proteins. A number of similar structural features were observed.

**Keywords:** 5-methylcytosine, DNA methyltransferase, oncogene, sulphonium, methylation, Pro-Cys, modelling,

To Mum and Dad

## Acknowledgements

The author wishes to thank Professor Malcolm Stevens for his support throughout the period of this research.

I would also like to thank my colleagues Dr. Clare Sansom, for her help in the use of Quanta and BIPED; Dr. Paul Yates for solving the dipeptide crystal structure; Dr. Paul Hepburn for his work in assaying the compounds; and Dr. Carl Schwalbe for his advice and discussion with regard to molecular modelling work. I am also very grateful to Professor R.Ramage for allowing me to work in his laboratories and for the use of the solid phase peptide synthesiser and to Dr. Dave Nicholls for his invaluable help in the proof reading of this thesis.

Finally I would like to thank the staff, postgraduate and undergraduate students from the Department of Pharmaceutical Sciences for making my period of research very enjoyable.

Many thanks are also extended to Dr. Mike Perry for running the high resolution NMR spectra and Karen Farrow for conducting the mass spectral analyses.

<b>Contents</b>		<b>Page</b>
Title Page		1
Abstract		2
Dedication		3
Acknowledgements		4
Contents		5
List of Figures		12
List of Schemes and Plates		16
List of Tables		17
Abbreviations		19
<b>Section 1: Introduction</b>		<b>22</b>
Chapter 1	Introduction	23
	1.1 Overview	23
	1.2 Occurrence of 5-methylcytosine	26
	1.3 Mechanism of methylation	27
	1.4 Eukaryotic methyltransferase enzymes	28
	1.5 Methods to determine methylated bases	31
	1.6 Possible roles of 5-methylcytosine	33
Chapter 2	Cytosine methylation and gene expression in eukaryotic cells	34
	2.1 Introduction	34
	2.2 Experimental evidence linking DNA methylation and gene control	35
	2.2.1 Correlation experiments	35

	2.2.2 Expression of genes <i>in vitro</i> using methylation inhibitors	36
	2.3 Aberrant methylation	39
Chapter 3	Cytosine methylation and cancer	41
	3.1 Introduction	41
	3.2 Hypomethylation in tumour tissue	42
	3.3 Effects of methylation inhibitors on malignancy	44
	3.4 Inhibition of DNA methylation by carcinogens	45
	3.5 Methylation, ageing and cancer	46
Chapter 4	Rationale for synthetic molecules	47
	4.1 Introduction	47
	4.2 Methylation inhibitors	48
	4.3 Potential synthetic targets	49
<b>Section 2:</b>	<b>Results and discussion</b>	<b>54</b>
Chapter 5	Restriction enzyme studies on simple sulphonium compounds	55
	5.1 Experiments to determine the sites of alkylation in DNA of simple sulphonium compounds using restriction enzymes	55
Chapter 6	Molecular modelling studies	62
	6.1 Introduction	62
	6.2 Modelling methods	63
	6.2.1 Chem-X	63
	6.2.2 QUANTA	63
	6.2.3 Modelling data	64
	6.3 Conformational studies on the dipeptide Pro-Cys	65

	6.3.1	Minimum energy conformation of Pro-Cys	66
	6.3.2	Interaction of the energy minimised Pro-Cys dipeptide with a GC base pair	69
	6.3.3	Structural comparisons between a protected Pro-Cys dipeptide and Pro-Cys sequences when present in proteins	71
	6.4	Structure prediction of S-adenosylmethionine	77
	6.5	Protein structure prediction	81
	6.5.1	Introduction	81
	6.5.2	Results of predictions and discussions	83
Chapter 7		Synthesis of sulphonium compounds	94
	7.1	Synthesis of simple sulphonium salts	94
	7.1.1	Instability of sulphonium salts	96
	7.2	Synthesis of S-[ <sup>3</sup> H-methyl]methionine sulphonium iodide	98
	7.2.1	Purification of S-[ <sup>3</sup> H-methyl]methionine sulphonium iodide	99
	7.2.2	Synthetic methods	101
	7.3	Synthesis of peptidic precursors to sulphonium-thiol molecules	107
	7.3.1	Synthesis of N-CBZ-S-(Bzl)cysteinylmethionine methyl ester (18)	109
	7.3.2	Synthesis of N-CBZ-methionyl-S-(Trt)cysteine methyl ester (10)	109



7.3.3	Synthesis of N-[2,2-dimethyl-3-formyl-4-thiazolidinecarboxy] methionine methyl ester ( <b>19</b> )	111
7.3.4	Synthesis of bis(N-CBZ-methionyl)cystine dimethyl ester ( <b>22</b> )	113
7.3.5	Synthesis of [N-Fmoc-S-(Trt)cysteinamido]-4-thioanisole ( <b>25</b> )	114
7.4	Synthesis of sulphonium-thiol molecules	115
7.4.1	Chemical stability of sulphonium-thiol molecules	115
7.4.2	Approaches to the alkylation of thiol-containing peptide molecules	116
7.5	Synthesis of peptidic sulphonium compounds	119
7.5.1	Alkylation of synthetic peptide precursor molecules	119
7.5.2	Alkylation of bis(N-CBZ-methionyl)cystine dimethyl ester ( <b>22</b> )	121
7.5.3	Alkylation of [N-2,2-dimethyl-3-formyl-4-thiazolidinecarboxy] methionine methyl ester ( <b>19</b> )	124
7.5.4	Alkylation of [N-Fmoc-S-(Trt)cysteinamido]-4-thioanisole ( <b>25</b> )	125
7.6	Deprotection of selected sulphonium-thiol peptidic molecules	126

	7.6.1	Cleavage of the trityl group from N-CBZ-S-methylmethionyl-S-(Trt)cysteine methyl ester sulphonium iodide (29)	126
	7.6.2	Cleavage of other sulphonium-thiol molecules	129
	7.7	Biological testing	129
Chapter 8		Synthesis of peptides to activate the C5 position of cytosine	131
	8.1	Introduction	131
	8.2	Synthesis of target molecules	133
	8.2.1	Synthesis of Pro-Cys derivatives	133
	8.2.2	Synthesis of N-Boc-prolyl-S-(N-methylcarbamoyl)cysteiny glycine ethyl ester (40)	135
	8.3	Design and synthesis of complex peptides with potential C5 activating ability	140
	8.3.1	Methodology	142
	8.3.2	Synthesis of target peptides	146
	8.4	Cleavage and deprotection of peptides 53-55 and 37	149
	8.4.1	Cleavage and deprotection of Cys-Ala-Pro-Cys-Leu-Pro-Ala-Gly (53)	149
	8.4.2	Cleavage and deprotection of Gly-Phe-Pro-Cys-Pro-His-Phe-Ser (54)	152
	8.4.3	Cleavage and deprotection of Ser-Pro-Cys-Met-Arg-Leu-Ile-His-Leu (55)	156
	8.4.4	Cleavage and deprotection of	159

Pro-Cys-His-Met-Asp (37)

	8.4.5 Purity of peptides for biological evaluation	162
Chapter 9	Development of an assay system for C5 methylating agents	164
	9.1 Requirement for a new assay system	164
	9.2 Experiments to determine the methylation sites of both S-methylmethionine sulphonium iodide and S-adenosylmethionine	166
	9.2.1 Experiments to determine whether S-methylmethionine sulphonium iodide is capable of acting as the methyl donor for DNA MTASE	166
	9.2.2 Experiments to determine whether S-methylmethionine sulphonium iodide is capable of independently alkylating cytosine at position C5	167
	9.3 Experiments to determine the ability of synthetic peptides to facilitate the alkylation of cytosine at position C5	170
	9.4 Cytotoxicity assays	172
Chapter 10	Conclusions and further work	176
	10.1 Development of small molecules as C5 methylating agents	176
	10.2 Identification of the active site sequence of eukaryotic DNA MTASE	177
	10.3 Structural homology of peptides	177
	10.4 Methylating ability of sulphonium compounds	178

<b>Section 3: Materials and Methods</b>	180
Chapter 11 Experimental	181
11.1 Chemistry experimental	181
11.1.1 Synthesis of sulphonium salts	181
11.1.2 Synthesis of amino acid derivatives	187
11.1.3 Synthesis of peptidic compounds	188
11.2 Stability study on sulphonium centre in the presence of thiol by 300 MHz <sup>1</sup> H-NMR	198
11.3 Study on methylation reaction to produce S-methylmethionine sulphonium iodide	198
11.4 Solid phase peptide synthesis	199
11.5 Biological methods	203
11.5.1 DNA MTASE assay	203
11.5.2 DNA hydrolysis	204
11.5.3 HPLC analysis	204
11.5.4 Assay for agents methylating cytosine at position 5	204
11.6 Cytotoxicity assays	206
11.6.1 Cells	206
11.6.2 Methods	206
Bibliography	207
Appendix 1 Crystal structure of N-CBZ-prolyl-S-(Bzl)cysteine methyl ester (5)	219

## List of Figures

	<b>Page</b>
Figure 1: Structures of DNA bases	23
Figure 2: Proposed method of gene control by DNA methylation	25
Figure 3: Inverted repeat sequences	27
Figure 4: Catalytic mechanism of DNA MTASE	29
Figure 5: 5-Azacytidine	37
Figure 6: Ethionine	37
Figure 7: The O <sup>6</sup> position and N <sup>7</sup> position of guanine, potential sites for alkylation	41
Figure 8: Structure of S-isobutyryladenine: an inhibitor of DNA MTASE	48
Figure 9: An carbonionic intermediate formed after reaction of DNA MTASE with the C6 position of cytosine	50
Figure 10: The absolute configuration of the sulphonium centre of S-adenosylmethionine	51
Figure 11: General structures proposed for target molecules	53
Figure 12: The principles behind separation by agarose gel electrophoresis	56
Figure 13: Structures of the simple sulphonium compounds used in the restriction enzyme studies	57
Figure 14: Proposed binding pocket for S-adenosylmethionine in DNA MTASE	59
Figure 15: Restriction enzyme cutting pattern for S-methylmethionine sulphonium iodide treated $\lambda$ -phage DNA	60
Figure 16: Standard <i>cis</i> and <i>trans</i> amide bonds showing phi and psi torsion angles	67

Figure 17:	A Ramachandran map comparing the phi and psi angles for proline from the crystal structure of N-CBZ-prolyl-S-(Bzl)cysteine methyl ester ( <b>5</b> ) with the phi and psi angles for proline from 10 proteins selected from the BIPED database	72
Figure 18:	A Ramachandran map comparing the phi and psi angles for cysteine from the crystal structure of N-CBZ-prolyl-S-(Bzl)cysteine methyl ester ( <b>5</b> ) with the phi and psi angles for cysteine from 10 proteins selected from the BIPED database	74
Figure 19:	The conversion of SAM to SAH as a result of methyl donation	77
Figure 20:	Atomic numbering scheme for the molecular model of S-adenosylmethionine	79
Figure 21:	Graphical displays of the Chou-Fasman and Garnier-Osguthorpe-Robson structure predictions for peptide <b>53</b>	85
Figure 22:	Graphical displays of the Chou-Fasman and Garnier-Osguthorpe-Robson structure predictions for peptide <b>54</b>	86
Figure 23:	Graphical displays of the Chou-Fasman and Garnier-Osguthorpe-Robson structure predictions for peptide <b>55</b>	88
Figure 24:	Graphical display of the Garnier-Osguthorpe-Robson structure prediction for EcorII	89
Figure 25:	Graphical display of the Garnier-Osguthorpe-Robson structure prediction for MspI	90
Figure 26:	Graphical display of the Garnier-Osguthorpe-Robson structure prediction for BsprI	91
Figure 27:	Mechanism of elimination that may occur at sulphonium centres	97
Figure 28:	HPLC purification profile of S-[ <sup>3</sup> H-methyl]methionine	102

	sulphonium iodide	
Figure 29:	<sup>1</sup> H NMR stack plot of the synthesis of S-methylmethionine sulphonium iodide showing the region δ 2.0-3.0 ppm for the first 17 hrs of reaction	103
Figure 30:	Graphical display of S-methylmethionine sulphonium iodide formation monitored by <sup>1</sup> H NMR	105
Figure 31:	Structures of possible precursors to sulphonium-thiol molecules	108
Figure 32:	Two possible decomposition routes of sulphonium thiol molecules	115
Figure 33:	Structure of a thiosulphonium salt	121
Figure 34:	Sites in bis(N-CBZ-methionyl)cystine dimethyl ester ( <b>22</b> ) that are potential targets for alkylation	122
Figure 35:	Proposed fragmentation pattern of N-CBZ-S-methylmethionyl-N-CBZ-methionylcystine dimethyl ester ( <b>30</b> ) as suggested by the FAB mass spectrum	123
Figure 36:	Sites in N-[2,2-dimethyl-3-formyl-4-thiazolidinecarboxy]methionine methyl ester ( <b>19</b> ) that are potential targets for alkylation	124
Figure 37:	Mechanism of alkylation of aromatic sulphides by methyl iodide	125
Figure 38:	<sup>1</sup> H NMR spectrum of N-CBZ-S-methylmethionyl-S-(Trt)cysteine methyl ester sulphonium iodide ( <b>29</b> )	128
Figure 39:	<sup>1</sup> H NMR spectrum of N-CBZ-S-methylmethionyl-S-(Trt)cysteine methyl ester sulphonium iodide ( <b>29</b> ) after incubation for 30 mins with 95% TFA and anisole	130
Figure 40:	Structures of protected derivatives of prolinylcysteine peptides synthesised	134
Figure 41:	The benzyloxybenzylalcohol linker molecule, an acid cleavable linker molecule for solid phase peptide synthesis	143

Figure 42:	Method flow diagram for solid phase synthesis of peptides 53-55 and 37	144
Figure 43:	Synthetic pathways for the solid phase synthesis of peptides 53-55 and 37 on an Applied Biosystems 430A peptide synthesiser	145
Figure 44:	UV deprotection profile, monitoring the removal of the Fmoc group after peptide coupling during the solid phase synthesis of peptide 53	147
Figure 45:	HPLC profile of crude peptide 53 after 4 hrs incubation with 95% TFA:anisole:ethylmethysulphide (91:4.5:4.5)	151
Figure 46:	HPLC profiles of crude peptide 54 after 1, 3 and 4 hrs incubation with 95% TFA: anisole:ethylmethysulphide (91:4.5:4.5)	153
Figure 47:	HPLC profile of crude Peptide 54 after complete cleavage from the resin and deprotection by 95%TFA:anisole:ethylmethysulphide (91:4.5:4.5)	155
Figure 48:	Structure of 2,2,5,7,8-pentamethylchromanarginine	156
Figure 49:	HPLC profile of crude peptide 55 after complete cleavage from the resin and deprotection by 95%TFA:anisole:thioanisole (91:4.5:4.5)	158
Figure 50:	Proposed fragmentation pathway of the peptidic compound (58) isolated by HPLC possessing a $R_t = 10.67$ mins from the cleavage reaction of peptide (55)	160
Figure 51:	$^1\text{H}$ NMR spectrum of prolyl-S-(Trt)cysteinylnhistidinylmethionylaspartate (38)	161
Figure 52:	The effect of S-methylmethionine sulphonium iodide and SAM on the methylation of <i>M. lysodeikticus</i> DNA in the presence and absence of DNA MTASE	168



## List of Schemes and Plates

Scheme		Page
1	Synthesis of N-CBZ-S-(Bzl)cysteinylmethionine methyl ester ( <b>18</b> )	109
2	Synthesis of N-CBZ-methionyl-S-(Trt)cysteine methyl ester ( <b>10</b> )	110
3	Synthesis of N-[2,2-dimethyl-3-formyl-4-thiazolidinecarboxy] methionine methyl ester ( <b>19</b> )	112
4	Synthesis of bis(N-CBZ-methionyl)cystine dimethyl ester ( <b>22</b> )	113
5	Synthesis of [N-Fmoc-S-(Trt)cysteinamido]-4-thioanisole ( <b>25</b> )	114
6	Synthesis of N-CBZ-prolyl-S-(Bzl)cysteinyglycine methyl ester ( <b>43</b> )	136
7	Synthesis of N-CBZ-prolyl-S-(Trt)cysteinyglycine methyl ester ( <b>47</b> )	138
8	Synthesis of N-CBZ-prolyl-S-(Nmc)cysteinyglycine ethyl ester ( <b>40</b> )	139
Plate		
1	Crystal structure of N-CBZ-prolyl-S-(Bzl)cysteine methyl ester ( <b>5</b> )	68
2	Comparison of a computer generated energy minimised structure of prolylcysteine (A) with an energy minimised structure generated from the crystal structure of N-CBZ-prolyl-S-(Bzl)cysteine methyl ester (B)	68
3	Docking of an energy minimised structure of prolinylcysteine into a GC base pair	70
4	Comparison of an energy minimised structure of prolylcysteine with the crystal structure of N-CBZ-prolyl-S-(Bzl)cysteine methyl ester ( <b>5</b> )	70
5	Comparison of 4 tetra-peptide sequences from proteins extracted from the Brookhaven database	76

## List of Tables

Table		Page
1	Restriction enzymes	32
2	Cleavage sites of MspI and HpaII	55
3	Phi and psi angles of four protease enzymes identified from the Brookhaven database containing a proline residue adjacent to a cysteine residue	75
4	Final energies and relevant torsion angles of a computer generated structure of SAM	80
5	Amino acid sequences from bacterial MTASE's on which protein structure prediction calculations were performed	83
6	Conditions for the formation of alkyl sulphonium salts of methionine	95
7	HPLC conditions for the separation of S-methylmethionine sulphonium iodide and methionine	100
8	Conditions for the synthesis of S-methylmethionine sulphonium iodide	104
9	Sulphonium generation in N-CBZ-S-(Bzl)cysteinylmethionine methyl ester ( <b>18</b> ) with methyl iodide	117
10	Sulphonium formation in N-CBZ-S-(Bzl)cysteinylmethionine methyl ester ( <b>18</b> ) with other alkylating agents	118
11	Conditions for the synthesis of sulphonium peptides with methyl iodide	120

12	Prokaryotic methylene transfer enzymes	132
13	Conditions for the synthesis of prolylcysteine derivatives	133
14	Active site regions in DNA MTASEs	141
15	Target peptide sequences	142
16	Details of peaks from the analytical HPLC purification of peptide <b>54</b>	154
17	Major peaks present in the HPLC chromatogram of crude peptide <b>55</b>	157
18	Purity of peptides for biological evaluation	163
19	Assay of synthetic peptides to determine their ability to aid alkylation of DNA in presence of [ <sup>3</sup> H]-S-adenosylmethionine	171
20	Assay for inhibition of A549 cell growth by sulphonium compounds	174

## Abbreviations

ABI	Applied Biosystems International
AcO-AAF	2-N-Acetoxyacetylaminofluorene
A	Adenine
Ala	Alanine
Arg	Arginine
Asp	Aspartate
5-AzaC	5-Azacytidine
b	Broad
BIPED	Brookhaven interactive protein engineering database
Boc	Tertiarybutyloxycarbonyl
tBu	Tertiarybutyl
Bzl	Benzyl
C	Cytosine
Ci	Curies
C5	Carbon 5
C6	Carbon 6
$\mu$ Ci	Microcuries
cDNA	Complementary DNA
Conc	Concentrated
CPM	Counts per minute
CBZ	Benzyloxycarbonyl
Cys	Cysteine
$^{\circ}$ C	Degree celcius
DCC	Dicyclohexylcarbodiimide
DIC	Diisobutylcarbodiimide
DMF	Dimethylformamide
DMAP	Dimethylaminopyridine
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
EDIC	1-Ethyl-2-2-diisopropylcarbodiimide
eV	Electron volts
FAB	Fast atom bombardment
FEL	Friend erythroleukaemia
Fig.	Figure

Fmoc	Fluorenylmethoxycarbonyl
G	Guanine
Gly	Glycine
Glu	Glutamic acid
Gln	Glutamine
hrs	Hours
His	Histidine
HL60	Hairy cell leukaemia
HOBT	Hydroxybenzotriazole
HPLC	High pressure liquid chromatography
HSA	Hexanesulphonic acid
Ile	Isoleucine
IR	Infra-red
IC <sub>50</sub>	Concentration required to decrease growth by 50%
Kcal	Kilocalories
Kda	Kilodaltons
lit.	Literature
Leu	Leucine
mCi	Millicuries
μCi	Microcuries
μl	Microlitres
mol	Mole
M	Molar
mmol	Millimole
pmol	Picomole
mM	Millimolar
Me	Methyl
Met	Methionine
MHz	Megahertz
min	Minute(s)
ml	Millilitres
m.p.	Melting point
MTASE	Methyltransferase
MNNG	N-Methyl-N'-nitro-N-nitroso-guanidine
MNU	Methyl nitrosourea
5-MC	5-Methylcytosine

MOPAC	Molecular orbital package
MSA	Methanesulphonic acid
Nhs	N-Hydroxysuccinimide
nm	Nanometres
Nmc	N-Methylcarbamoyl
NMR	Nuclear magnetic resonance
NOBA	Nitrobenzaldehyde
p	Phosphate
ppm	Parts per million
pet ether	Petroleum ether
Phe	Phenylalanine
Pmc	2,2,5,7,8-Pentamethylchroman
PPO	2,5 Diphenyloxazole
POPOP	1,4 Di-2-(5-phenyloxazolyl)benzene
Pro	Proline
PTFE	Polytetrafluoroethylene
RP	Reverse phase
R <sub>t</sub>	Retention time
SAM	S-Adenosylmethionine
SAH	S-Adenosylhomocysteine
SCX	Strong cation exchanger
SERC	Science and Engineering Research Council
Ser	Serine
T	Thymine
Temp	Temperature
Tk	Thymidine kinase
TLC	Thin layer chromatography
TFA	Trifluoroacetic acid
Thr	Threonine
THF	Tetrahydrofuran
trityl	Triphenylmethyl
Trt	Trityl
tRNA	Transfer ribonucleic acid
UV	Ultraviolet
V/V	Volume for volume
λ	Wavelength

## **Section 1: Introduction**

## CHAPTER 1

### Introduction

#### 1.1 Overview

The aim of the work described in this thesis is an ambitious one—that is, to examine possible chemical strategies to mimic biological methylation of the genome. By way of an introduction, the biological role of DNA methylation and its relevance to genomic control will be presented. After defining the goals of the thesis, the chemistry used to achieve the proposed objectives will be discussed, in particular the fields of sulphonium and peptide chemistry.

DNA contains four major bases, adenine(A), thymine(T), cytosine(C), and guanine(G) (Fig.1), and a number of minor bases such as 5-methylcytosine (5-MC), N<sup>6</sup>-methyladenine and N<sup>4</sup>-methylcytosine.

DNA from prokaryotes can contain all the minor bases. In eukaryotes 5-MC is the predominant modified base.

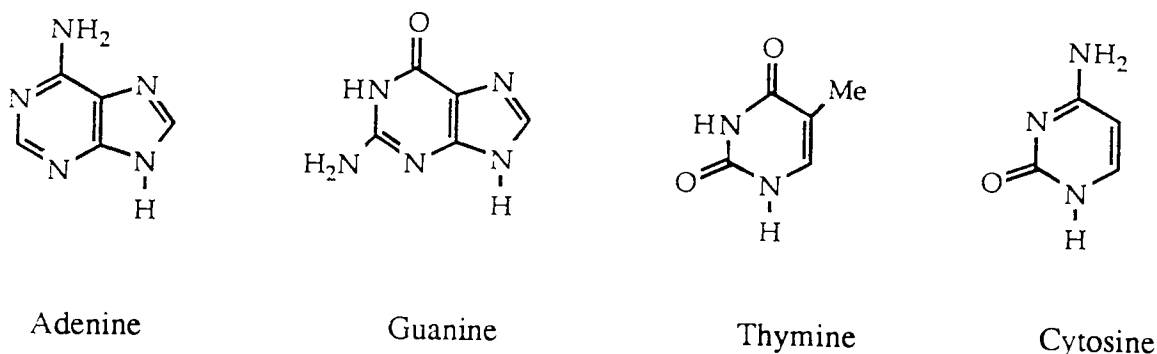


Figure 1 Structures of DNA bases



The existence of 5-MC was first observed in 1948 by R.D.Hotchkiss.<sup>1</sup> Methylation was then disregarded as having any biological significance. Twenty-seven years later in 1975 Holliday proposed the hypothesis that methylation is heritable, and passed from generation to generation as cells divide.<sup>2</sup>

He suggested from this that the three base coding system for protein synthesis using the four deoxyribonucleotide bases might be supplemented by the presence of 5-MC. A growing body of evidence has been accumulated linking the modified nucleotide 5-MC with eukaryotic gene expression. In general quiescent (silent) genes are found to be hypermethylated at the 5'-end or in the promoter region, whereas active genes are hypomethylated in these regions (Fig.2). This phenomenon has been shown in some but not all genes.

Cancer cells are known to exhibit an aberrant pattern of gene expression in certain tumours. A number of oncogenes have been found to be relatively hypomethylated in malignant tissue when compared with the same gene examined in normal tissue.

In mammalian DNA up to 90% of the 5-MC residues are found in the dinucleotide CpG, therefore opposite strands are symmetrically methylated. Patterns of DNA methylation are maintained from cell division to cell division by DNA methyltransferase (MTASE) . All biochemical evidence indicates DNA is methylated in an early post-replicative step.<sup>3</sup> Only the newly synthesised strand is methylated.

In nature cytosine is methylated by the enzyme cytosine-5 DNA MTASE in the presence of the methyl donor S-adenosylmethionine(SAM) .

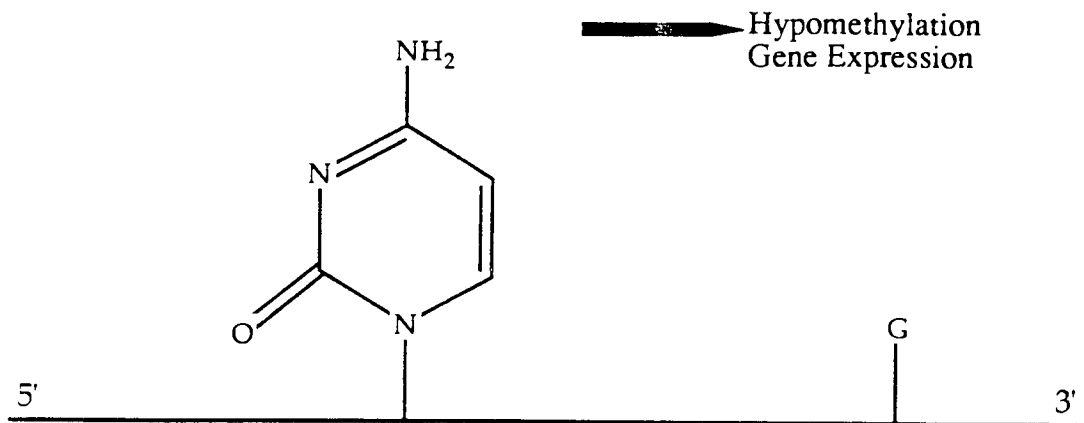
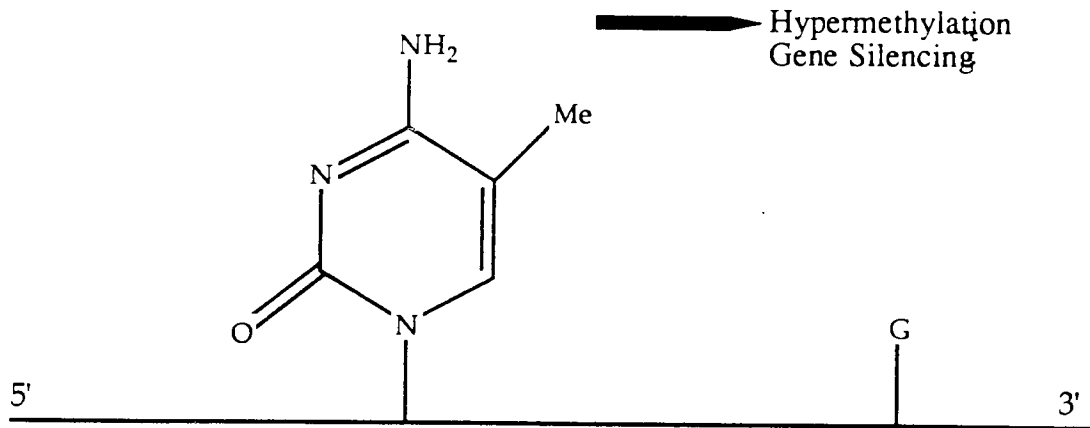


Figure 2 Proposed method of gene control by DNA methylation

The aim of the work described in this thesis is to test the hypothesis that small synthetic molecules can accomplish the methylation of the 5-position of cytosine.

## 1.2 Occurrence of 5-methylcytosine

The distribution of 5-MC throughout the genome suggests that it may have a biological function. 5-MC is present in only small amounts throughout the genome. Vertebrates have 3-7% of their DNA cytosine in a methylated state. This figure tends to rise as the evolutionary scale is ascended. An exception to the rule occurs in plants which normally possess around 30% of their total cytosine in the form of 5-MC.<sup>4</sup> As mentioned earlier 5-MC is normally found in the dinucleotide CpG. Unmethylated CpG tends to be found clustered in GC rich 'islands',<sup>5</sup> mostly towards the 5'-region of the gene. The dinucleotide is found to be statistically underrepresented in eukaryotes. On analysis one would expect to find 4% of all base pairs to be CpG: the actual figure found is 1%. One possibility is that the underrepresentation of CpG is due to the propensity of 5-MC to undergo deamination to thymidine, resulting in mutation. However this underrepresentation is not seen in prokaryotes.<sup>6</sup>

Chargaff *et al.* were the first to draw attention to the non-random distribution of the 5-MC throughout the genome in prokaryotes and eukaryotes.<sup>7</sup> Generally repetitive DNA sequences are found to have a higher 5-MC content than normal DNA. Inverted repeat sequences (Fig.3) within DNA are methylated to a significantly greater extent than ordinary repetitive sequences. Other areas of high 5-MC content include centromeric regions.<sup>8</sup>

GGAATTXXXTTAAGG

X = any base

Figure 3. Inverted repeat sequences

DNA in mammalian sperm and early embryos is usually highly methylated but there are some notable exceptions. Methylation free 'islands' found towards the 5'-end of 'housekeeping' genes and repetitive DNA in sperm tend to be found largely unmethylated. Tissue specific variations have also been found to exist. In the  $\beta$ -globin gene intron methylation levels in brain and sperm were shown to be 80 and 100% respectively.<sup>9</sup> Some species specificity has also been observed when examining the methylation patterns of tissues from rabbit and mouse. The non-random distribution of 5-MC in eukaryotes suggests the modified base does have a crucial role to play in cellular information.

### 1.3 Mechanism of methylation

DNA MTASE catalyses the methylation of 2'-deoxycytidine residues in DNA to 5-methyl 2'-deoxycytidine with a concomitant conversion of SAM to S-adenosylhomocysteine (SAH). A number of chemical and enzymatic studies have revealed that enzymes which catalyse the transfer of 1 carbon units to the 5-position (C5) of pyrimidine nucleotides proceed by a similar mechanism.<sup>10</sup> The mechanism of this class of enzymes has been established most thoroughly for thymidylate synthetase,<sup>11</sup> an enzyme whose role is similar

to that of DNA MTASE. The carbon at the 5-position of the pyrimidine ring is insufficiently nucleophilic for it to react with biological methyl donors of low electrophilicity. Therefore, there is a requirement for the C5 position of the pyrimidine ring to be made susceptible to methylation prior to methyl transfer.

The catalytic mechanism of the reaction is believed to involve two steps (Fig.4):

The first step involves a nucleophilic moiety, probably a cysteine residue in the enzyme, which is able to attack the carbon 6 (C6) position of the pyrimidine ring. The result of this attack is the opening of the double bond between carbon 5 (C5) and C6 of cytosine, forming a transient Michael-type adduct with anionic character at the C5 position.

The second step consists of an addition of an electrophile (i.e. methyl group from SAM ) to the C5 position. A  $\beta$ -elimination follows resulting in the enzyme leaving the C6 position and the double bond reforming. It is not known whether the two steps are concerted or sequential.

The stereospecificity of the reaction is determined by which face of the pyrimidine accepts the substituents. It is thought that attack by the enzyme and addition of the methyl group proceed in a stereospecific manner.

#### **1.4 Eukaryotic methyltransferase enzymes**

The structure and function of a wide variety of prokaryotic enzymes has been well documented.<sup>12</sup> The study of eukaryotic MTASE's has lagged somewhat behind. Only one eukaryotic enzyme has been sequenced to date. Two different types of DNA methylating activities have been proposed to exist in higher eukaryotes.<sup>13</sup>

## DNA METHYLATION

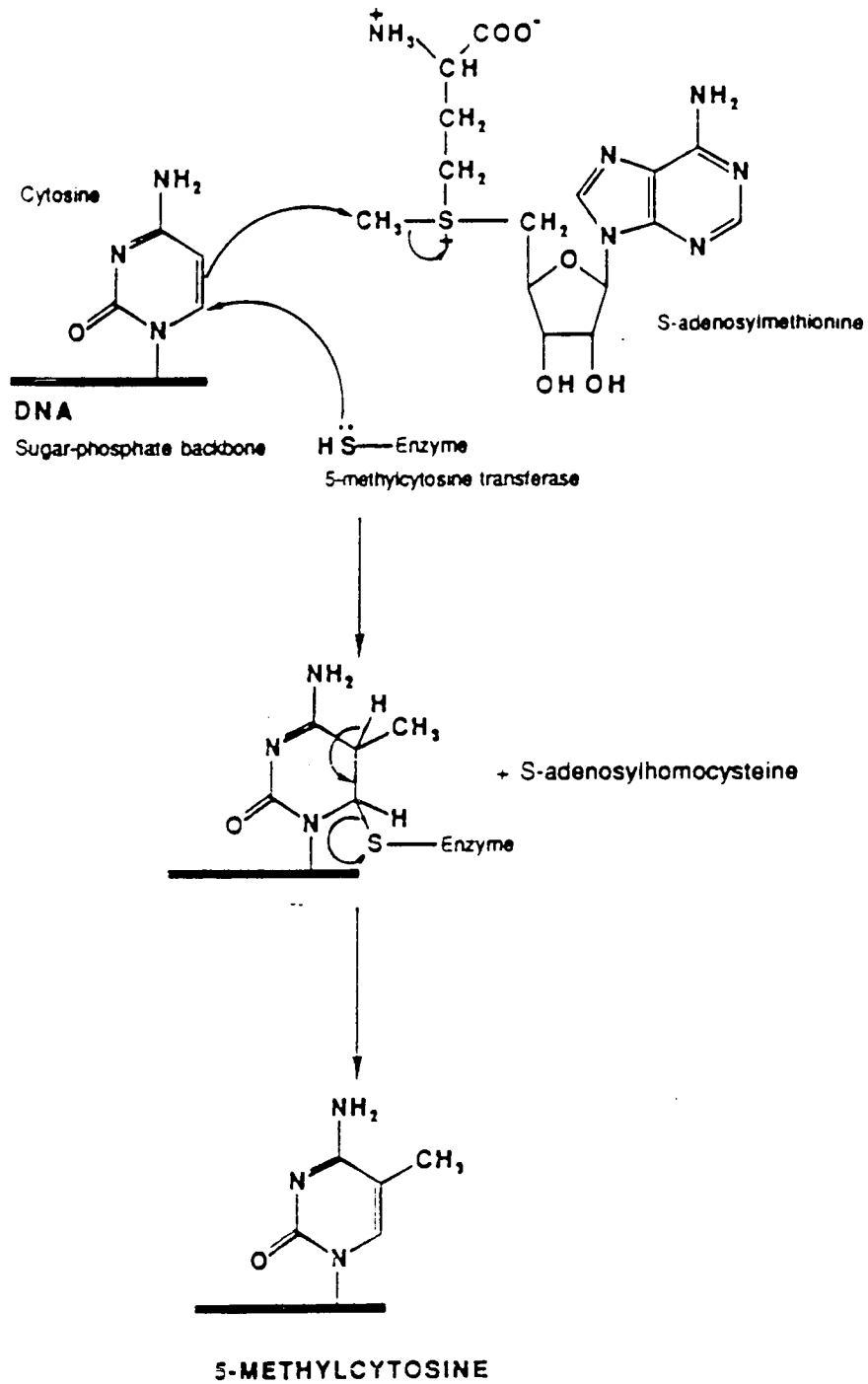


Figure 4 Catalytic mechanism of DNA MTASE

These are:

A *de novo* function which methylates previously unmethylated sites in a symmetrical manner on both strands of DNA.

A maintenance type function, which recognises hemimethylated sites in DNA and alkylates the opposite strand to produce a symmetrical methylation pattern.

The *de novo* function is thought to be a rare event *in vivo*.<sup>14</sup> However *in vitro* mammalian MTASE tends to show *de novo* rather than maintenance activity. All eukaryotic enzymes studied so far are able to methylate both double stranded and single stranded DNA *in vitro*. Another unique feature of the mammalian enzymes is their ability to methylate homologous DNA.<sup>15</sup> This possibly suggests activity may differ in *in vivo* and *in vitro* states.

The eukaryotic enzyme is a large enzyme over 100 Kda in size and is found predominantly in the cell nucleus associated with linker DNA in condensed regions of chromatin. The enzyme activity varies markedly throughout the cell cycle. Activity is low in the initial resting phase G<sub>1</sub>, increases to a maximum during the S phase (synthetic) and then gradually decreases during G<sub>2</sub> and M phases.

There is a high degree of homology present amongst the prokaryotic enzymes. This may be extended to the eukaryotic enzymes when more information is available.<sup>16</sup> In particular cytosine specific MTASEs contain a highly conserved amino acid sequence Pro-Cys-Gly thought to be at the active site.<sup>17</sup>

## 1.5 Methods to determine methylated bases

Several sensitive techniques are available for the detection of 5-MC in DNA. A number of chromatographic techniques can be employed including paper, column and thin layer chromatography.<sup>18</sup> Gas chromatography,<sup>19</sup> mass spectrometry,<sup>20</sup> and high pressure liquid chromatography (HPLC)<sup>21</sup> have been applied to detect down to microgram amounts of 5-MC. All the above methods require DNA to be extracted from the cell nucleus and acid hydrolysed to its constituent bases. The sensitivity of the HPLC method may be enhanced further by using radiolabelled DNA. 5-MC covalently bound to bovine serum albumin has been used to raise antibodies against 5-MC in rabbits. This approach has proved useful in the detection of small amounts of the modified base, but is limited by lack of antibody specificity.<sup>22</sup>

By far the most commonly used system for determining methylation patterns involves the use of restriction endonucleases, combined with electrophoretic techniques. Endonucleases operate by cutting DNA at sequence specific sites. Some restriction enzymes are prevented from working if their target sequence is methylated. An incidental discovery by Waalwijk and Flavell was of decisive importance.<sup>9</sup> They discovered a restriction endonuclease pair HpaII and MspI which operate on the same target sequence, CCGG. However if the internal cytosine residue is methylated MspI can cut, whereas HpaII cannot. Conversely if the external cytosine residue is methylated, MspI is rendered refractory. A number of the commonly used restriction endonucleases are listed in Table 1.



Table 1                      Restriction Enzymes

Enzyme	Source	Recognition sequence	Modified sequence refractory to cleavage
HhaI	<i>Haemophilus</i> <i>haemolyticus</i>	GCGC	GmCGC
HpaII	<i>Haemophilus</i> <i>parainfluenza</i>	CCGG	CmCGG
MspI	<i>Moraxella</i> species	CCGG	mCCGG
BSuRI	<i>Bacillus subtilis</i>	GGCC	GGmCC

The isochizomeric pair MspI and HpaII with complementary refractoriness, is the most commonly used system for elucidating the pattern of methylation in specific genes.<sup>9</sup> Restriction enzyme analysis is somewhat limited as it is only possible to detect a small number of the total methylated cytosines. This is due to restriction enzymes being restricted to their own specific target sequences.

The presence of 5-MC can also be detected using the Maxam-Gilbert nucleotide sequencing method. 5-MC does not appear in pyrimidine sequencing patterns, as it reacts relatively slowly with hydrazine compared with cytosine and thymine. The presence of 5-MC then can be detected by sequencing the complementary strand.<sup>23</sup> This method is not used generally.

## 1.6 Possible roles of 5-methylcytosine

The ubiquity of 5-MC suggests that it fulfills an important function. A number of hypotheses have been proposed to suggest its possible role.

Borek and Srinivisan<sup>24</sup> claimed that DNA methylation is involved in cell differentiation and that aberrant methylations may be involved in oncogenesis. These points will be discussed at greater length in Chapter 3.

Riggs<sup>25</sup> proposed that sequence specific DNA methylation may be a factor in the inactivation of the mammalian X-chromosome.

Holliday<sup>2</sup> suggested that damaged DNA, when corrected by repair synthesis, may not always be adequately methylated resulting in an altered methylation pattern giving rise to a malignant phenotype.

In prokaryotes protection of DNA from restriction endonucleases is the most obvious function for 5-MC.<sup>26</sup> In almost all instances restriction enzymes are prevented from cutting the DNA polymer if the bases are methylated. This also suggests that 5-MC may play a role in affecting DNA-protein interactions. The methylation patterns normally displayed in eukaryotic DNA argues against the restriction protection theory.

By far the most common view held, is that methylation plays a role in gene expression.

## CHAPTER 2

### Cytosine methylation and gene expression in eukaryotic cells

#### 2.1 Introduction

In the literature the body of evidence linking 5-MC and gene expression is substantial. Considering this evidence there seems good reason *a priori* to believe that chemical agents which can methylate DNA cytosine at position C5 might affect gene expression. The possibility also exists that molecules that can methylate cytosine C5 may be able to silence oncogenes and hence may prove attractive targets for novel chemotherapeutic agents. In order to support these proposals, it is necessary to review the evidence linking DNA methylation firstly to gene expression and then to cancer.

The regulatory principles of the genetic code are complicated and not yet fully understood. The first speculations relating DNA methylation to the control of gene activity in eukaryotes were formulated in 1975.<sup>27</sup> Since then the idea has been strongly supported. Razin and Riggs described a number of the salient points:<sup>25</sup>

- (i) specific methylation patterns exist.
- (ii) methylation is symmetrical in both strands.
- (iii) methylation patterns are clonally heritable.
- (iv) methylation patterns are tissue specific.

Genes can be divided into 3 classes: permanently activated, permanently inactivated and inactive genes which are occasionally reactivated. DNA methylation at highly specific sites is thought to be involved in long term inactivation. The possibility that only a few specific sites need to be unmethylated to affect gene expression is supported by a coordinated

appearance of deoxyribonuclease (DNase) I hypersensitive sites 5' to genes. DNase hypersensitivity is normally associated with increased transcriptional activity and a loss of methyl groups.<sup>28</sup> Methylation may act in several possible ways to control gene expression. It could possibly act as an on/off switch or as a modulator, controlling a number of transcription factors which can bind to DNA. Another possible role for methylation is to act as a centre from which non-transcribed genes can be packaged into inactive chromatin.

## **2.2 Experimental evidence linking DNA methylation and gene control**

### **2.2.1 Correlation experiments**

Inverse correlations between the extent of DNA methylation and the level of gene expression have been observed both in eukaryotic genes and viral systems. Generally the experimental approach has involved restriction enzyme studies using MspI and HpaII. Mandel and Chambon, whilst studying the chicken ovalbumin gene, found a correlation between hypomethylation and gene activity.<sup>29</sup> They observed three classes of sites in all tissues studied: namely, non-methylated, fully methylated and variably methylated. The variable sites tended to show less methylation in tissue where the gene was active. Studies on rabbit and human globin genes produced similar conclusions.<sup>30,31</sup> The phenomenon has also been observed in viruses where methylation has been studied more intensively. An inverse correlation between the degree of methylation of retro-viral proviruses and levels of transcription has been seen in a number of cases.<sup>32</sup>

More specifically studies have been carried out on individual genes before and after transformation. The thymidine kinase (Tk) gene of the herpes simplex virus, when in the Tk+ (expressed) form possesses only unmethylated Tk genes. The Tk- (unexpressed)

form is normally unstable and frequently converts to the Tk<sup>+</sup> form. However in one exceptional case an unusually unstable Tk<sup>-</sup> gene was found to be highly methylated.<sup>33</sup>

There are however a number of exceptions to the general trend. At least some CpG sites unmethylated in actively transcribed genes, are unmethylated in tissues where the gene is quiescent. Some actively transcribed genes contain many methylated CpG sites. Also the pattern of methylation of some CpG sites accessible to examination may not be altered when genes are activated or shut off. A specific example of the second point is shown by the oestrogen-inducible vitellogenin genes A1 and A2 in hepatocytes of *Xenopus laevis*. These genes are found to be heavily methylated irrespective of whether they are transcribed or not.<sup>34</sup> The information gained from correlation studies suggests a proposal that methylation may be linked to expression, but cannot support a cause-effect relationship, linking DNA methylation and gene expression.

### 2.2.2 Expression of genes *in vitro* using methylation inhibitors

More conclusive evidence has been uncovered by studying genes *in vitro* using methylation inhibitors. In a number of systems it has been possible to activate dormant genes by using agents thought to act in a manner so as to hypomethylate the DNA.

5-Azacytidine (5-AzaC) is one of these agents and has been a major tool in the study of gene expression (Fig.5).

5-AzaC can be incorporated into replicating DNA in place of 2'-deoxycytidine. In the presence of the enzyme DNA MTASE binding occurs as it would with cytosine, the resultant enzyme-5-AzaC complex formed is stable, thus the enzyme cannot leave the DNA being irreversibly bound and further DNA methylation is inhibited.

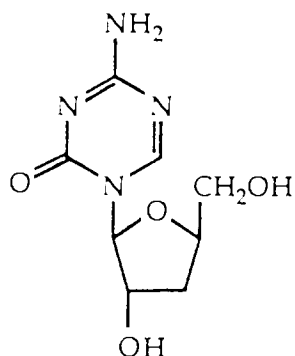


Figure 5 5-Azacytidine

The action of 5-AzaC is rapid; hemimethylated DNA can be observed after one S phase of the cell cycle following incubation with the drug.<sup>34</sup> 5-AzaC has been used in a number of different systems to study gene expression. Taylor and Jones<sup>35</sup> treated the mouse cell line 10 T<sub>1/2</sub> with 5-AzaC over a short period of time. After several generations foci of differentiated cells were seen. In order to eliminate mutagenesis as a cause of any effect, known mutagens were incubated with the cells. However the mutagens were unsuccessful in inducing differentiation. The effect was also observed in Hairy cell leukaemia HL60 cells.<sup>36</sup> 5-AzaC cannot activate all dormant genes. In the Ad-12 transformed hamster cell line (T637), inactive and hypermethylated late viral genes could not be activated although an alteration in methylation pattern was observed. A disadvantage of using 5-AzaC is that saturation levels can be reached which results in inhibition of its hypomethylatory effect.

Another methylation inhibitor used in studying gene control is ethionine, a methionine analogue (Fig.6).

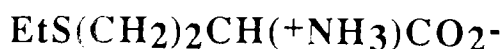


Figure 6 Ethionine

Ethionine in the form of S-adenosylethionine is a competitive inhibitor of a variety of MTASE reactions. Ethionine was shown to be an effective inducer of globin gene expression in Friend erythroleukaemia (FEL) cells.<sup>37</sup> The cells showed a tendency to become more hypomethylated as they matured. Several pieces of evidence support this hypothesis:

- (i) Hypomethylation was detected as early as 24 hours(hrs) after exposure to the inducing agent. The fraction of cells committed and the degree of hypomethylation increased with exposure time.
- (ii) FEL cell DNA does not become hypomethylated when cells are exposed to inducing agents under conditions where differentiation does not occur.

The experiments described above although open to interpretation were the first indication that hypomethylation and gene expression had definite links. Other inducers of differentiation have also been used to study methylation effects. Murine erythroleukaemia cells can be induced to differentiate by a variety of agents, for example; hexamethylenebisacetamide. On induction a transient genome wide hypomethylation occurs. If dexamethasone, which inhibits events critical for commitment to differentiation is administered, commitment is inhibited but hypomethylation is not prevented. This experiment highlights the point that hypomethylation may be present in silent genes.

The opinion is generally held that DNA hypomethylation is a necessary but not sufficient pre-condition for gene activation.<sup>6</sup> The control of gene activity being a mechanism of great importance is likely to be under the control of a multifaceted regulatory system of which methylation is probably one important part.

### 2.3 Aberrant methylation

A number of suggestions have been put forward to propose how methylation patterns may become altered naturally. Following replication a methylation pattern is normally copied semi-conservatively by maintenance activity where only the newly synthesised DNA strand is methylated by the enzyme. A few sites escape methylation but may be 'silenced' later by *de novo* activity; where both strands are methylated symmetrically. If however a site fails to be methylated in the new strand, an alteration in methylation pattern will occur. There are several mechanisms by which DNA methylation may become aberrant :

- (i) Inactivation of DNA MTASE.
- (ii) Non-appropriate cofactors for the enzyme.
- (iii) Modification of DNA probably at, or adjacent to, the methylation site.

All the above mechanisms would result from passive demethylation, but this does not correlate with the observations of Wilks *et al.*<sup>34</sup> When studying the induction of the pitellogenin gene by oestrogen in chick liver, the authors saw a rapid demethylation suggesting some active method of losing methyl groups. Razin tried to suggest a mechanism for this reaction.<sup>25</sup> When witnessing the observation of loss of methyl groups from FEL cells following induction, he discounted the possibility that the demethylation could be enzymatically controlled. A process involving the replacement of 5-MC moieties at CpG sequences by cytosine appeared the most likely event.

As discussed earlier the most significant effect of aberrant methylation is an effect on gene expression. Other possible effects of these alterations in methylation pattern include an alteration in the rate at which mutations may occur and an effect on the fidelity of repair mechanisms. Substituting 5-MC for cytosine also will affect the melting temperature of



DNA.<sup>38</sup> The most significant of these other effects was suggested by Holliday in 1975.<sup>27</sup> Holliday proposed that damaged DNA if incorrectly repaired could give rise to a malignant phenotype.

## CHAPTER 3

### Cytosine methylation and cancer

#### 3.1 Introduction

Evidence resulting from studies on methylation of DNA and gene expression has linked hypomethylation and carcinogenesis. If DNA methylation plays an important role in gene expression, then it follows that aberrations within this controlling mechanism may be implicated in abnormal gene expression which causes cancer. This raises the intriguing prospect that a C5 methylating agent may be able to act as an oncogene silencing agent.

Cellular oncogenes have been implicated in the induction of the malignant transformation in some model systems *in vitro* and may be related to malignancies *in vivo* of vertebrate species. A specific target for a molecule capable of methylating the C5 position of cytosine would be the promoter regions of GC rich oncogenic sequences in an effort to silence the cancer related genes.

DNA alkylation and carcinogenesis have long been associated, and a large class of chemotherapeutic agents are known to act by alkylating DNA. Alkylation normally occurs at the N<sup>7</sup> and O<sup>6</sup> positions of guanine residues (see Fig.7) but not at cytosine position C5.

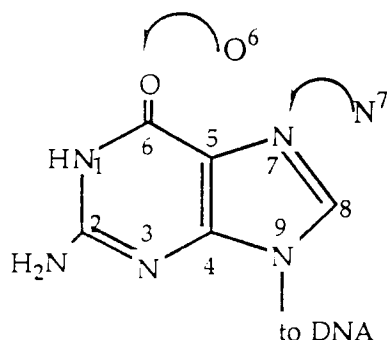


Figure 7 The O<sup>6</sup> position and N<sup>7</sup> position of guanine, potential sites for alkylation

The evidence linking gene expression and C5 methylation led several investigators to examine methylation patterns in transformed cells.<sup>39,40</sup> It was conceived that methylation could play a role in cancer development in two possible ways :

- (i) It may activate particular genes likely to induce transformation of normal cells .
- (ii) It may silence tumour suppressor genes.

A large body of evidence now exists showing that methylation levels and patterns are deranged in tumour cells.<sup>41</sup>

### **3.2 Hypomethylation in tumour tissue**

Most of the studies of hypomethylation in cancer tissue have been conducted using either restriction endonucleases or chromatographic methods, for example HPLC, to detect levels of 5-MC.

There have been relatively few studies of hypomethylation in primary human cancers. Gama Sosa *et al.* demonstrated an approximate 6% overall reduction in the average genomic 5-MC content when comparing a large number of primary malignancies of various types to normal tissues.<sup>42</sup> DNA from metastases in 20 patients showed an even greater average decrease (11%) in their mean 5-MC content relative to that of benign tumours and normal tissues. Feinberg and Vogelstein conducted a similar study using probes made of complimentary DNA (cDNA) from the  $\gamma$  and  $\alpha$ -globin genes to study colon and lung carcinomas.<sup>43</sup> They found substantial hypomethylation in several specific genome

regions, when compared to normal tissues. However hypomethylation in tumours freshly excised from children with a variety of neoplasms is not always observed.<sup>44</sup>

A number of later studies were fuelled by the observation that the cellular proto-oncogenes C-int1, C-myc, C-src, C-fos, and C-Ha-ras are associated with 'islands' rich in GC,<sup>45</sup> and are therefore likely to be controlled by methylation. These studies concentrated on looking at methylation levels of specific oncogenes rather than at the whole genome. Genes of the C-Ha-ras family in colon and lung carcinomas were undermethylated in 6 of 8 tumours when compared to normal mucosa.<sup>43</sup> Hypomethylation was also observed at a specific site in the C-myc gene in human tumour cell lines.<sup>46</sup>

The phenomenon of oncogenic activity being related to cancer is not universal however. Expression of the C-Ha-ras and C-myc genes is known to be enhanced in hepatocellular carcinoma. In restriction enzyme studies results indicated that some CCGG sites within or near the second exon of the C-myc gene are methylated.<sup>47</sup> If hypomethylation is linked to gene silence, this suggests the possibility that the C-myc gene does not play a role in hepatocellular carcinoma.

It is possible that other specific genes may play a role in carcinogenesis. Baylin *et al.*<sup>47</sup> when studying the calcitonin gene found that hypermethylation of this gene is a general property of tumour cells in patients with human lymphoid malignancies and patients with acute non-lymphocytic leukaemias. It was also found that in over 80% of a group of human small cell lung cancers, increased numbers of CCGG sites are methylated in the 5' region. It is possible that methylation at CG is an event characteristic to malignant or transformed cells.

More direct evidence for a link between methylation and cancer has been produced from experiments in which methylation patterns are altered artificially in tumour cells. The experiments suggest a strong causal relationship between methylation changes and increased phenotypic diversity in tumour cells.

### 3.3 Effects of methylation inhibitors on malignancy

5-AzaC, a known methylation inhibitor has proved a useful tool in studying methylation levels in malignancy.

5-AzaC is particularly well suited to studying malignancy, as several studies have shown that the agent is not demonstrably mutagenic in eukaryotic cells.<sup>48</sup> Therefore mutations do not interfere with interpretation of results and direct effects of methylation can be seen. Treatment of murine tumour cells *in vitro* with 5-AzaC can greatly alter the metastatic and tumorigenic phenotypes of the treated cells.<sup>49</sup> The drug is capable of transforming cultured cells and is tumorigenic in whole animals.<sup>50</sup> The *in vivo* administration of 5-AzaC has been shown to induce the expression of certain cellular and viral genes within a variety of tissues.<sup>3</sup> Another methylation inhibitor, ethionine, has been used to study the effect of methylation and differentiation in cancer cells. Ethionine is unique in the fact that it is a competitive inhibitor of the enzyme, DNA MTASE. The drug was able to cause HL60 cells *in vivo* to differentiate within 3 to 5 days of administration,<sup>51</sup> suggesting that inhibition of methylation may lead to the activation of genes capable of inducing differentiation in these cells. The effect seen is not restricted to one cell type. Ethionine is also capable of inducing maturation in FEL cells.<sup>37</sup> Another possible link between methylation and cancer results from studies using chemotherapeutic agents. Several alkylating agents (known carcinogens) have been shown to have an effect on the methylation of a variety of DNA substrates *in vitro*.

### 3.4 Inhibition of DNA methylation by carcinogens

A vital clue establishing a link between methylation and cancer was uncovered when it was found that some known carcinogens actually affected levels of cytosine C5 methylation. Studies by Drahovsky and Wacker,<sup>52</sup> showed that N-methyl-N'-nitro-N-nitrosoguanidine (MNNG) was able to inhibit rat liver DNA MTASE. After looking at a number of carcinogens they found that MNNG was the only one to have an effect on methylation activity levels. On further investigation they found the effect was prevented by dithiothreitol suggesting the enzyme contained an important sulphhydryl residue. It was proposed that enzyme inactivation was most likely caused by MNNG binding to the active site of the enzyme. A similar compound methylnitrosourea (MNU), was later shown to have a similar effect.<sup>53</sup> The effect was observed both *in vivo* and *in vitro*.

Other agents that alkylate DNA at sites other than cytosine residues also possess the ability to inhibit DNA MTASE. 2-N-acetoxyacetylaminofluorene (AcO-AAF) reacts with DNA at guanine to yield N-(deoxyguanosine-8-yl)AAF.<sup>54</sup>

In experiments comparing levels of methylation in the presence and absence of AcO-AAF methylation levels were lower in the presence of the carcinogen. Therefore AcO-AAF inhibits methylation to some degree. A possible explanation of this can be found if one considers that on binding to DNA the enzyme moves along the DNA strand to methylate cytosine residues. On encountering a guanine adduct this movement would be stopped. The degree of inhibition of the enzyme is much less than that produced by MNU and so it can be considered that the compound AcO-AAF has no direct effect on the enzyme.

### 3.5 Methylation, ageing and cancer

Cancer is generally recognised as a disease of old age. Methylation levels may well decrease with ageing. Although no direct relationship has been established between alteration of methylation patterns due to ageing and cancer, it is still worthy of consideration. Cell lines grown in culture exhibit small irregular changes in their pattern of DNA methylation but overall the level remains generally constant.<sup>55</sup> Differentiated cells however steadily lose 5-MC as their proliferation potential declines.<sup>56</sup> The rate of loss can be correlated with the rate of transformation of the malignant phenotype.<sup>57</sup> This phenomenon was observed in all species, and was inversely correlated to life span. The rate of loss of methyl groups was greatest in mice, and smallest in humans. A good consistency was shown between incidence of cancer and rate of loss of methyl groups. Agents that increase the rate of demethylation would be expected to increase transformation rate, 5-AzaC and herpes simplex virus and other carcinogens may act in this manner.<sup>58</sup>

The evidence from a number of pieces of research points towards DNA methylation and malignancy being linked. Whether DNA demethylation at cytosine produces a malignant phenotype or whether 5-MC demethylation occurs as a result of a malignant phenotype being expressed is still unclear. The evidence for the former seems to be stronger. A small molecule capable of methylating cytosine would therefore provide a most valuable tool for the study of this relationship as well as providing a possible chemotherapeutic agent. The rationale behind the synthesis of such a molecule is discussed in the following Chapter.

## CHAPTER 4

### Rationale for synthetic molecules

#### 4.1 Introduction

As mentioned earlier the aim of this thesis is to test the possibility that a small molecule may be able to function as a methylator of cytosine at position C5 in DNA and hence act as an anticancer agent. Although a mechanism has been proposed for the action of the methylating system in nature, little is known about the three dimensional transition state that is thought to exist when methylation occurs. In view of the mechanism discussed in Chapter 1, a molecule with potential C5 methylating ability should have the following important features.

The molecule should possess:

- (i) The ability to enter the nucleus and recognise the DNA target.
- (ii) A site capable of interacting with the C6 position of cytosine resulting in the activation of the C5 position.
- (iii) A methyl donating centre.

Before discussing the possible features in a molecule, a brief overview of other alternative approaches to methylation therapy will be discussed.



## 4.2 Methylation inhibitors

Biological transmethylation reactions which utilise SAM as the methyl donor have already proved a target for drug therapy. There are a number of MTASE enzymes present in the cell, concerned with a variety of functions. Work to date has concentrated on the fact that a general feature of most SAM-dependant MTASE's is that they are inhibited by SAH.

Two general approaches have been employed:

- (i) Synthesis of SAH analogues that inhibit directly a particular MTASE.<sup>59</sup>
- (ii) Synthesis of SAH hydrolase inhibitors which cause an elevation in levels of SAH and subsequently produces inhibition of MTASEs.<sup>60</sup>

Some of the molecules in category (i) have been observed to have oncostatic properties,<sup>61</sup> for example, S-isobutyryladenine. (Fig.8)

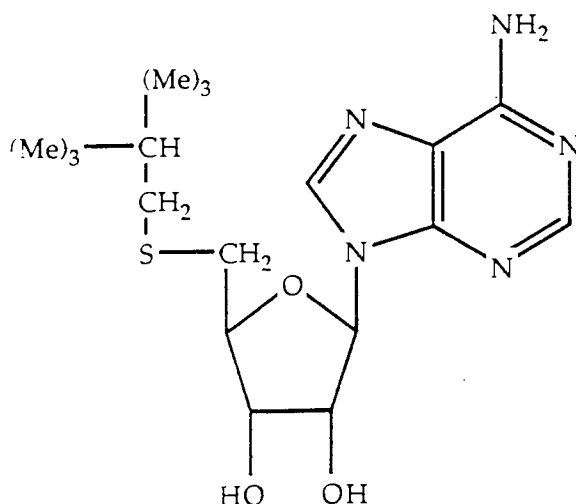


Figure 8 Structure of S-isobutyryladenine: an inhibitor of DNA MTASE

The main targets for these molecules have included catechol O-MTASE, and a number of tRNA MTASEs.<sup>62</sup> There are no reports on the effect of these molecules on DNA methylation.

The idea of trying to produce C5 methylating molecules as potential inhibitors of cancer cell growth is novel and as yet unexplored. In the remainder of this thesis the feasibility of the idea will be addressed.

### 4.3 Potential synthetic targets

Normal electrophilic methylating agents of the type Me-X (X is a suitable leaving group) are unsuitable as methylating agents of carbon nucleophiles in the presence of more strongly nucleophilic oxygen and nitrogen containing moieties. Therefore any molecule required to specifically methylate cytosine at position C5 will require some means of interacting with the C6 position of the pyrimidine ring to render the C5 position susceptible to an incoming electrophile.

The most convenient starting point for approaching this problem would involve attempts to mimic the action of the enzyme. A proposal for the mechanism of activation by intramolecular catalysis suggests that the enzyme acts by initially attacking the C6 position of the cytosine ring. The attack results in an intermediate with high electron density at the C5 of the pyrimidine ring (Fig.9).

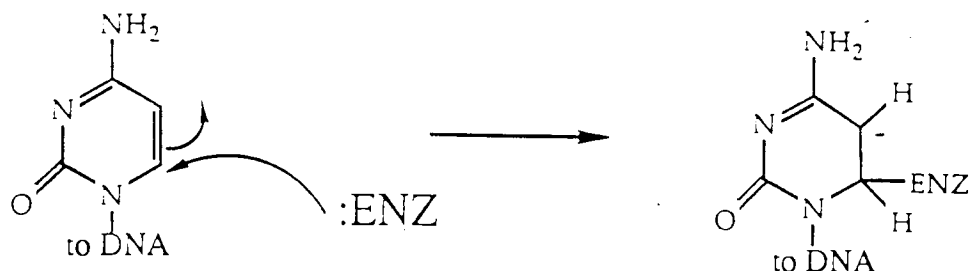


Figure 9 An carbanionic intermediate formed after the reaction of DNA MTASE with the C6 position of cytosine

The only information pertinent to the mode of action of DNA MTASE is that the active site of the enzyme is known to contain an important thiol residue.<sup>63</sup> Therefore a thiol functionality is likely to be a requirement of an active site mimic.

The second important feature of any C5 methylating molecules is the methyl donation fragment. The classical studies of Cantoni on biological transmethylation reactions culminated in the proposal that a sulphonium compound SAM was the active methyl donor;<sup>64</sup> this has since been confirmed. A large amount of research has been centred around the role of SAM in transmethylation reactions. Before adopting a potential candidate for a methyl donating moiety, some of the important points that have emerged about SAM will be discussed.

Kinetic isotope studies on the catalysis of the methyltransferase catechol-O-MTASE revealed that some of the catalytic power of the enzyme was associated with the compression of partial bonds about the transferring methyl group in the transition state.<sup>65</sup> Whether DNA MTASE operates by a similar mechanism is not known. The importance of configuration of the sulphonium centre of SAM for enzyme activity was reported in 1959.<sup>66</sup>

SAM possesses 2 chiral centres, one at the sulphonium moiety and the other at the  $\alpha$ -carbon of the amino acid moiety. Using degradation techniques Carnforth *et al.* isolated S-carboxymethylmethionine from SAM and then by crystallisation separated the diastereoisomers of S-carboxymethylmethionine.<sup>67</sup> X-ray crystallography was used to determine the absolute configuration of the sulphonium centre. The conclusion reached was that SAM possessed the S configuration at the sulphonium centre (Fig.10).

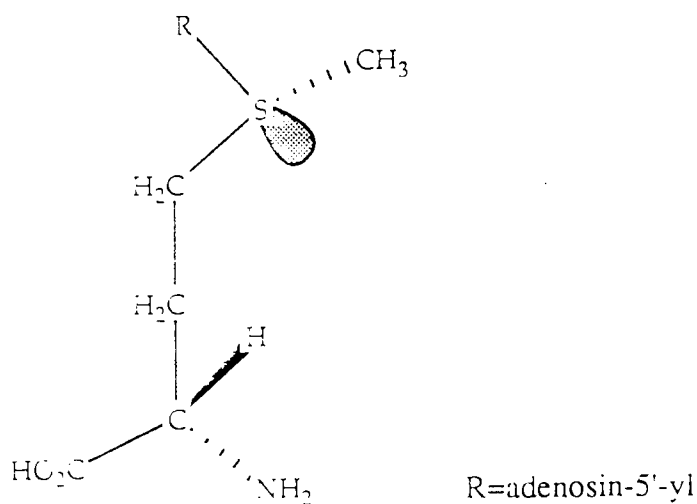


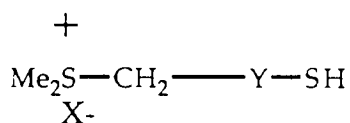
Figure 10 The absolute configuration of the sulphonium centre of S-adenosylmethionine

Coward, using an enzymic model system to look at methylation of  $sp^3$  carbon suggested that the methyl transfer involved in MTASE reactions is a classic  $S_N2$  reaction requiring a linear arrangement of nucleophile, methyl carbon and thioether.<sup>68</sup> He proposed that MTASEs might facilitate attack by providing an area around the active site of very low electron density. The hybridisation state for C5 in the DNA MTASE-2'deoxycytidine complex is thought to lie somewhere between  $sp^2$  and  $sp^3$ , so whether a linear transition state is required is unclear. A linear transition state would be vital if a molecule was required to act in place of SAM as a methyl donor for the enzyme. However a molecule which could methylate cytosine independently would not be constricted by any enzyme binding site and therefore may not require a linear transition state, in order to transfer a methyl group. Despite this conclusion an obvious choice for the methyl donating compound would be a molecule possessing a methyl sulphonium group. To this end sulphonium salts might possibly be able to act as a methyl donor molecule for the enzyme and were therefore potential synthetic targets.

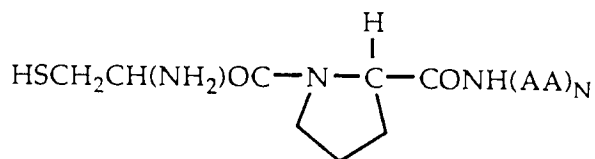
In 1988 the sequence of a eukaryotic DNA MTASE was published.<sup>69</sup> The sequence of the mammalian DNA MTASE contained three Pro-Cys sequences within its length. Peptidic molecules based around the conserved Pro-Cys sequence motif, which is known to be present in the active site of prokaryotic DNA MTASEs would provide another synthetic target area. These peptides may be able to act as C5 position activating molecules.

A third target group of molecules exist which would possess both activating and methylating properties. The methylating fragment of the molecule would be provided by a sulphonium salt. The activating portion of the molecule would be provided by a thiol containing molecule, possibly peptidic. Target molecules of this group would be expected

to methylate the C5 position of cytosine in the absence of the enzyme and any natural methyl donor. The four target groups of molecules are shown in Fig. 11.

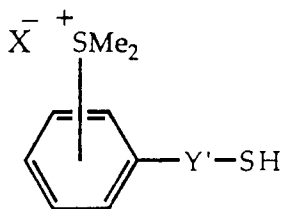


(A)

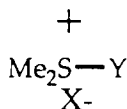


(B)

AA = any amino acid



(C)



(D)

Y and Y' = any peptidic molecule

(A)+(C)---- Activating/methylating molecules

(B)---- A Pro-Cys mimic

(D)---- A simple sulphonium salt.

Figure 11 General structures proposed for target molecules.

## **Section 2: Results and discussion**

## CHAPTER 5

### Restriction enzyme studies on simple sulphonium compounds

#### 5.1 Experiments to determine the sites of alkylation in DNA of simple sulphonium compounds using restriction enzymes

An initial interest in the mechanism of methylation of DNA MTASE prompted some experiments to investigate whether it would be possible to replace the natural methyl donor, SAM, by another simpler sulphonium salt. The results from these experiments are presented below and were the fuel behind the hypothesis that the enzyme methylating system could actually be completely replaced by a smaller molecule. These preliminary studies were carried out using the isochizomeric pair of restriction enzymes HpaII and MspI (Table 2).

Table 2 Cleavage sites of MspI and HpaII

Restriction endonuclease	Cleavage sites
MspI <sup>a</sup>	↓ ↓ C.mC.G.G and C.C.G.G.
HpaII <sup>a</sup>	↓ C.C.G.G

a) Neither endonuclease cleaves mCCGG



Experiments were carried out on  $\lambda$ -phage DNA to investigate the ability of simple sulphonium compounds to act as C5 methylating agents. Compounds were tested both in the presence and the absence of the enzyme DNA MTASE. Following restriction enzyme digestion the DNA fragments were separated by agarose gel electrophoresis. Electrophoresis separates the DNA fragments in order of molecular weight, lightest fragments show bands at the top of the gel, heavier fragments appear towards the bottom of the gel (Fig.12).

Direction of electric current

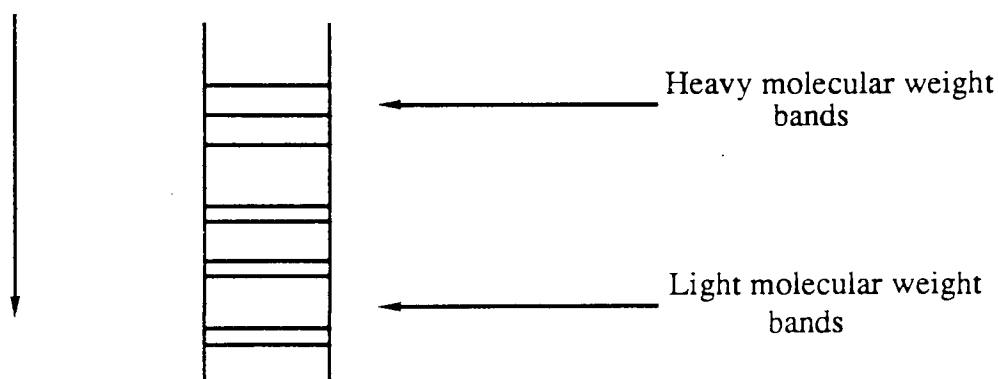
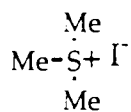


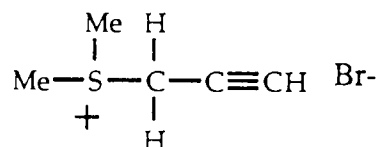
Figure 12 The principles behind separation by agarose gel electrophoresis

The complementary nature of the two restriction enzymes allows detection of sites of methylation by comparing the range of fragments produced by both HpaI and MspII on a particular DNA.  $\lambda$ -phage DNA was the chosen substrate as it has a relatively low molecular weight ( $32 \times 10^6$  da) and also possesses relatively low levels of 5-MC compared to other DNA's eg. calf thymus DNA.

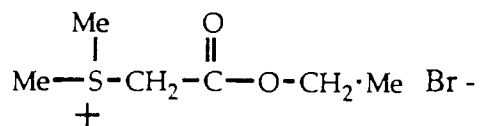
As introduced earlier, investigation of the methylating ability of some simple sulphonium compounds (Fig.13) was carried out using the restriction enzyme isochizomeric pair MspI and HpaII which cleave the sequence CCGG depending on whether or not the internal



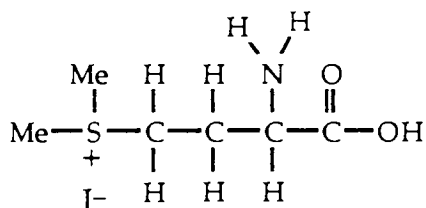
Trimethylsulphonium iodide (1)



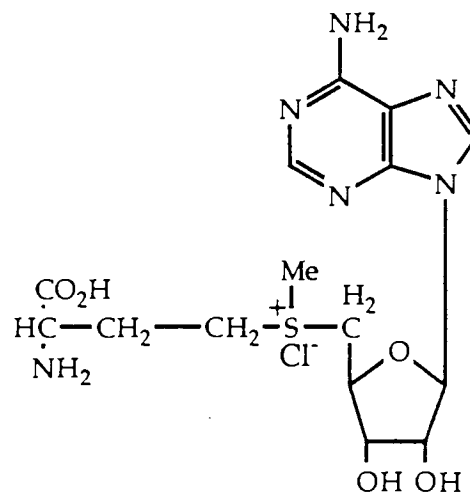
Dimethylpropargylsulphonium bromide (2)



Carboethoxymethyldimethyl sulphonium bromide (3)



S-methylmethionine sulphonium iodide (4)



S-adenosylmethionine chloride

Figure 13 Structures of the simple sulphonium compounds used in the restriction enzyme studies

cytosine is methylated. Agarose gel electrophoresis of the resulting fragmented DNA was used to determine whether any methylation had occurred. The sulphonium salts were incubated with  $\lambda$ -phage DNA, a DNA containing 48,502 base pairs. Initially the most simple sulphonium compound trimethylsulphonium iodide (**1**) was incubated with  $\lambda$ -phage DNA. The resulting gel pattern showed that as the incubation time was increased, the restriction fragments increased in molecular weight. This effect was only observed at high concentrations of trimethylsulphonium iodide. For incubations of greater than 0.5 hrs, an increase in molecular weight of the fragments was seen for both restriction enzymes and no new high molecular weight bands were seen for MspI when compared to the pattern produced by HpaII. From this it was concluded that trimethylsulphonium iodide cannot methylate cytosine at position C5.

For the sulphonium compounds dimethylpropargylsulphonium bromide (**2**) and carboethoxymethyl dimethyl sulphonium bromide (**3**) only restriction analysis by HpaII was carried out. HpaII is sensitive to methylation at both cytosines in the CCGG recognition sequence and therefore would be sufficient to detect whether any C5 methylation at this site had occurred.

Both compounds were analysed in the presence and absence of DNA MTASE to determine whether they can play a role as substrates for the enzyme or can act to methylate cytosine independently. For **2** an increase in the molecular weight of restriction enzyme fragments was seen at 75 millimolar (mM) in the presence of DNA MTASE. At a concentration of 100mM both in the presence and absence of DNA MTASE at least 3 new bands of higher molecular weight were detected. For **3** at corresponding concentrations no new bands were detected. Concentrations of 0.5molar (M) and above were required before any heavy molecular weight fragments were detected on the agarose gel. From these results it was concluded that the two sulphonium compounds assayed were not effective methylators of

cytosine at the C5 position. Methylation did occur at high concentrations of the compounds but whether the assay was reliable at such high concentrations was not clear.

The only amino acid derived sulphonium compound assayed by restriction enzyme analysis was S-methylmethionine sulphonium iodide (4). This compound produced some interesting results (Fig.14). At concentrations of 50mM and 75mM, a single band of higher molecular weight appears when in the presence of enzyme but not when the enzyme is absent. From the postulated method of enzyme binding by SAM,<sup>70</sup> it was not unfeasible that S-methylmethionine sulphonium iodide may bind to the same enzyme pocket and therefore act as a weak substrate for DNA MTASE (Fig.14).

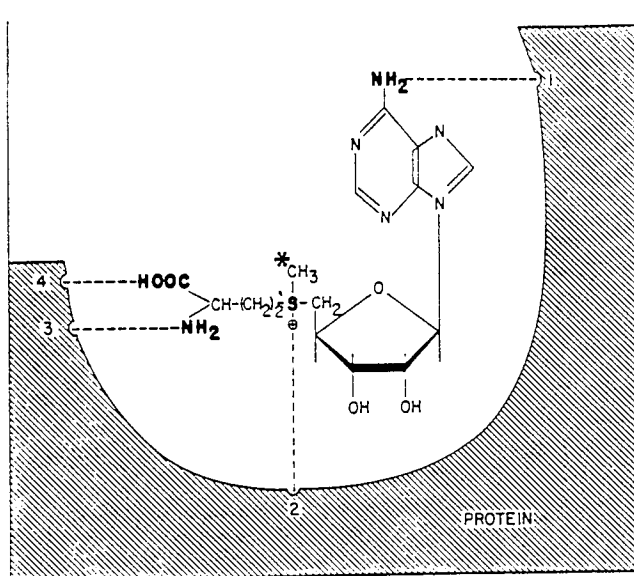


Figure 14 Proposed binding pocket for SAM in DNA MTASE

This result suggested the possibility that a molecule based upon S-methylmethionine sulphonium iodide may well be able to act as a cytosine C5 methylating molecule. At this stage the only information known about the activating mechanism of DNA MTASE was that the enzyme's activity was highly dependent on a cysteine thiol residue present in the active site. Considering both these pieces of information, the idea was suggested that

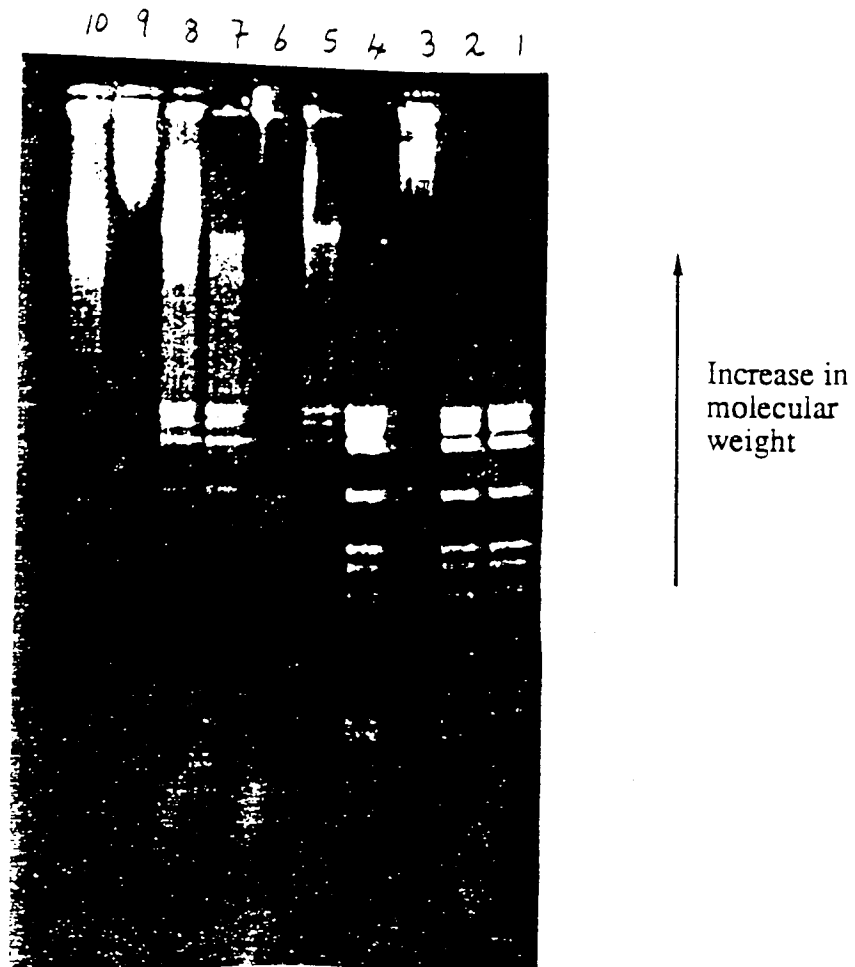


Figure 15 Restriction enzyme cutting pattern for S-methylmethionine sulphonium iodide treated  $\lambda$ -phage DNA

Incubation with drug was for 0.5hrs.

Lane: 1) control MspI; 2) control HpaII; 3) 12.5mM + MTASE; 4) 12.5mM, HpaII; 5) 12.5mM, HpaII + MTASE; 6) 25mM, MspI + MTASE; 7) 25mM HpaII; 8) 25mM, HpaII + MTASE; 9) 50mM + MTASE; 10) 75mM + MTASE

DNA MTASE was capable of transferring 0.4pmol of methyl groups under the standard assay conditions.

perhaps a sulphonium molecule based upon the dipeptide Met-Cys would provide a suitable first synthetic target.

## CHAPTER 6

### Molecular modelling studies

#### 6.1 Introduction

Little structural information is known about the enzyme DNA MTASE and its mechanism, however the three-dimensional structure of the enzyme is considered to be crucial. A smaller peptidic molecule capable of mimicking the MTASE's activation mechanism would be thought to require a similar structure to the active site of DNA MTASE. The Pro-Cys sequence highly conserved in the active site of a number of prokaryotic MTASEs may play a critical role in determining the structure of the active site region and hence was an important target. The structural conformation the sequence produces may render the cysteine thiol group more reactive.

In an effort to more clearly understand how the molecular features of the active site region of DNA MTASE facilitate the opening of the C5-C6 double bond in cytosine and to determine the degree the dipeptide Pro-Cys influences this structure, some modelling studies were undertaken. These studies would provide valuable information as to whether target peptidic molecules could adopt the conformation required to react with cytosine.

Molecular modelling acts as a complementary technique to X-ray crystallography in studying the conformation of potentially biologically active molecules. X-ray crystallography provides a representation of a molecule in its lowest energy state in a solid crystal. In a biological situation i.e. an aqueous solution at pH 7.4 and temperature of 37<sup>0</sup>C, the molecule may well adopt different conformations dependent on solvent interactions. Using the theoretical calculations of molecular modelling along with crystal

structure data, it may be possible to obtain a more accurate representation of the biological situation.

At this point it must also be noted that theoretical calculations predict properties for molecules in given ideal situations and therefore application of these calculations must be carefully considered for specific problems.

## 6.2 Modelling methods

### 6.2.1 Chem-X

Chem-X is a suite of programs produced by Chemical Design, Oxford,<sup>71</sup> which allows the construction and manipulation of molecules on which a variety of chemical calculations can be performed, for example, energy minimisation calculations. Molecules may be displayed in a variety of formats. Simple molecular mechanics calculations are able to define energy contour maps with respect to individual parameters allowing a basic conformational analysis of individual molecules to be performed. In the studies to be described, Chem-X was used to build and manipulate peptidic structures. Chem-X manipulations and calculations were performed on a cluster of two VAX 11/8650 computers. Within Chem-X the molecular orbital program MOPAC 5.0 can be run.<sup>72</sup> MOPAC 5.0 was used for calculating low energy conformations of molecules.

### 6.2.2 QUANTA

QUANTA is a program produced by the Polygen corporation<sup>73</sup> and was run on an Iris 3130 Unix workstation produced by Silicon Graphics. Structures can be built within



QUANTA or imported from other molecular modelling programs. The program allows dynamic display of up to 4500 atoms at one time and also the docking of two or more molecules together, which allows the visualisation of substrate-receptor interactions. Docking experiments are facilitated by the ability of the program to display interaction energies between atoms in real time. Within QUANTA itself, molecular distances between the two molecules to be docked can be displayed and the program also possesses the ability to identify possible hydrogen bonds that may be likely to form between the structures. The program possesses a library of molecules, in particular amino acids, and a facility which allows generation of a peptide backbone from individual amino acids. Docking studies were carried out using QUANTA. The program is also able to perform energy minimisation calculations but lacks the power to deal with large molecules effectively therefore all calculations were undertaken in Chem-X, which is able to use the power of a main frame computer.

### 6.2.3 Modelling Data.

The data used for these modelling studies was obtained from BIPED (Brookhaven integrated protein engineering database), the Cambridge crystallographic database and from our own data (see Appendix 1). Other structures were built and modified within the Chem-X program, using amino acid data contained within the Chem-X program.

### 6.3 Conformational studies on the dipeptide Pro-Cys

Small peptidic molecules when in solution may take up a large number of conformations. In larger peptides the steric bulk of the molecules act to restrict rotation about amide bonds and therefore much more rigid structures are formed. Using X-ray crystallography, molecular modelling and protein structure prediction programs it may be possible to determine whether smaller peptides can adopt the conformation of similar sequences present in much larger molecules.

The objectives of these studies were as follows:

- (i) Firstly to generate a structure in Chem-X as close as possible to the actual minimum energy structure of the dipeptide Pro-Cys. The minimised structure could then be docked into a GC base pair to show whether interaction of the two molecules was likely.
- (ii) The second objective was to compare the phi and psi angles of the crystal structure of N-benzyloxycarbonyl(CBZ)-prolyl-S-benzyl(Bzl)cysteine methyl ester (**5**) with the phi and psi angles of Pro-Cys sequences present in a number of randomly selected proteins from the Cambridge crystallographic data base. The aim of this study was to determine whether the bulky protecting groups were capable of constraining the protected peptide in a similar manner to that seen in large proteins.
- (iii) The third objective was to study the homology of structure exhibited by sequences containing the Pro-Cys motif from a number of different proteins.

### 6.3.1 Minimum energy conformation of Pro-Cys

Two possible approaches were considered to start to model the lowest energy conformation of the dipeptide Pro-Cys. The first approach was to build and manipulate the structure entirely within Chem-X. The second approach used a crystal structure solved at Aston of the protected dipeptide N-CBZ-prolyl-S-(Bzl)cysteine methyl ester (**5**) (Plate 1) (Appendix 1) and then manipulate the structure using Chem-X. The first approach involved using the crystal structures of the free amino acids proline and cysteine obtained from the Cambridge crystallographic database. These were linked within Chem-X by an amide bond with the phi and psi angles being equivalent to those seen in a standard *cis* conformation (Fig.16). Using the MNDO (moderate neglect of differential overlap) parameter set within the semi-empirical molecular orbital program MOPAC 5.0, the constructed dipeptide was energy minimised to give a structure **A** (Plate 2). The second approach involved the crystal structure data of **5**. The co-ordinates from the crystal structure were imported into Chem-X. The protecting groups were then removed using computer manipulation techniques, necessitating the addition of a hydroxyl and an amino functional group to generate a structure **B** (Plate 2). Structures **A** and **B** were then compared directly by transferring the energy minimised co-ordinates into QUANTA and overlaying the two structures (Plate 2). The amide backbones of the two structures were overlaid as closely as possible which revealed a high degree of structural overlap. The similarity between the two energy minimised structures suggests that the computed conformations are likely to be approaching the global minimum of the free dipeptide, rather than being two distinct local minima. Either of these structures appeared suitable to be used as a minimum energy conformation for **5**. Structure **B** was chosen as representing the minimum energy conformation, as this structure was actually built from the crystal structure of **5**, rather than being completely computer generated

**CIS AMIDE**

**TRANS AMIDE**

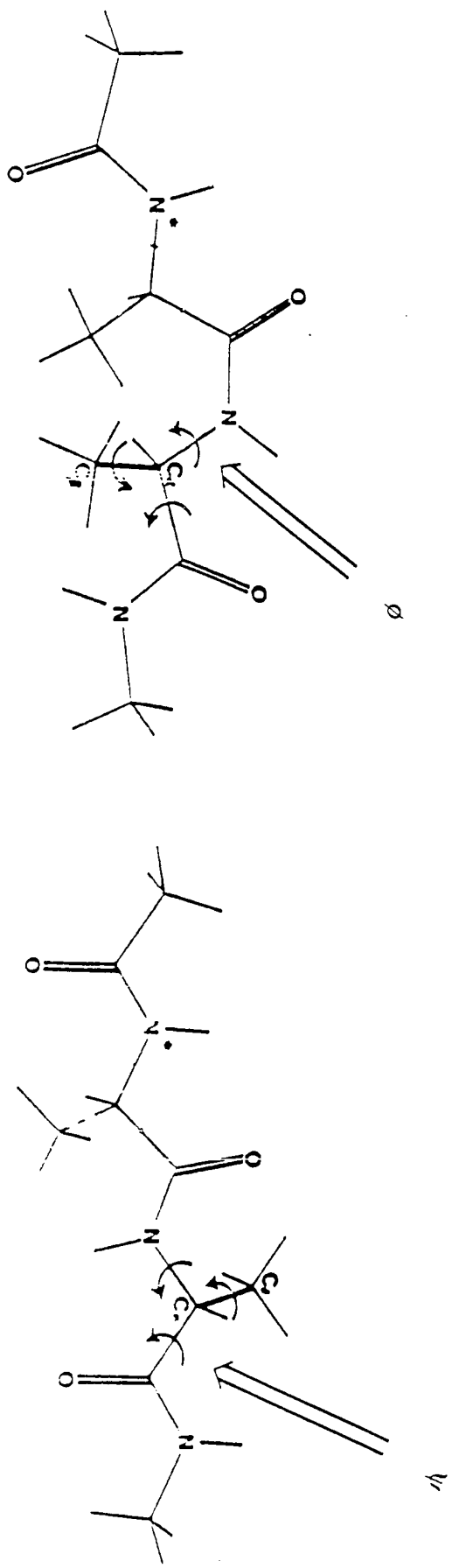


Figure 16 Standard *cis* and *trans* amide bonds showing phi ( $\phi$ ) and psi ( $\psi$ ) torsion angles

**Content has been removed for copyright reasons**

Plate 1 Crystal structure of N-CBZ-prolyl-S-(Bzl)cysteine methyl ester (5)

**Content has been removed for copyright reasons**

Plate 2 Comparison of a computer generated energy minimised structure of prolylcysteine (A) with an energy minimised structure generated from the crystal structure of N-CBZ-prolyl-S-(Bzl)cysteine methyl ester (B)

### 6.3.2 Interaction of the energy minimised Pro-Cys dipeptide with a GC base pair

Present opinion on the mechanism of methylation by DNA MTASE suggests the involvement of a cysteinyl thiol residue within the active site of the enzyme which is thought to interact with cytosine C6. The close proximity of the thiol and C6 of cytosine was therefore considered vital for any reaction to occur.

The aim of this study was to see whether the energy minimised structure of the Pro-Cys dipeptide would possess a conformation that would enable it to dock into a GC base pairing in such a manner as to allow the cysteine thiol to be in close proximity to the C6 position of cytosine.

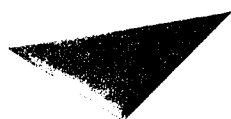
Within QUANTA an energy minimised structure of a GC base pair was used as the docking target. The Pro-Cys minimised structure **B** was docked so that the cysteine thiol of the enzyme and the C6 of cytosine were as close as possible without any overlap of Van der Waals radii of each of the atoms. From the docked structure (Plate 3) it can be seen that there is the possibility of hydrogen bonds forming which may be able to stabilise any structure required for the activation reaction. Some more powerful initial binding of the dipeptide to the DNA would also be required. This suggests that the energy minimised structure produced within Chem-X may be able to bind to cytosine within DNA and hence act to facilitate the opening the C5-C6 double bond of cytosine. The proposed hydrogen bonds would be sufficient to help to stabilise any docked structure once formed, but the dipeptide structure as it stands would not have a particularly high affinity for a GC base pair. Binding affinity could be improved by incorporating some charged features into the molecule or by using peptides of greater length thereby increasing the number of hydrogen bonding contacts.



Aston University

**Content has been removed for copyright reasons**

Plate 3 Docking of an energy minimised structure of prolylcysteine  
into a GC base pair



Aston University

**Content has been removed for copyright reasons**

Plate 4 Comparison of an energy minimised structure of prolylcysteine  
with the crystal structure of N-CBZ-prolyl-S-(Bzl) methyl ester (5)

### 6.3.3 Structural comparisons between a protected Pro-Cys dipeptide and Pro-Cys sequences when present in proteins.

The energy minimised structure **A** (**A** was used rather than **B** to ensure at least some conformational difference in the structures existed) was compared with the crystal structure of the protected dipeptide (Plate 4), to examine whether any structural similarities exist. The crystal structure was not energy minimised using MOPAC 5.0. Overlap of these two structures shows that the protecting groups may play a role in altering the dipeptide structure, the proline rings are flipped in opposite directions and the cysteine side chains protrude in differing directions.

Although the difference in conformation may be due to crystal packing forces, the possibility existed that the bulky protecting groups were significantly altering the conformation and hence may simulate the conformations shown by a Pro-Cys sequence when in proteins. The most practical approach to this investigation was to carry out a comparison of the phi and psi angles of the Pro-Cys dipeptide from the crystal structure with the phi and psi angles from a number of proteins which contain Pro-Cys sequences and for which the crystal structures are known. This comparison could most easily be carried out by the use of Ramachandran maps.<sup>74</sup>

Phi and psi angle information for proteins was extracted from the BIPED database. (The database is able to provide secondary structural information for all the protein crystal structures held on the Brookhaven protein databank). Initially the information from ten selected proteins was plotted on a Ramachandran map together with the phi and psi angle information for **5** (Fig.17).



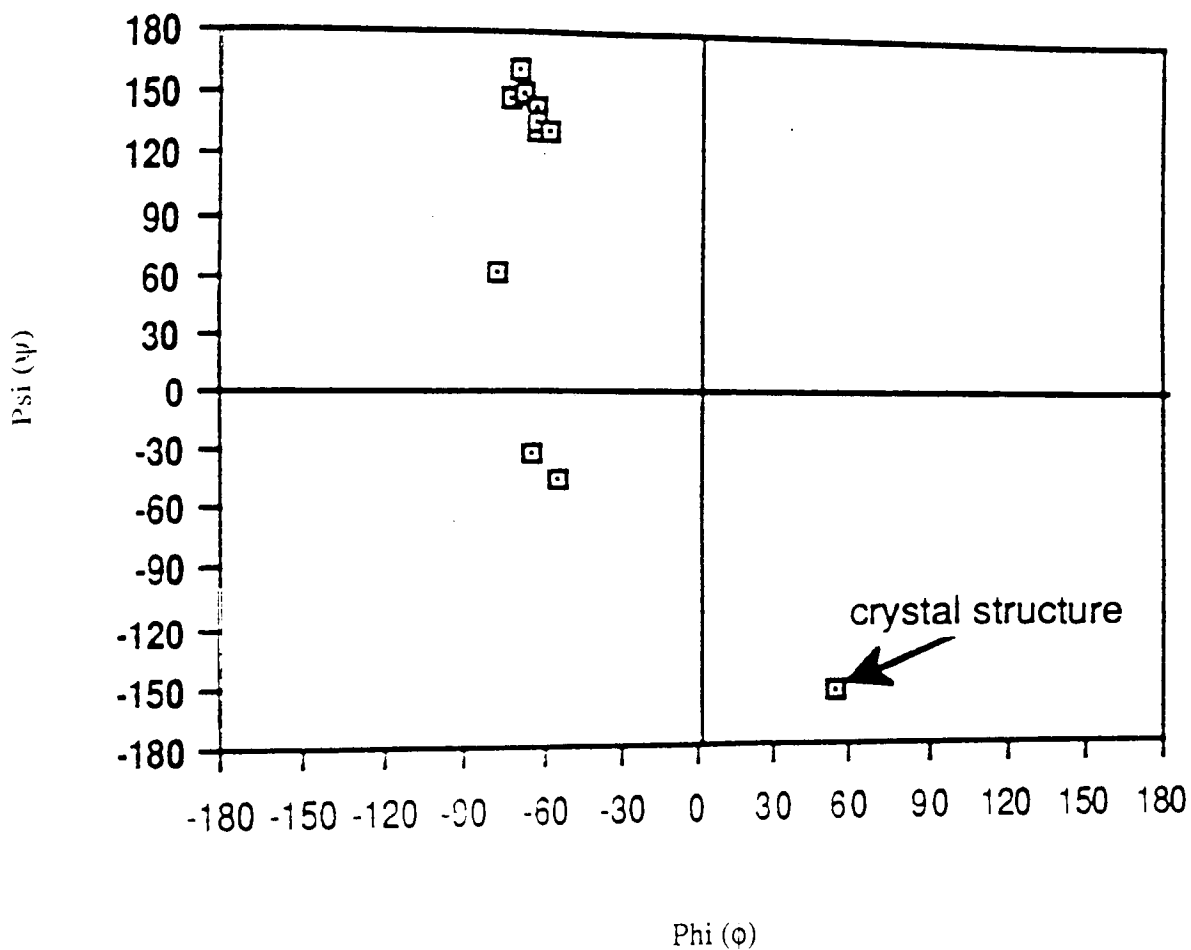


Figure 17 A Ramachandran map comparing the phi and psi angles for proline from the crystal structure of N-CBZ-prolyl-S-(Bzl)cysteine methyl ester (5) with the phi and psi angles for proline from 10 proteins selected from the BIPED database

From the torsion angle map of proline it can be seen that little structural similarity exists between the protected dipeptide and the Pro-Cys sequences present in the protein structures selected. The Pro-Cys sequences present in the proteins show good homology, the majority being found in the top left hand quadrant of the Ramachandran map. The protected dipeptide however shows a very different set of torsion angles and is found in the lower region of the bottom right hand quadrant. The cysteine torsion angle map showed a slightly closer relationship (Fig.18). Nine of the proteins appear very closely together in the top left hand corner of the Ramachandran map again showing good structural homology. The crystal structure appears in the top right-hand quadrant. Considering both maps it can be seen that the amide bond in the crystal structure of **5** is very differently constrained to that seen in the proteins chosen. Therefore it can be concluded that the attachment of sterically bulky groups to the dipeptide influences the conformation of the amide backbone in a different manner to that seen when the dipeptide sequence is present within the bulk of a protein. The fact that a number of proteins extracted from the BIPED database appeared to show homology of structure supported the proposal that similar amino acid sequences in different proteins are likely to adopt similar structures.

This suggested that the only way to mimic the active site of DNA MTASE would be to use a larger peptidic molecule.

As an extension to the above studies, all the proteins on the BIPED database with a Pro-Cys within their sequences were identified. Eighty separate records were identified: these could be associated with 19 different proteins, of which 10 were identified as proteases. No information was available for any MTASE-type enzymes on this database, therefore a more direct comparison between the structure of DNA MTASE and that of the Pro-Cys dipeptide could not be performed. It was not possible to ascertain from the database the location of the Pro-Cys sequences within the proteins, and consequently whether or not the

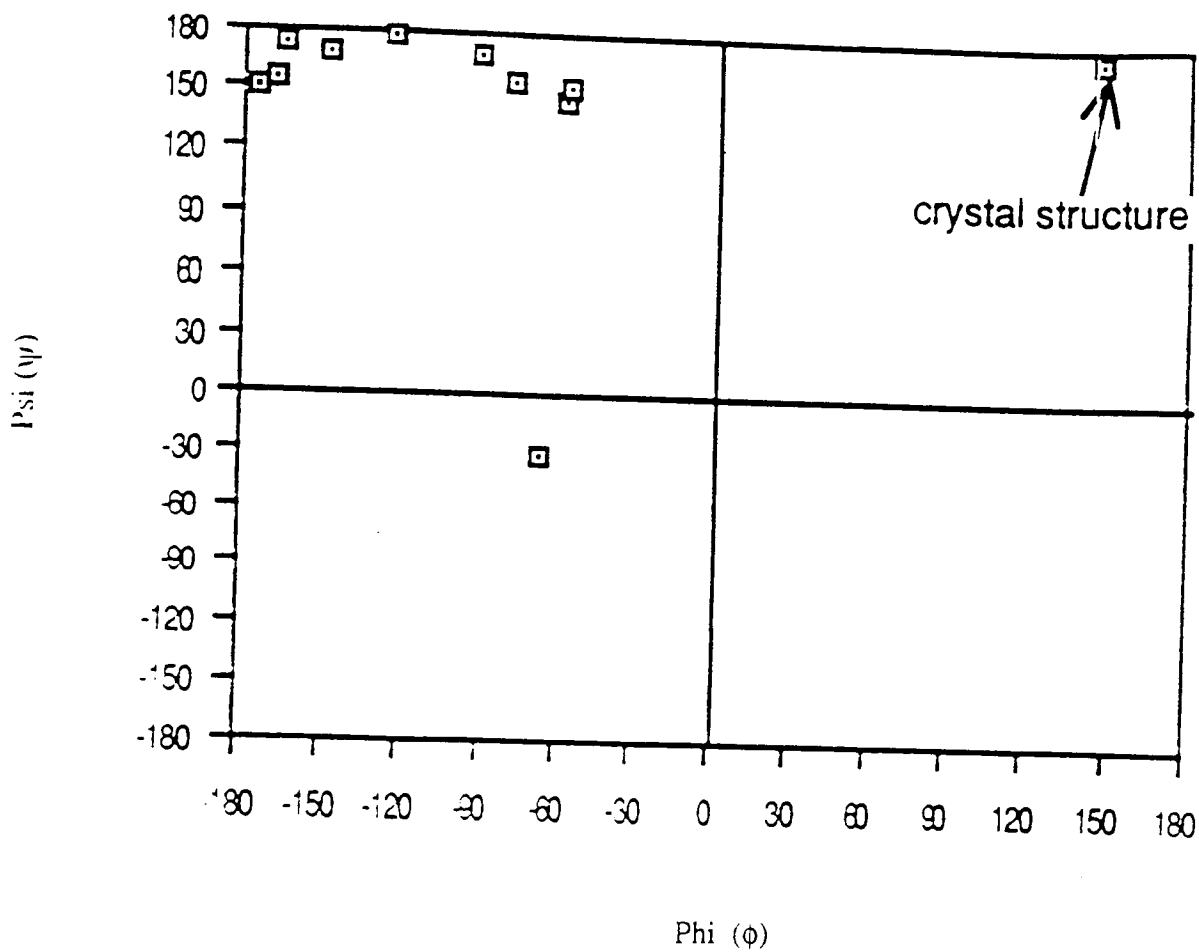


Figure 18 A Ramachandran map comparing the phi and psi angles for cysteine from the crystal structure of N-CBZ-prolyl-S-(Bzl)cysteine methyl ester (5) with the phi and psi angles for cysteine from 10 proteins selected from the BIPED database

cysteine residues identified were actually associated with any specific mechanistic function within the proteins. The coordinates of four of the proteins with similar activities (all proteases), were randomly selected from the ten identified and were used for the following investigation (Table 3).

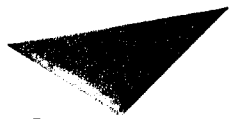
A tetrapeptide sequence containing the specific Pro-Cys sequences was extracted from each of the proteins. The tetrapeptide structures were then superimposed on each other within QUANTA. Examination of the amide backbone revealed a high degree of overlap between the four structures (Plate 5). Such studies were unable to evaluate whether similar three dimensional structures within peptides correlated to similar biological activities, but provide further evidence to suggest that amino acid homology can be related to structural homology.

Table 3 Phi and psi angles of four protease enzymes identified from the Brookhaven database containing a proline residue adjacent to a cysteine residue

Sequence No.	Amino Acids	Protein	Phi	Psi
55	Pro-Cys	Actinidin <sup>a</sup>	-72.3	148.8
198	Pro-Cys	$\alpha$ -Chymotrypsin <sup>b</sup>	-114.9	152.4
126	Pro-Cys	Elastase <sup>b</sup>	-71.6	118.1
152	Pro-Cys	Papain <sup>a</sup>	-69.2	161.7

a) Sulphydryl protease

b) Serine protease



Aston University

**Content has been removed for copyright reasons**

Plate 5 Comparison of 4 tetra-peptide sequences from proteins extracted  
from the Brookhaven database



The homocysteine portion of the molecule was then replaced in the model by the crystal structure of S-methylmethionine sulphonium chloride obtained from the Cambridge crystallographic database.<sup>76</sup>

These molecules have very similar conformations apart from the degree of rotation about the C5-C6 bond.

Manipulations were carried out within Chem-X. The two molecules SAH and S-methylmethionine sulphonium chloride were overlaid superimposing appropriate main chain carbon atoms so the two structures as a whole resembled SAM : then the section was deleted (Fig.20). The resulting structure possessed a number of short intramolecular contacts principally involving the C9 methyl group. The structure was optimised using the MM2 molecular mechanics program to give a minimum energy structure. The conformation of the ribose moiety in the molecule is likely to be dependent upon the relative positions of the attached groups. The bond between SP1 and C38 was considered to be the most likely bond about which any free rotation would occur, along with the torsion angle about C5-C6 which may possibly vary when rotation occurs about SP1-C38. Considering this a conformational analysis was performed in Chem-X with torsion angles about SP1-C38 and C5-C6 as variables.

The energy calculations took into consideration a number of factors for each possible minimum energy conformation, the sum of the Van der Waal's contacts, any electrostatic attractions and repulsions that may occur and any torsional strain that may exist when a particular conformation forms. Appropriate torsional barriers for each bond were input into the program and the atom SP1, the sulphur atom of the sulphonium centre, was given a formal charge of +1. The structure was then given a full molecular mechanics optimisation in MM2 ( Method B, Table 4).

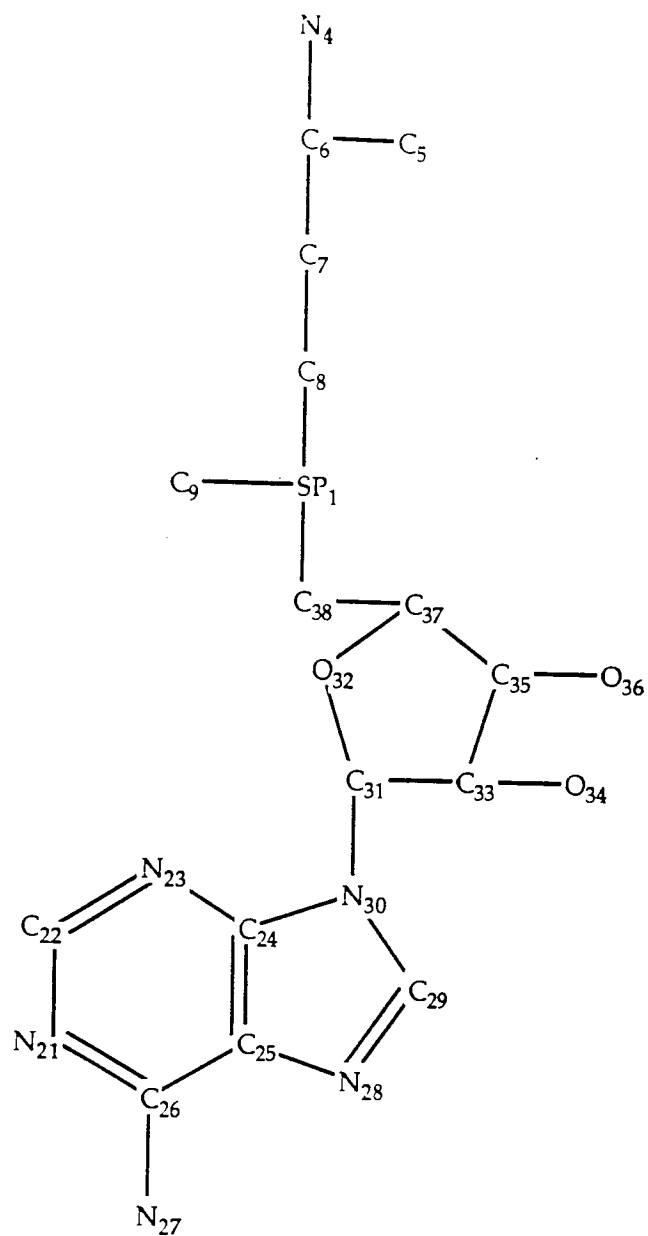


Figure 20 Atomic numbering scheme for the molecular model of S-adenosylmethionine



Table 4 Final energies and relevant torsion angles of a computer generated structure of SAM

Computation Method	Torsion angles <sup>a</sup>				Energies <sup>b</sup>
	1 <sup>c</sup>	2 <sup>d</sup>	3 <sup>e</sup>	4 <sup>f</sup>	
A	-10.7	117.0	-61.9	-170.6	14.85
B	-10.0	127.7	-65.6	-166.4	12.57
C	-9.5	126.8	25.6	-166.3	15.55
D	-9.8	131.0	175.1	-164.3	14.17

a) Torsion angles in degrees

b) Energies in Kcal mol<sup>-1</sup>

c) O3-C5-C6-N4

d) C9-SP1-C38-C37

e) SP1-C38-C37-O32

f) C24-N30-C31-O32

Further manipulations of the structure involved rotation about the bonds C37-C38 and N30-C31. The resultant minimum energy structure optimised in MM2 (Method C, Table 4) was of a higher energy than the two earlier conformations examined. One interesting feature of this structure was the change in the angle SP1-C38-C37-O32. A plot of potential energy against this torsion angle was obtained using Chem-X which showed a minimum at approximately 25°. A second slightly higher energy minimum was seen at around 180° possessing an energy value of 14.17 Kcalmol<sup>-1</sup>. These data indicated that the most

energetically favourable structure results from the crystal structure combination of SAH and S-methylmethionine sulphonium chloride with only a slight twist in torsion angles.

This study considered only a few parameters, further manipulations may be able to identify a lower energy structure. The information from this study showed that SAM would be unlikely to possess an unusual structural conformation which may aid its role as a methyl-donating molecule and that a compound such as S-methylmethionine sulphonium chloride or another similar compound may be able to replace SAM as a methyl donor for DNA MTASE.

## **6.5 Protein structure prediction**

### **6.5.1 Introduction**

The molecular modelling studies suggested, as expected, that larger peptides (ie greater than 4 amino acids) would be required to mimic the active site structure of DNA MTASE. As well as using modelling methods and energy minimisation programs to predict structural characteristics of molecules, mathematical methods have also been developed specifically to deal with secondary structure predictions of peptides and proteins. Using these methods it was possible to determine whether the larger target peptides identified were likely to adopt a similar structure to that exhibited by active sites of prokaryotic and eukaryotic DNA MTASE enzymes.

There are two methods which are commonly used to predict protein structures:

- i) The Chou-Fasman method
- ii) The Garnier-Osguthorpe-Robson method.

The Chou-Fasman method was introduced first in 1974.<sup>77</sup> Using statistical analysis, Chou-Fasman developed a system that rated the potential of all 20 amino acids to form helix or  $\beta$ -sheet conformations in an hierarchical order. From this order a set of empirical rules was devised governing the folding of secondary structural regions in proteins. This method was improved on by Garnier, Osguthorpe and Robson in 1978.<sup>78</sup> They employed a method, again based on statistical analysis, but each sequence analysed was regarded as being part of a larger sequence, so the prediction is based on the amino acid sequence of interest incorporated within a larger statistically predicted sequence of amino acids. The flanking amino acids are suggested by the program from statistical analysis considering the appearance of the sequence in a large number of proteins. Using this method a more accurate view of the actual peptide or protein structure is thought to be achieved.

Both of these methods are employed in the PROTYLZE program. This program is able to produce both graphical and numerical probability data for the structure a particular sequence may adopt along with a crude diagrammatic representation of the likely structure the peptide or protein will display.

Using the PROTYLZE structure prediction program both Chou-Fasman and Garnier-Osguthorpe-Robson prediction calculations were carried out on a number of target peptide sequences identified and on the active site regions of a number of bacterial MTASE enzymes (Table 5). A disadvantage with such protein structure prediction programs is their limited accuracy particularly when looking at very short sequences.

Table 5 Amino acid sequences from bacterial MTASEs on which protein structure prediction calculations were performed.

Bacterial MTASE	Amino acid sequence
EcoRII	Gly-Phe-Pro-Cys-Gln-Pro-Phe-Ser-Leu-Ala-Gly
MspI	Gly-Phe-Pro-Cys-Gln-Pro-Phe-Ser-His-Ile-Gly
BspI	Gly-Phe-Pro-Cys-Pro-Gly-Phe-Ser-Glu-Ala-Gly

### 6.5.2 Results of predictions and discussion

The objectives of this study were to determine whether target peptides may show secondary structural characteristics similar to those exhibited by the active site regions of prokaryotic DNA MTASE enzymes and also to examine whether a similar structure appears in the mammalian DNA MTASE.

From examining the amino acid sequence of mouse mammalian MTASE and the amino acid sequences of the active site regions of a number of prokaryotic MTASE enzymes, the sequences of three target peptides were identified:

<sup>1</sup>Gly-Phe-Pro-Cys-Pro-His-Phe-Ser<sup>8</sup> (53)

<sup>1</sup>Cys-Ala-Pro-Cys-Leu-Pro-Ala-Gly<sup>8</sup> (54)

<sup>1</sup>Ser-Pro-Cys-Met-Arg-Leu-Ile-His-Leu<sup>9</sup> (55)

All three peptides were subjected to Chou-Fasman and Garnier-Osguthorpe-Robson analysis. The peptide **53** was predicted to be a highly turned structure by both prediction methods. The Chou-Fasman method predicted two turns between Phe (2) and His (6). The Garnier-Osguthorpe-Robson method predicted a turn at Gly (1) and another between Pro (5) and Ser (8) (Fig.21). These results suggest it is highly likely that the cysteine residue will be adjacent to a turn in the structure.

For the peptide **54** the Garnier-Osguthorpe-Robson method predicted that the structure would adopt a sheet-like structure containing 1 or 2 turns (Fig.22). The peptide is probably far too small to adopt a sheet-like structure and therefore only a turn in the peptide would be structurally significant. The Chou-Fasman method predicted a turn structure particularly between Ala (2) to Leu (5) ie. around the Cys residue (Fig.22).

The structure prediction for the peptide **55** again showed evidence of a turn being present in the sequence.

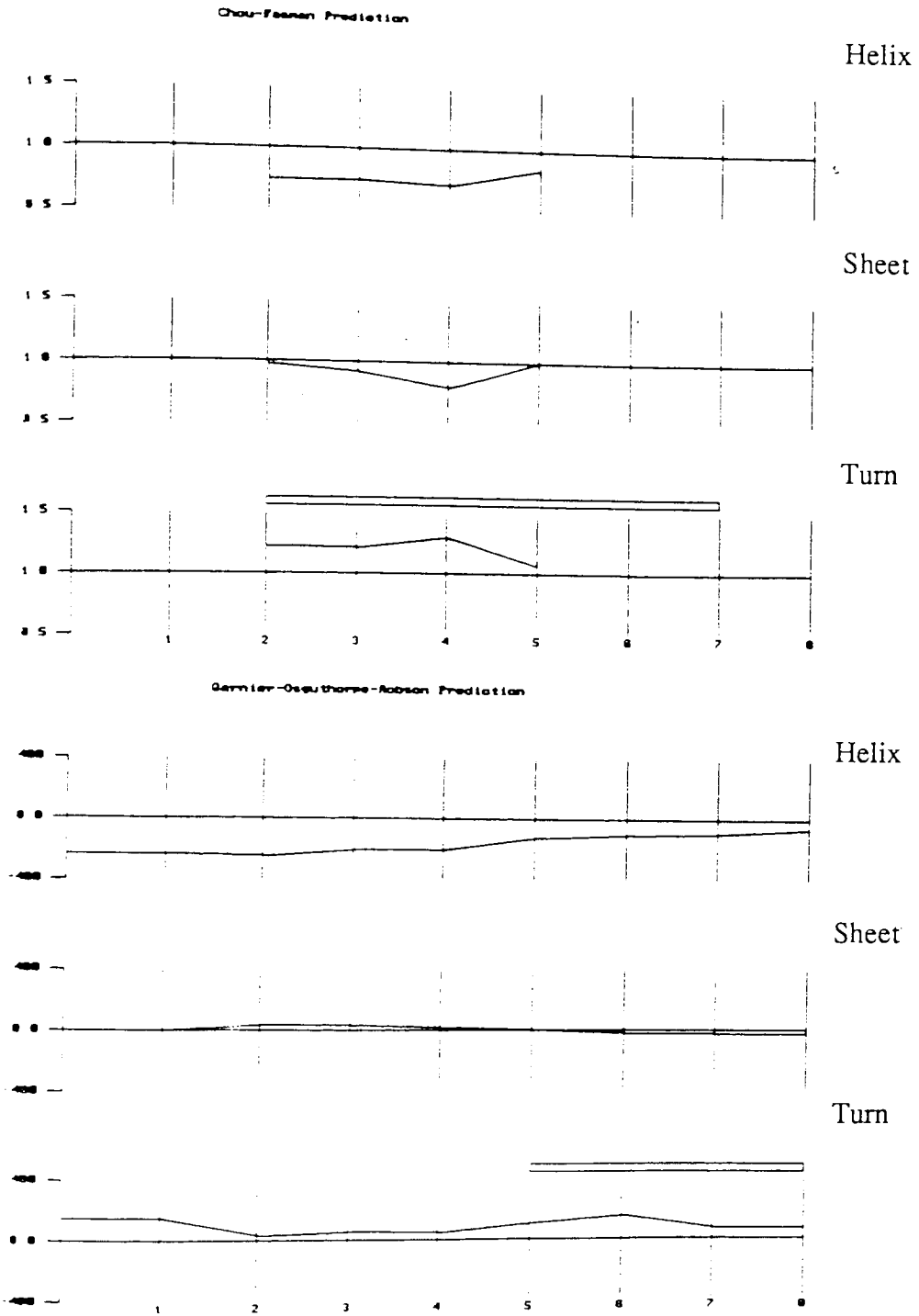


Figure 21 Graphical displays of the Chou-Fasman and Garnier-Osguthorpe-Robson structure predictions for peptide 53

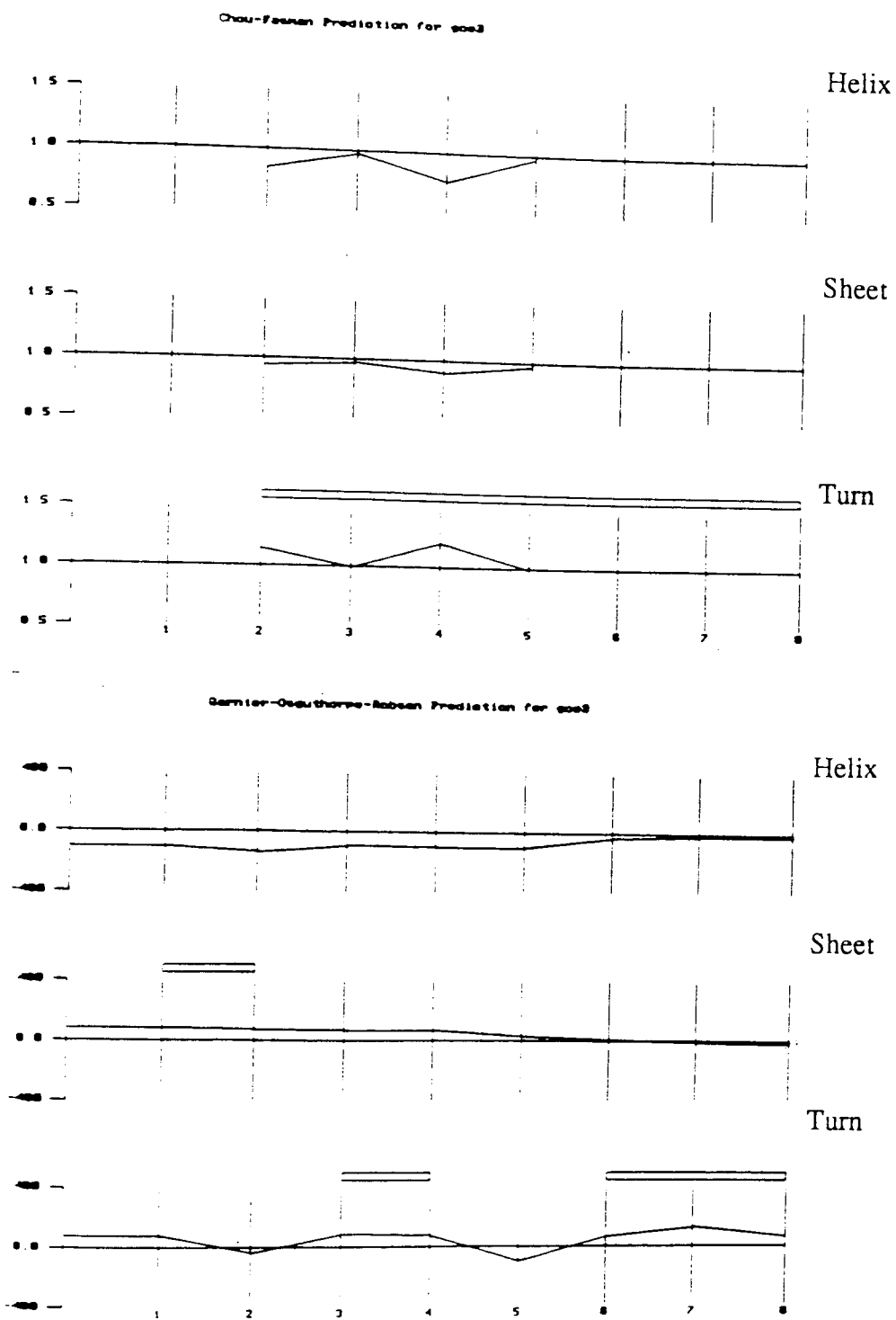


Figure 22 Graphical displays of the Chou-Fasman and Garnier-Osguthorpe-Robson structure predictions for peptide 54

That a turn structure was being predicted in all of these sequences was not unexpected as it is widely accepted that proline residues are responsible for producing turn structures in protein secondary structure.<sup>78</sup> However considering peptide **55** as a whole, the two predictions produced using the Garnier-Osguthorpe-Robson and Chou-Fasman methods differ considerably (Fig.23). The Garnier-Osguthorpe-Robson method suggested that, within a protein, the peptide sequence would be more likely to adopt the structure of an  $\alpha$ -helix than a  $\beta$ -sheet, whereas the Chou-Fasman method predicted the peptide sequence is only likely to adopt a  $\beta$ -sheet conformation. Which prediction was correct is unclear.

The predicted structures of the peptides **53-55** were also compared with structure predictions of the active site regions of a number of bacterial DNA MTASEs. Sequences of up to 11 amino acids in length were taken from around the active site cysteine of the bacterial enzymes and applied to the structure prediction calculations. The results from this study could not be compared with actual structures as only one three dimensional structure has been elucidated for the enzyme EcoRII but to very poor resolution.<sup>79</sup> The Garnier-Osguthorpe-Robson structure prediction for the enzyme EcoRII (Fig.24) suggests a highly turned structure similar to the peptide **53** possessing 3 turns in the short 11 amino acid sequence. A very similar prediction was produced for the prokaryotic enzymes MspI (Fig.25) and BSprI (Fig.26) again possessing structures containing 3 turns in a short sequence were predicted. The secondary structure of the active site region of the enzyme BSprI appears very similar to that proposed for peptide **55** possessing a single turn at the beginning of the sequence. However the overall prediction for the structure differs between the two sequences .

The amino acid sequence of the mouse DNA MTASE was also subjected to Chou-Fasman and Garnier-Osguthorpe-Robson analysis. The prediction for the active site sequence in the mammalian enzyme suggests that the Pro-Cys sequence lies at a turn structure and followed by a region of  $\alpha$ -helix. Viewing the DNA MTASE as a whole, the enzyme



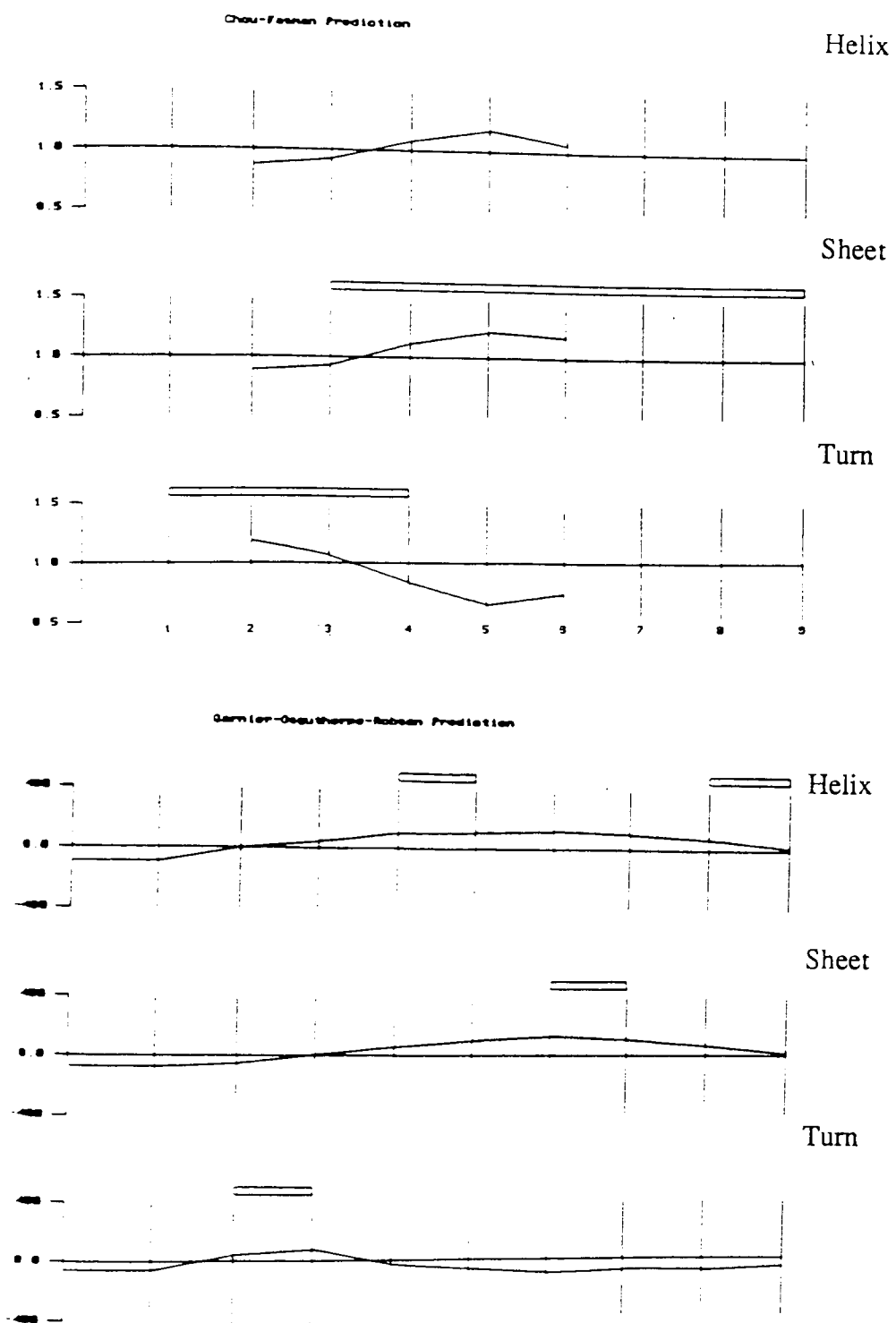


Figure 23 Graphical displays of the Chou-Fasman and Garnier-Osguthorpe-Robson structure predictions for peptide 55

Garnier-Osguthorpe-Robson Prediction for ecorII

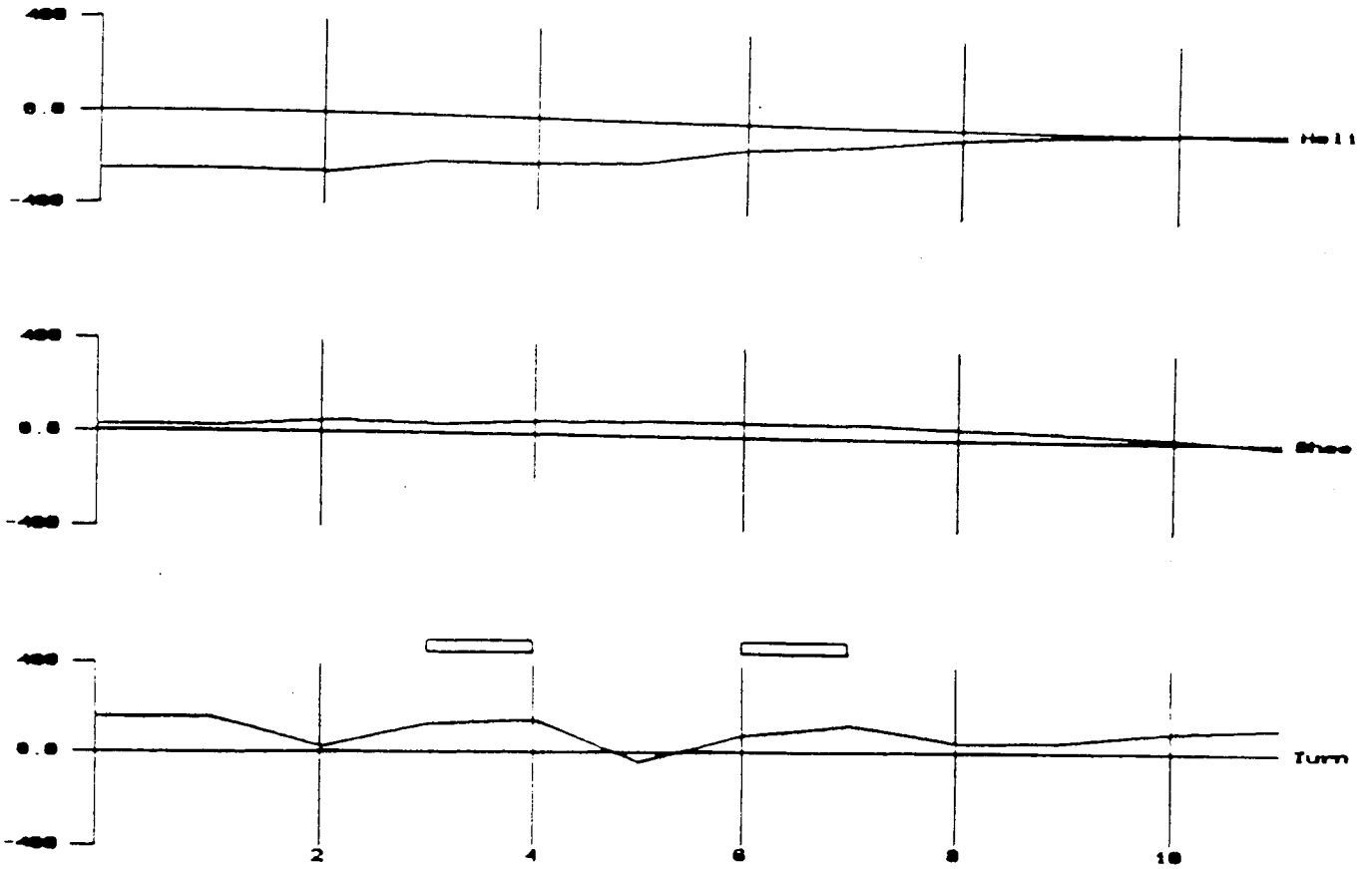


Figure 24 Graphical display of the Garnier-Osguthorpe-Robson structure prediction for EcorII

Garnier-Osguthorpe-Robson Prediction for  $Msp1$

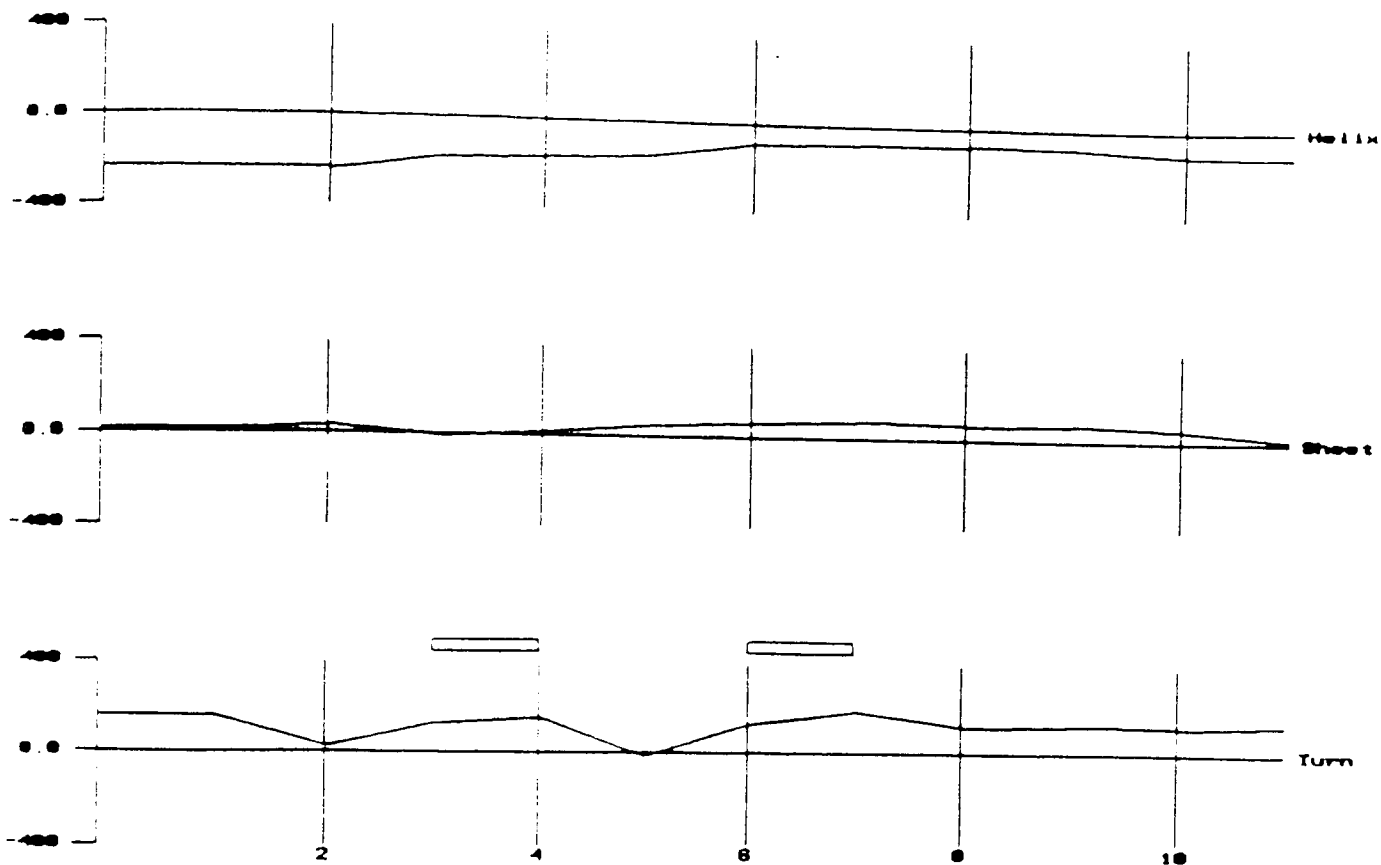


Figure 25 Graphical display of the Garnier-Osguthorpe-Robson structure prediction for  $Msp1$

Garnier-Osguthorpe-Robson Prediction for bspr1

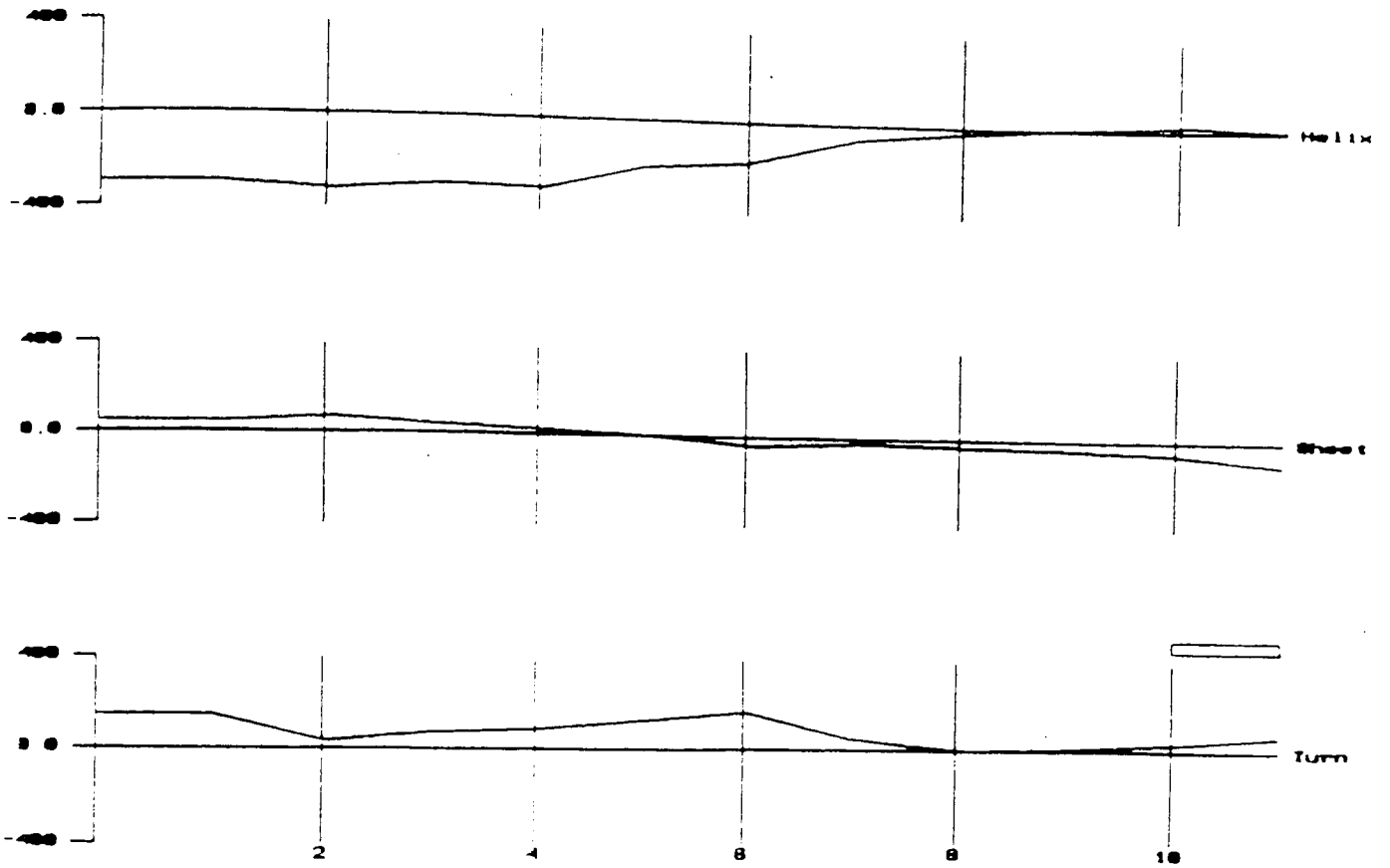


Figure 26 Graphical display of the Garnier-Osguthorpe-Robson structure prediction for Bspr1

appears to show some helical characteristics in the carboxy terminal domain of the enzyme: this would seem feasible for a DNA binding enzyme. The prediction methods do however differ somewhat in the nature of the structure they predict. The Garnier-Osguthorpe-Robson prediction suggests a right-handed turned structure for the enzyme, the Chou-Fasman prediction suggests a much more open sheet like structure. Both structure predictions appear to show one large binding pocket, with possibly another smaller binding pocket being present. The larger binding pocket is likely to be the active site region with the smaller pocket possibly being the site at which SAM binds. The graphical information suggests that from amino acid 1 to 500 the enzyme will most probably adopt a coil or turned structure. From amino acid 1000 to 1500 the enzyme is unlikely to adopt a helical conformation, but will exhibit a structure consisting of turns and  $\beta$ -sheet. The graphical data is inconclusive but for the first 500 amino acids a turned helical structure is predicted, and for the rest of the enzyme a highly turned sheet structure is the most likely structure for the enzyme to adopt. From this information it can be suggested that the active site region must exist somewhere in the region between amino acids 1-500.

The results of these studies suggest that the active site region will contain amino acids in a highly turned structure but the analysis is unable to specify the handedness or the degree of turn produced. Whether the predicted turn conformation in the active site structures is sufficient to ensure the interaction of the active site cysteine thiol and the target cytosine residue is unclear. These prediction calculations do show that the structural requirements for a cytosine C5-position activating molecule will be complex. Crystal structure data will be required to determine the active site structure accurately: However these predictions suggest that smaller peptides may be able to adopt conformations seen in much larger proteins.

The studies also suggest a next potential synthetic target. A peptide sequence known to adopt an  $\alpha$ -helical structure attached to one of the synthesised peptides (peptides **53-55**),

so that overall the helical structure will be unaffected would be an attractive molecule that would exhibit DNA binding characteristics and MTASE activity.

## CHAPTER 7

### Synthesis of sulphonium compounds

From the investigations outlined in Chapter 5, two distinct types of target molecule were selected as potential alkylating agents of the C5 position of cytosine.

These are:

- (i) Simple sulphonium compounds, which might be expected to show a non-specific alkylating ability with some ability to alkylate cytosine at position C5.
- and
- (ii) Sulphonium-thiol compounds which would possess both a sulphonium centre and a thiol group and might be expected to show a more specific alkylating ability for position C5 of cytosine.

#### 7.1 Synthesis of simple sulphonium salts

In initial biological studies S-methylmethionine sulphonium chloride had shown an ability to alkylate cytosine at position C5 (Chapter 5), which was enhanced in the presence of the enzyme DNA MTASE. This compound was therefore used as a specific lead molecule. A series of S-alkyl methionine sulphonium molecules was prepared to investigate whether the size of the alkyl group attached to the sulphonium centre affected the ability of the methionine sulphonium salt to donate its alkyl group. In addition these studies would provide important information on potential synthetic strategies for the more complex sulphonium molecules.

There exists a large number of synthetic methods for producing sulphonium compounds.<sup>80</sup> The simplest of which involves the alkylation of a sulphide with methyl iodide. A number of methylating agents and reaction conditions leading to S-alkylmethionine sulphonium compounds were explored the results of which are summarised in Table 6.

Table 6 Conditions for the formation of alkyl sulphonium salts of methionine

Methylating agent	Conditions	Temp/°C	yield(%)	Ref
MeI	acetic/formic acid	25	68	81
Et I	acetic/formic acid	25	NR	NA
Et I	water	45	51	82
BuI/Bzl I	acetic/ formic acid	25	NR	NA
BuI/Bzl I	water	45	NR	82

NR- no reaction

NA- not applicable

S-methylmethionine sulphonium iodide (4) was synthesised in 68% yield following the method of Toennies and Kolb,<sup>81</sup> using methyl iodide in acetic/formic acid. Attempts to synthesise S-ethylmethionine sulphonium iodide (6) using ethyl iodide as the alkylating agent employing the same acetic/formic acid solvent conditions failed. Using ethyl iodide under aqueous conditions at 45°C resulted in a 51% yield of S-ethylmethionine sulphonium iodide. The synthesis of S-butyl and S-benzylmethionine sulphonium iodide has been

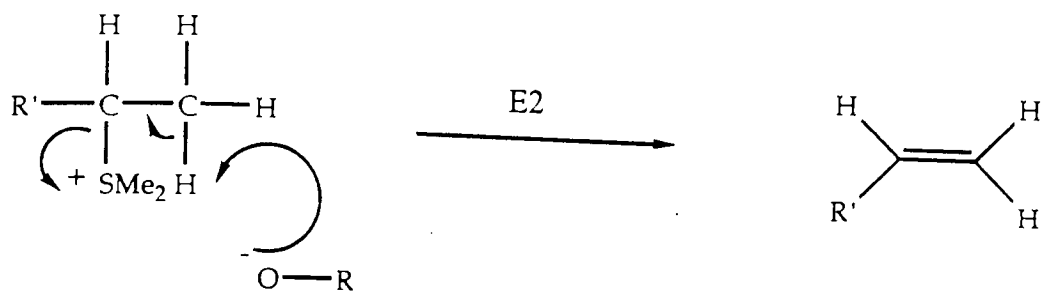


reported in the literature,<sup>82</sup> however all previous attempts to isolate these compounds as pure products failed. Attempts to synthesise these compounds using either acetic/formic acid or aqueous conditions were unsuccessful. In order to assess the likelihood of donation of a butyl group from a sulphonium centre in comparison to a methyl or ethyl group, the simple butyl sulphonium salt, tributyl sulphonium perchlorate (**7**) was prepared by the reaction of butan-1-ol with n-butylsulphide in 70% perchloric acid.<sup>83</sup>

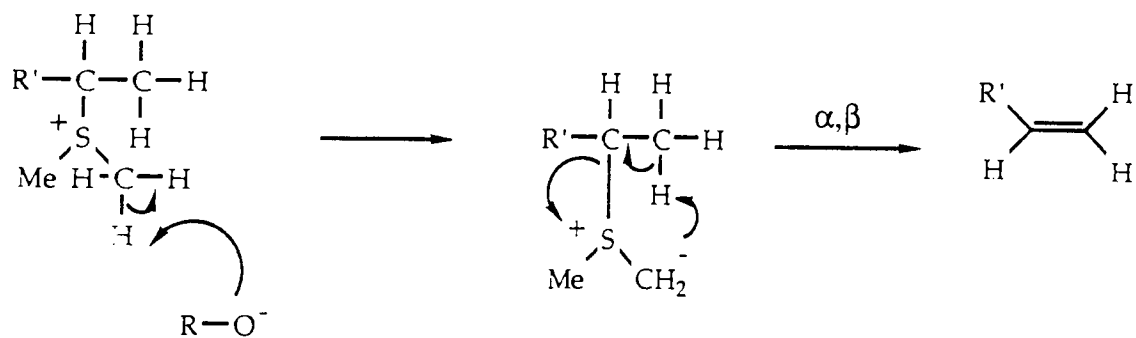
Methyl-*p*-toluene sulphonate was also investigated as an alkylating agent for the synthesis of sulphonium salts. The reaction of N-CBZ-methionine-4-nitrophenyl ester with methyl-*p*-toluene sulphonate in ethyl acetate resulted in a 22% yield of the alkylated product, N-CBZ-S-methylmethionine-4-nitrophenyl ester sulphonium iodide (**8**).<sup>84</sup> This method was inefficient and not readily reproducible and was therefore considered unsuitable as a general synthetic method for sulphonium salts.

#### 7.1.1 Instability of sulphonium salts

Methyl iodide did not prove universally useful for the synthesis of sulphonium salts, attempts to alkylate other methionine-containing molecules using methyl iodide in acetic/formic acid produced thick red oils from which sulphonium compounds were not recoverable. That the odour of dimethyl sulphide, a decomposition product of sulphonium compounds, could be detected, suggested sulphonium compounds may have formed and then subsequently decomposed. A possible explanation of this may lie in the fact that the iodide ion is strongly nucleophilic and hence may facilitate the breakdown of the cationic sulphonium centre. The lability of sulphonium salts in solution is well documented and is favoured by an anion of high nucleophilicity.<sup>80</sup> The presence of dimethyl sulphide would suggest the sulphonium compounds are undergoing an elimination reaction (Fig.27). The



E2 elimination



α,β elimination

Figure 27 Mechanism of elimination that may occur at sulphonium centres

odour of dimethyl sulphide was less apparent in the reactions where perchlorate sulphonium salts were synthesised. This may be a consequence of the lower nucleophilicity of the perchlorate ion.

## 7.2 Synthesis of S-[<sup>3</sup>H-methyl]methionine sulphonium iodide

Restriction enzyme analysis has been used widely to monitor cytosine C5 methylation in DNA, however it has been reported that the results from inhibition studies using the methylation sensitive restriction enzymes MspI and HpaII may be unreliable. Two problems exist with the restriction enzyme analysis approach:

- (i) The recognition sequence of HpaII and MspI (CCGG) limits the usefulness of these enzymes in that only 15% of those methylated cytosines can be identified using these two enzymes isochizomerically, as a consequence of the distribution of GC throughout the genome.
- (ii) It has been reported that alkylations at other sites in DNA, eg. O<sup>6</sup> of guanine affects the cutting pattern of HpaII and MspI.<sup>85</sup>

The information afforded by the restriction enzyme studies therefore needed clarification. A more accurate assay system was required to support the preliminary results reported in Chapter 5. The system proposed involved incubation of DNA with a labelled sulphonium compound, followed by digestion of the DNA with a DNase and separation of the products by HPLC (Chapter 9). A synthetic target identified was S-[<sup>3</sup>H-methyl]methionine sulphonium iodide. Prior to attempting the synthesis of the labelled compound it was necessary to develop a purification procedure.

### 7.2.1 Purification of S-[<sup>3</sup>H-methyl]methionine sulphonium iodide

One commonly used purification method for radiolabelled compounds is HPLC. To date there is no published HPLC method for the separation of S-methylmethionine sulphonium salts and methionine, and therefore a separation method had to be developed.

Initially it was necessary to determine a working ultraviolet (UV) wavelength for the detection of methionine and S-methylmethionine sulphonium iodide. The UV spectrum of S-methylmethionine sulphonium iodide showed a peak at  $\lambda_{\text{max}}$  230nm and methionine was also weakly absorbing in this region, 234nm was chosen as an optimum wavelength for detection of both compounds. Three HPLC methods were investigated, conventional reverse phase (RP), ion-pair and ion-exchange chromatography (Table 7)

The first of these methods was conventional RP chromatography employing a C18 bonded column (Method A, Table 7). Under these conditions methionine and S-methylmethionine sulphonium iodide were too polar to bind to the column, had short retention times and were not sufficiently separated. Attempts to increase the separation by increasing the water concentration in the solvent system proved unsuccessful. The second method made use of the ion-pair reagent hexane sulphonic acid (HSA). The principle behind the use of these reagents is that the anionic HSA can ion-pair with cationic species and hence form an overall neutral ion-pair complex. The neutral ion-pair is therefore retained more efficiently by the column. Ion-pair chromatography was effective

Table 7 HPLC conditions for the separation of S-methylmethionine sulphonium iodide and methionine

Method	Column Functionality	Solvent system	Methionine $R_t$ /mins	S-methylmethionine sulphonium iodide $R_t$ /mins
A	RPC18	methanol-water 50:50	1.78	1.44
B	RPC18	10% methanol- 90% HSA <sup>a</sup>	1.80	1.33
C	Partisil 10-SCX <sup>b</sup>	5% methanol/ 95% (0.5% acetic acid)	3.30	1.65
D	Partisil 10-SCX <sup>b</sup>	5% methanol- 95%(0.001% acetic acid)	4.48	2.37

a) Hexane sulphonic acid in water

b) Strong cation exchanger

in increasing the  $R_t$  of methionine, but the more polar S-methylmethionine sulphonium iodide  $R_t$  was unaffected. Extending this idea of separating the methionine and S-methylmethionine sulphonium iodide on the basis of charge an ion-exchange HPLC column (Strong cation exchanger) was used. Using Method C retention of S-methylmethionine sulphonium iodide was observed. For ion-exchange chromatography to be successful, it was important that in both methionine and S-methylmethionine sulphonium iodide the amino acid portions of each molecule existed in a zwitterionic form. The  $pK_a$  (carboxyl) of methionine is 2.0 and the  $pK_b$  (amino) of methionine is 9.0. Values for S-methylmethionine sulphonium iodide were unavailable. Therefore a number of

separations were attempted in the pH range 4-7. Method D provided the best separation of methionine and S-methylmethionine sulphonium iodide.

UV detection would not be sufficiently sensitive for the detection of the small quantities of products in the radiolabelled synthesis of S-methylmethionine sulphonium iodide. Therefore the column eluate was collected as fractions at thirty second intervals and analysed by scintillation spectrometry. The product was ultimately isolated by lyophilisation. The radioactive peaks correlated well with the UV chromatograms obtained during method development (Fig.28).

### 7.2.2 Synthetic methods

The synthesis of labelled S-methylmethionine sulphonium iodide has been reported in brief.<sup>88</sup> Attempts to synthesise S-[<sup>3</sup>H-methyl]methionine sulphonium iodide using this literature method resulted in a poor radiolabelled yield of only 2.5%. The radiolabelled reaction was inefficient and therefore in an effort to improve the yield, the cold reaction was investigated.

The literature method employed the synthetic route devised by Toennies and Kolb using methyl iodide in acetic/formic acid.<sup>81</sup> This involved leaving the reaction for three days. A possible explanation for the poor radiolabelled yield may lie in that the products and/or reactants may undergo radiolysis over the three day reaction period. It was therefore important to determine the optimum time required for the reaction to reach completion. The 'cold' alkylation of a 1mM solution of methionine using methyl iodide in deuterated acetic acid at 25°C was monitored by <sup>1</sup>H nuclear magnetic resonance (NMR) (Fig.29). The reaction proceeded smoothly with the sulphonium methyl peak at  $\delta$ 3.0 parts per million

CPM Counts per minute

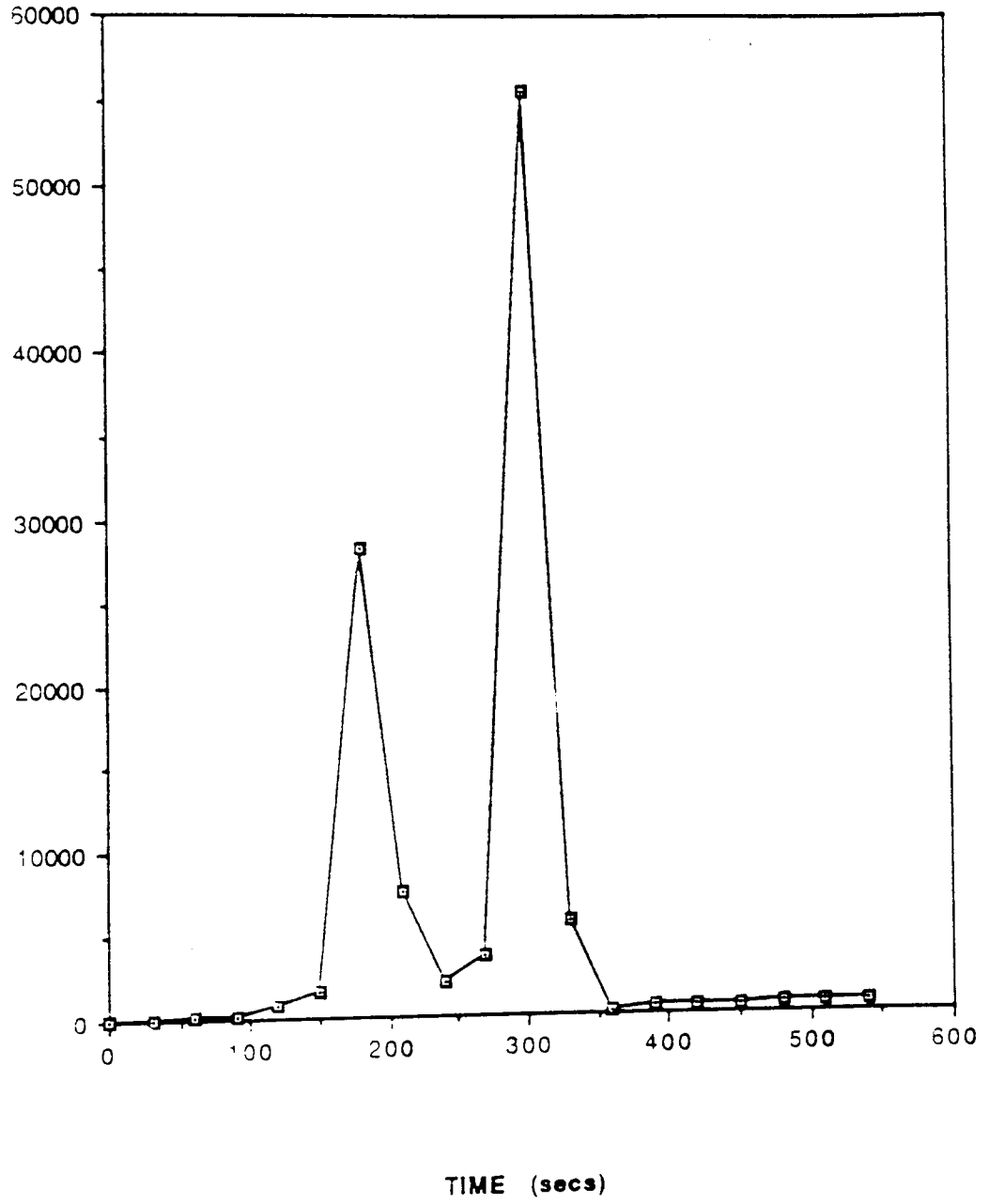


Figure 28 HPLC purification profile of S-[<sup>3</sup>H-methyl]methionine sulphonium iodide

Peak at  $\delta$  3.0ppm due to methyl sulphonium group

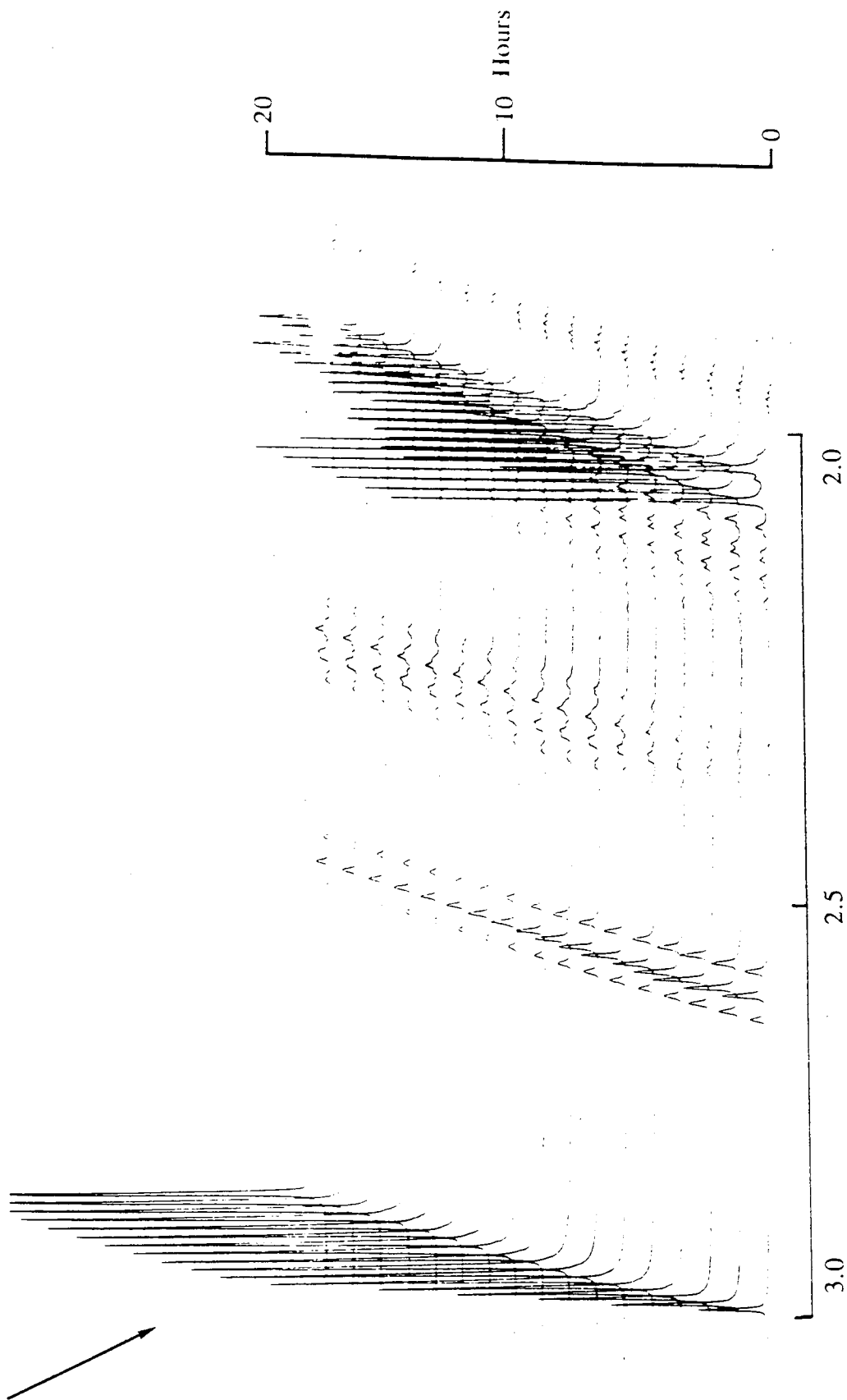


Figure 29  $^1\text{H}$  NMR stack plot of the synthesis of S-methylmethionine sulphonium iodide showing the region  $\delta$  2.0-3.0ppm for the first 17 hrs of reaction



(ppm) increasing in proportion to the rate of decrease of the methyl peak of methyl iodide at  $\delta 1.0$  ppm. The rate of formation of the sulphonium centre was approximately linear over the first 6 hrs (Fig.30). After which it appeared to increase reaching a maximum at 10 hrs. The spectrum showed that the reaction was essentially complete after 15 hrs, and that a reaction time of three days appeared unnecessary.

The labelled reaction was repeated over 15 hrs, but again a poor radiolabelled yield was achieved. The poor yields obtained using this method prompted an investigation of other literature synthetic methods. The methylating agents and conditions investigated are summarised in Table 8.

Table 8 Conditions for the synthesis of S-methylmethionine sulphonium iodide

Method	Methylating agent	Reaction conditions	Temp°C	% yield	reference
A	methanesulphonic acid	anisole/MSA	25	94	95
B	methanol	methanolic HCl	60	95	108
C	methanol	conc H <sub>2</sub> SO <sub>4</sub>	25	90	109
D	methyl iodide	water	40	70	82

Methods A to D produced good yields, however methods A and C were considered unsuitable for the small scale labelled synthesis because purification procedures for both these reactions would be difficult and inefficient. Method B appeared the most promising, the product being isolated as a solid after a simple workup procedure involving removal of solvent in *vacuo*.

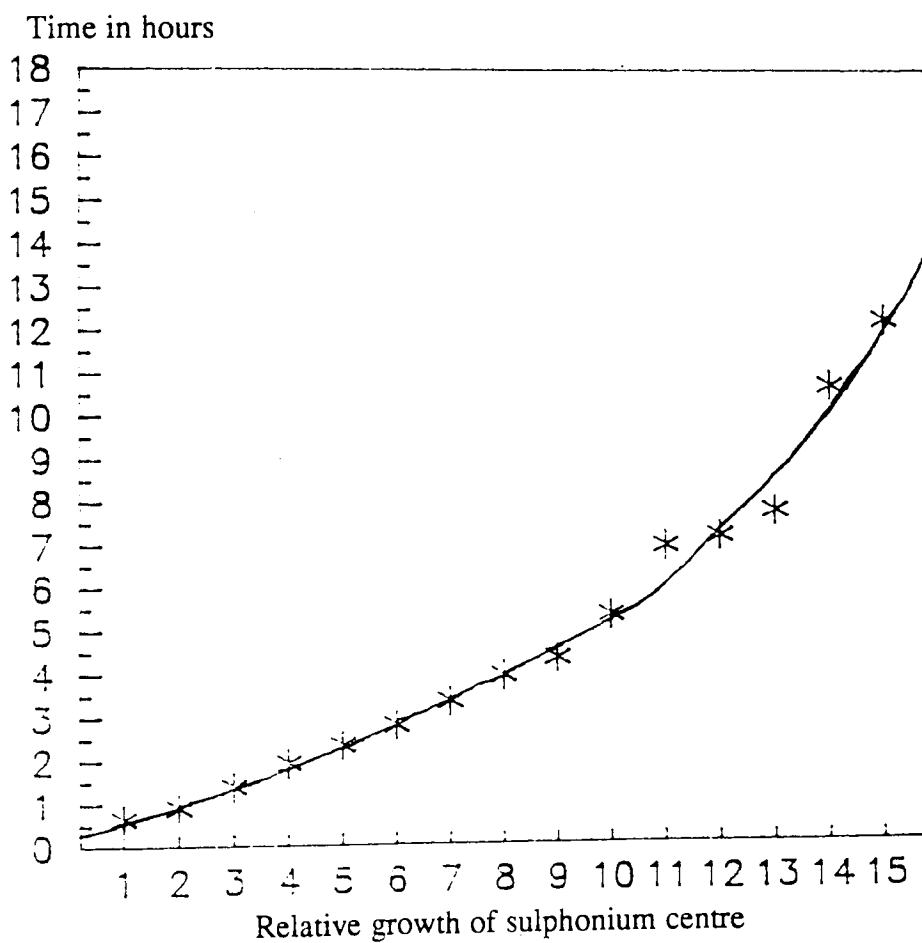


Figure 30 Graphical display of S-methylmethionine sulphonium iodide formation monitored by  $^1\text{H}$  NMR

The radiolabelled synthesis was carried out using  $^3\text{H}$ -methionine ( $340\mu\text{Ci}$ ,  $4.4\mu\text{mol}$ ) in a sealed vessel at  $60^\circ\text{C}$  for 24 hrs. The reaction was monitored by thin layer chromatography (TLC) (n-butanol, acetic acid, water (12:5:3)) which showed that after 24 hrs the reaction had not reached completion. Purification by preparative HPLC (see section 7.2.1) resulted in a poor radioactive yield (1.1%) of S-[ $^3\text{H}$ -methyl]methionine sulphonium iodide.

To investigate the suitability of method D, a reaction was carried out using methionine (1mmol) spiked with  $^3\text{H}$ -methionine ( $100\mu\text{Ci}$ ). Purification by HPLC produced S-[ $^3\text{H}$ -methyl]methionine sulphonium iodide (30%).

This result suggested that method D possessed a number of advantages over other methods considered:

- (i) improved radiochemical yield
- (ii) simple workup (lyophilisation)
- (iii) easier purification (no contaminants present only reactants and product)

Method D was selected as the most suitable for the synthesis of S-[ $^3\text{H}$ -methyl]methionine sulphonium iodide.

Aware of the earlier problems of poor radiolabelled yield the radiosynthesis of S-methylmethionine sulphonium iodide was performed on a larger scale using 1mCi of [ $^3\text{H}$ ]-methionine and afforded  $150\mu\text{Ci}$  of S-[ $^3\text{H}$ -methyl]methionine sulphonium iodide of radiochemical purity  $>90\%$ .

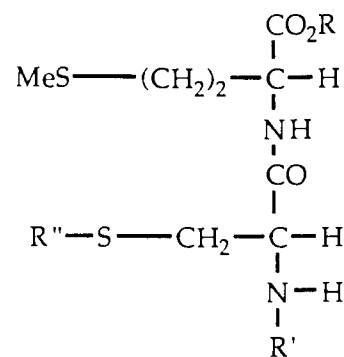
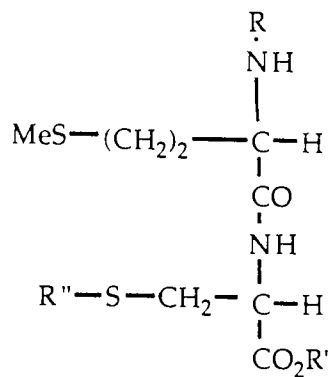
During the development of a synthetic method for labelled S-methylmethionine sulphonium iodide, it became clear that there was a problem with the stability of S-[<sup>3</sup>H-methyl]methionine sulphonium iodide in the reaction mixture. It would be interesting to investigate whether similar sulphonium salts exhibit similar stability properties. The synthesis of labelled S-methylmethionine sulphonium iodide using method D is an improvement on the reported literature method and provides material of suitable quality for biological assay.

### 7.3 Synthesis of peptidic precursors to sulphonium-thiol molecules

A number of simple di, and tetrapeptides were identified as possible precursors to sulphonium-thiol molecules (Fig.31). Two particular amino acids methionine and cysteine stood out as obvious choices to be included in any precursor molecule, methionine to provide a sulphide for generation of the sulphonium centre and cysteine to provide a thiol residue.

The strategy behind the synthesis of this type of molecule was complex as the molecule would contain two sulphur centres possessing similar susceptibilities to alkylating agents. This problem could be overcome by ensuring that in the planned synthesis the final two reaction steps would involve respectively :

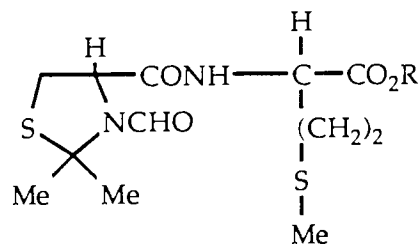
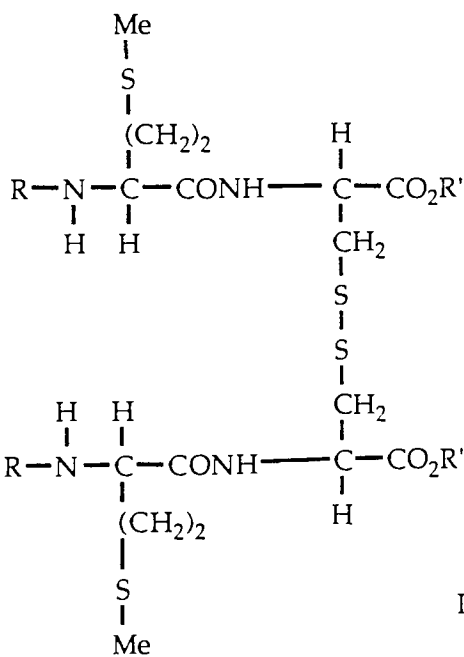
- 1) alkylation of a sulphide in a molecule also possessing a protected thiol group.
- 2) deprotection of the thiol under acidic conditions to ensure stability of the sulphonium centre.



(9) R=CBZ R'=methyl R''=Bzl

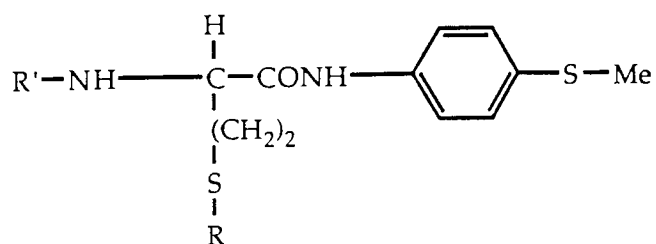
(18) R=methyl R'=CBZ R''=Bzl

(10) R=CBZ R'=methyl R''=Trityl



(19) R = Me

(22) R=CBZ R'=methyl

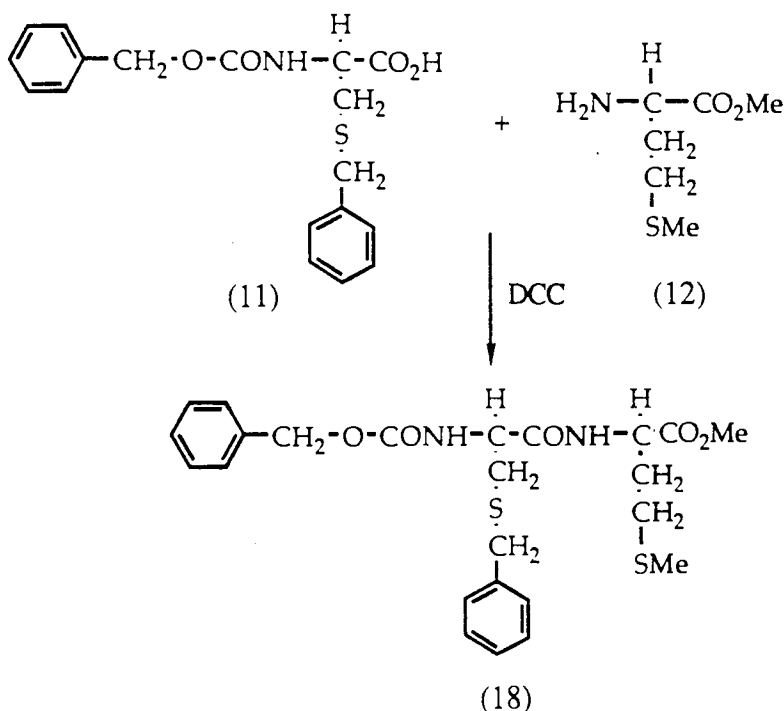


(25) R' = Fmoc R = trityl

Figure 31 Structures of possible precursors to sulphonium-thiol molecules

### 7.3.1 Synthesis of N-CBZ-S-(Bzl)cysteinylmethionine methyl ester (18)

N-CBZ-S-(Bzl)cysteinylmethionine methyl ester (18) was synthesised in 74% yield according to reaction scheme 1.

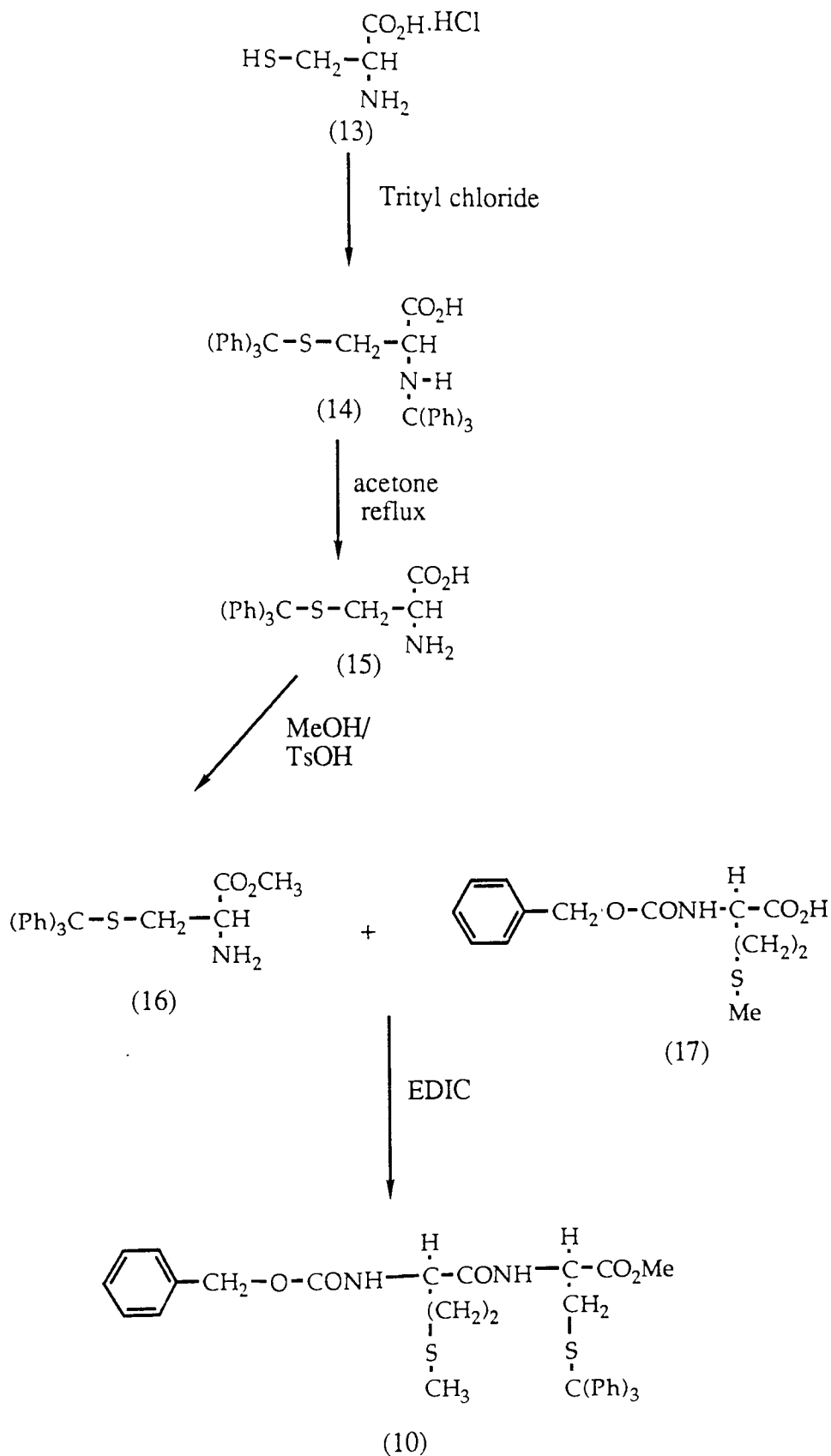


Scheme 1 Synthesis of N-CBZ-S-(Bzl)cysteinylmethionine methyl ester (18)

The coupling of N-CBZ-S-(Bzl)cysteine (11) with methionine methyl ester (12) using the ethyl chloroformate method initially proved successful but was poorly reproducible.<sup>128</sup> This might be accounted for by the tendency of ethyl chloroformate to oxidise readily on standing. Coupling with dicyclohexylcarbodiimide (DCC) proved to be more effective resulting in good yields of dipeptide, and was therefore the method of choice.<sup>98</sup>

### 7.3.2 N-CBZ-methionyl-S-(Trt)cysteine methyl ester (10)

N-CBZ-methionyl-S-(Trt)cysteine methyl ester (10) was synthesised according to scheme 2.



Scheme 2 Synthesis of N-CBZ-methionyl-S-(Trt)cysteine methyl ester (10)

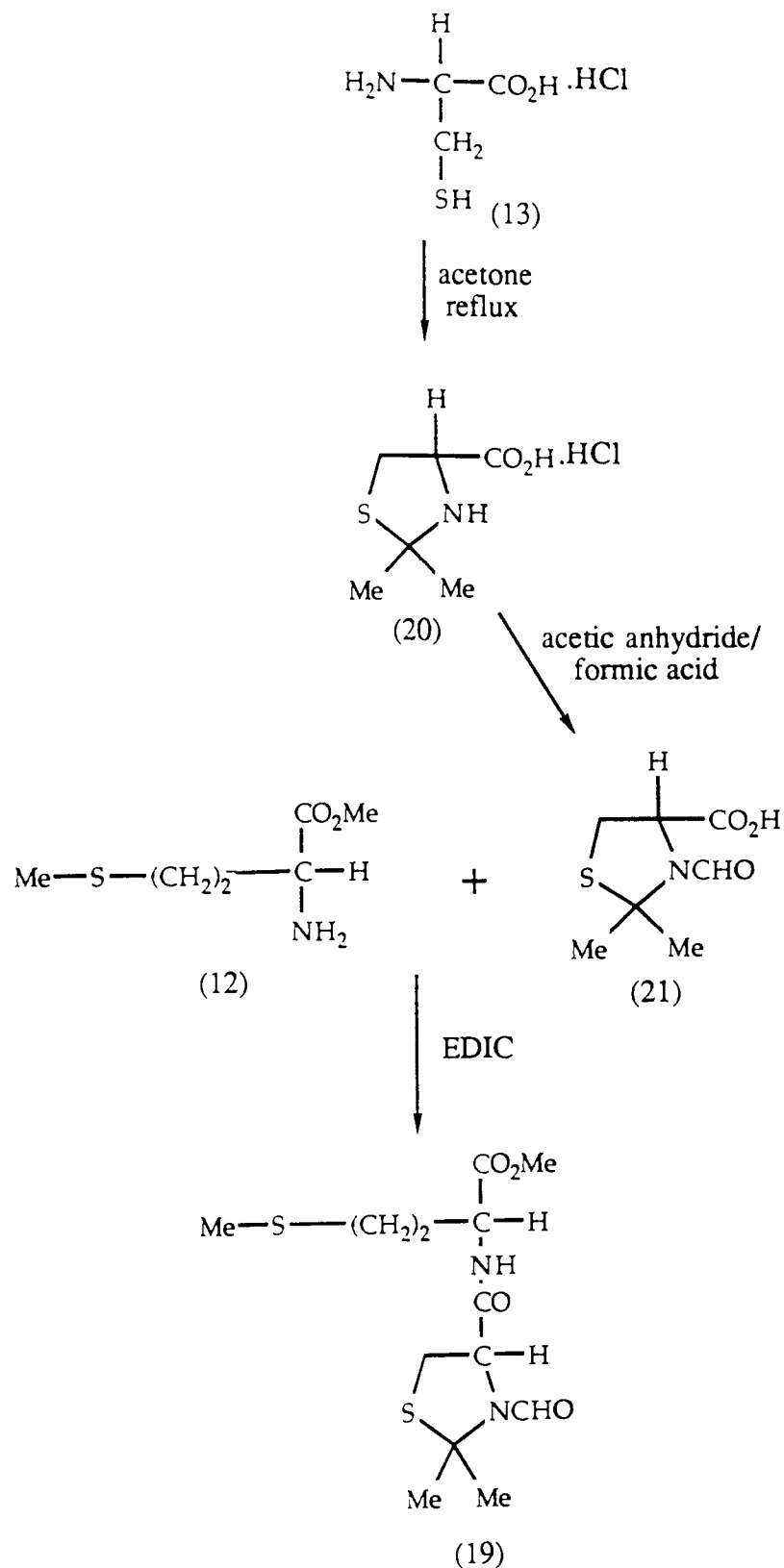
N,S-ditritylcysteine (**14**) was synthesised by the reaction of cysteine hydrochloride (**13**) with trityl chloride. The N-(Trt) group was removed by refluxing **14** in acetone. Esterification was achieved using methanol catalysed by p-toluenesulphonic acid to give S-(Trt)cysteine methyl ester (**16**) in 43% yield. **16** was then coupled to N-CBZ-methionine (**17**) with the aid of a water soluble coupling agent 1-ethyl-2,2-diisopropylcarbodiimide(EDIC)<sup>87</sup> to give **10**.

### 7.3.3 Synthesis of N-[2,2-dimethyl-3-formyl-4-thiazolidinecarboxy]methionine methyl ester (**19**)

N-[2,2-dimethyl-3-formyl-4-thiazolidinecarboxy]methionine methyl ester (**19**) was synthesised according to the reaction scheme 3.

In this synthesis the cysteine residue thiol was protected as a thiazolidine derivative as opposed to the S-trityl derivative. Release of the free thiol is afforded by hydrolysis of the ring under acidic conditions and was therefore of potential use in the synthesis of a sulphonium-thiol molecule. The thiazolidine derivative was readily prepared from cysteine hydrochloride by refluxing in acetone. The product formed 2,2-dimethylthiazolidinecarboxylic acid hydrochloride (**20**) has free carboxyl and amino functions, one of which requires protection before coupling. In this synthesis the amino function was protected as the formyl derivative by the reaction of **20** with acetic anhydride in formic acid to give 2,2-dimethyl-3-formylthiazolidinecarboxylic acid (**21**). The protected thiazolidine derivative **21** was then coupled with methionine methyl ester using EDIC to give **19** in 45% yield.

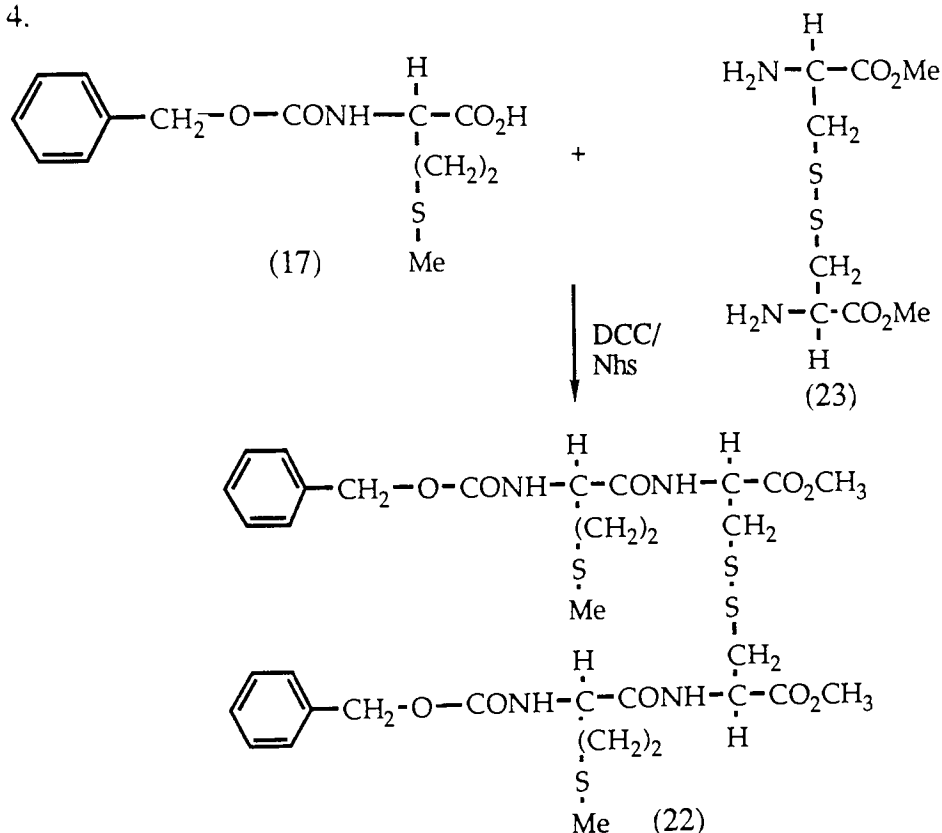




Scheme 3 Synthesis of N-[2,2-dimethyl-3-formyl-4-thiazolidinecarboxy]methionine methyl ester (19)

### 7.3.4 Synthesis of bis(N-CBZ-methionyl)cystine dimethyl ester (22)

Bis(N-CBZ-methionyl)cystine dimethyl ester (22) was synthesised in 31% yield according to scheme 4.



Scheme 4 Synthesis of bis(N-CBZ-methionyl)cystine dimethyl ester (22)

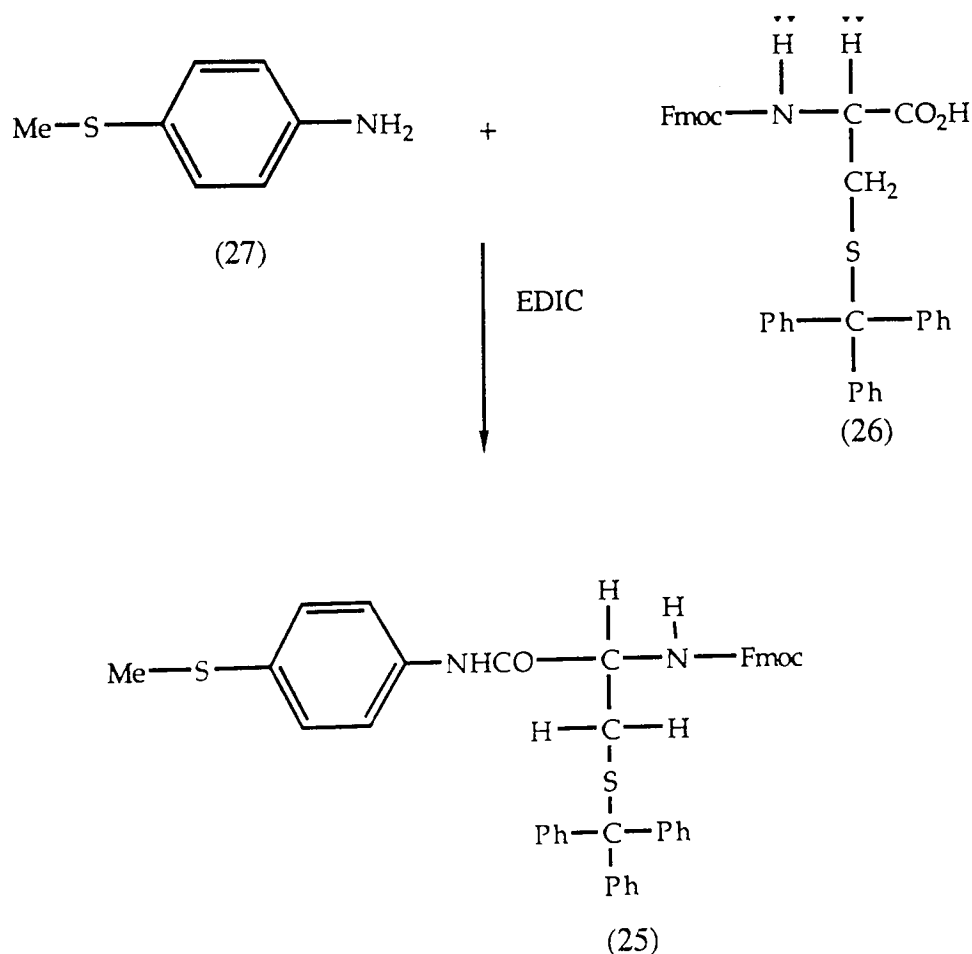
Several coupling methods were evaluated in the synthesis of 22 including the use of ethyl chloroformate and also EDIC. However coupling using DCC in the presence of the catalyst N-hydroxysuccinimide (Nhs) proved most successful. Reduction of the disulphide 22 would liberate 2 molecules of N-CBZ-methionylcysteine methyl ester (24).

All of the precursor compounds described above possess an aliphatic sulphide residue. It was thought interesting to compare the methyl-donating ability of a sulphonium centre possessing an aromatic residue with one possessing an aliphatic residue. With this in mind a precursor molecule was synthesised containing an aromatic sulphide.

### 7.3.5 Synthesis of [N-Fmoc-S-(Trt)cysteinamido]-4-thioanisole (**25**)

[N-Fmoc-S-(Trt)cysteinamido]-4-thioanisole (**25**) was synthesised according to scheme 5.

N-Fmoc-S-(Trt)cysteine (**26**) was coupled to 4-aminothioanisole (**27**) using EDIC. The reaction afforded **25** in 26% yield after purification by flash chromatography.



Scheme 5 Synthesis of [N-Fmoc-S-(Trt)cysteinamido]-4-thioanisole (**25**)

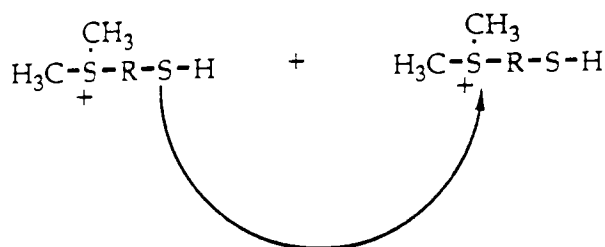
## 7.4 Synthesis of sulphonium-thiol molecules

The mechanism of DNA MTASE is thought to involve the opening of the C5-C6 double bond of cytosine followed by alkylation at the 5-position by a methyl group from SAM (Chapter 1, section 1.3). Synthetic sulphonium-thiol molecules might be expected to be able to perform both of these functions and hence independently alkylate cytosine at position C5. Before considering the syntheses of these molecules, the stability of these compounds needed to be addressed.

### 7.4.1 Chemical stability of sulphonium-thiol molecules

Decomposition following synthesis may occur via two possible routes (Fig.32). The possibility of transalkylation reactions between the sulphonium centre and the free thiol was of most concern.

#### 1) Intermolecular reaction



#### 2) Intramolecular reaction

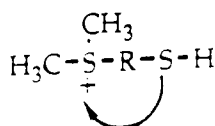


Figure 32 Two possible decomposition routes of sulphonium-thiol molecules

The susceptibility of a sulphonium-thiol molecule to intramolecular reaction could only be addressed once target molecules were synthesised. In order to determine the stability of sulphonium-thiol molecules to intermolecular attack, a model reaction was set up.

Intermolecular decomposition was modelled by monitoring the stability of a simple sulphonium salt in solution in the presence of a thiol containing molecule. The reaction of S-methylmethionine sulphonium iodide with cysteine under aqueous conditions at 37°C was monitored by <sup>1</sup>H-NMR for 24 hrs. No change in the relative integration of the sulphonium centre occurred and no new peaks appeared. Therefore it was concluded that in dilute solution there was no intermolecular reaction between the sulphonium centre and the free thiol. This was confirmed in a similar experiment where no intermolecular reaction was seen between the sulphonium salt, dimethylpropargyl sulphonium bromide (**2**) and cysteine. Having established that the target molecule was likely to be stable under neutral conditions, the synthesis of the dipeptide could be addressed.

#### 7.4.2 Approaches to the alkylation of thiol-containing peptide molecules

From earlier studies (see section 7.2.1) alkylation of a sulphide with methyl iodide was found to be the most suitable method for the synthesis of sulphonium salts.

This methylation reaction was studied further to assess its suitability for alkylation of sulphides contained within peptidic molecules. The dipeptide synthesised N-CBZ-S-(Bzl) cysteinylmethionine methyl ester (**18**) was used as a model target compound for alkylation. Results of the study are shown in Table 9.

Table 9 Sulphonium generation in N-CBZ-S-(Bzl)cysteinylmethionine methyl ester (**18**) with methyl iodide.

Reaction Number	Solvent	Temp/°C	Result <sup>a</sup>
1	CH <sub>2</sub> Cl <sub>2</sub>	25	-
2	EtOH	25	-
3	MeOH	25	+
4	None	25	-
5	CH <sub>3</sub> CO <sub>2</sub> H/HCO <sub>2</sub> H	25	+
6	EtOH	70	+ <sup>b</sup>

a) Reaction products and progress assessed by TLC, eluting solvent acetic acid/propan-2-ol/water (12:5:3)

b) Reaction performed in a sealed vial, complex reaction products.

That no reaction occurred in dichloromethane was not unexpected as the transition state required for the generation of a sulphonium centre is more polar than the starting materials and hence is not favoured by a non-polar solvent. A similar result was found using ethanol at room temperature; at 70°C a complex reaction mixture resulted possibly caused by the elevated temperature of the reaction. In methanol at room temperature when left overnight, a single polar product was observed by TLC.

This series of preliminary experiments suggested that methylation of peptidic compounds using methyl iodide was best performed in methanol at room temperature.

Other alkylation methods were also considered for alkylating the peptidic molecules (Table 10).

Table 10 Sulphonium formation in N-CBZ-S-(Bzl)cysteinylmethionine methyl ester (18) with other alkylating agents.

Alkylating agent	Conditions	Temp <sup>0</sup> C	Comments
Trimethyloxonium tetrafluoroborate	nitromethane/ 3 hrs	25	no reaction
Methyl- <i>p</i> -toluenesulphonate	ethyl acetate/ 24 hrs	25	product isolated decomposed rapidly on exposure to air
Methyl iodide	dichloromethane/ silver perchlorate 2 hrs	25	10% yield of <b>28</b>

No reaction was seen using trimethyloxonium tetrafluoroborate. On reaction with methyl-*p*-toluene sulphonate, a solid precipitated out from the reaction mixture, suggesting formation of product, however, the compound turned a lilac colour on exposure to air. The lilac colouration was exhibited by a number of sulphonium compounds after they had been in contact with air and was considered a marker of oxidative breakdown. Methyl iodide in the presence of the catalyst, silver perchlorate did produce some sulphonium compound. The reaction generally produced low yields but work up afforded a very clean product.

After further consideration of the methods investigated, methyl iodide in methanol or methyl iodide in dichloromethane in the presence of silver perchlorate were the most successful methods for the synthesis of peptidic sulphonium salts.

## **7.5 Synthesis of peptidic sulphonium salts**

### **7.5.1 Alkylation of synthetic peptide precursor molecules**

Several of the peptidic molecules prepared earlier (section 7.3) were then alkylated using both of the methods suggested above in an attempt to produce the required sulphonium salts. Table 11 summarises the successful methods.



Table 11 Conditions for the synthesis of sulphonium peptides with methyl iodide

Parent molecule	Conditions	Product	% Yield
N-CBZ-methionyl-S-(Trt)cysteine methyl ester (10)	MeOH/MeI 24 hrs	(29)	64 <sup>a</sup>
Bis(N-CBZ-methionyl)cystine dimethyl ester (22)	MeI/CH <sub>2</sub> Cl <sub>2</sub> / AgOCl <sub>4</sub> / 2 hrs	(30)	20 <sup>b</sup>
N-[2,2-dimethyl-3-formyl-4-thiazolidinecarboxy]methionine methyl ester (19)	MeI/CH <sub>2</sub> Cl <sub>2</sub> / AgOCl <sub>4</sub> / 24 hrs	no reaction	
[N-Fmoc-S-(Trt)cysteinamido]-4-thioanisole (25)	MeI/CH <sub>2</sub> Cl <sub>2</sub> / AgOCl <sub>4</sub> / 2 hrs	(31)	10 <sup>c</sup>

a) N-CBZ-S-methylmethionyl-S-(Trt)cysteine methyl ester sulphonium iodide (29) was isolated as a stable yellow solid.

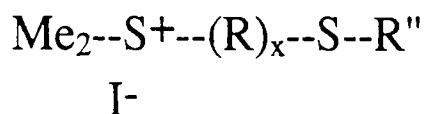
b) Product isolated as the monosulphonium salt N-CBZ-S-methylmethionyl-N-CBZ-methionylcystine dimethyl ester (30).

c) Product [N-Fmoc-S-(Trt)cysteinamido]-S,S-dimethylphenyl sulphonium iodide (31) isolated by flash chromatography.

Alkylation of other precursor molecules with methyl iodide in methanol was unsuccessful.

### 7.5.2 Alkylation of bis(N-CBZ-methionyl)cystine dimethyl ester (**22**)

Alkylation of disulphides normally results in the formation of thiosulphonium salts (Fig.33).<sup>90</sup>



R and R''- any chemical group

Figure 33 Structure of a thiosulphonium salt

**22** contains four possible sites at which alkylation may occur (Fig.34)

The tetra-peptide **22** was alkylated using methyl iodide in dichloromethane in the presence of the catalyst silver perchlorate.

The isolated product was the monosulphonium salt N-CBZ-S-methylmethionyl-N-CBZ-methioninylcystine dimethyl ester (**30**). Reaction in the presence of excess alkylating agent failed to result in alkylation at any other sites. Alkylation at one of the methionine residues was confirmed by examining the FAB mass spectrum. The spectrum shows the loss of dimethylsulphide which is characteristic of dimethylsulphonium compounds. (Fig.35).

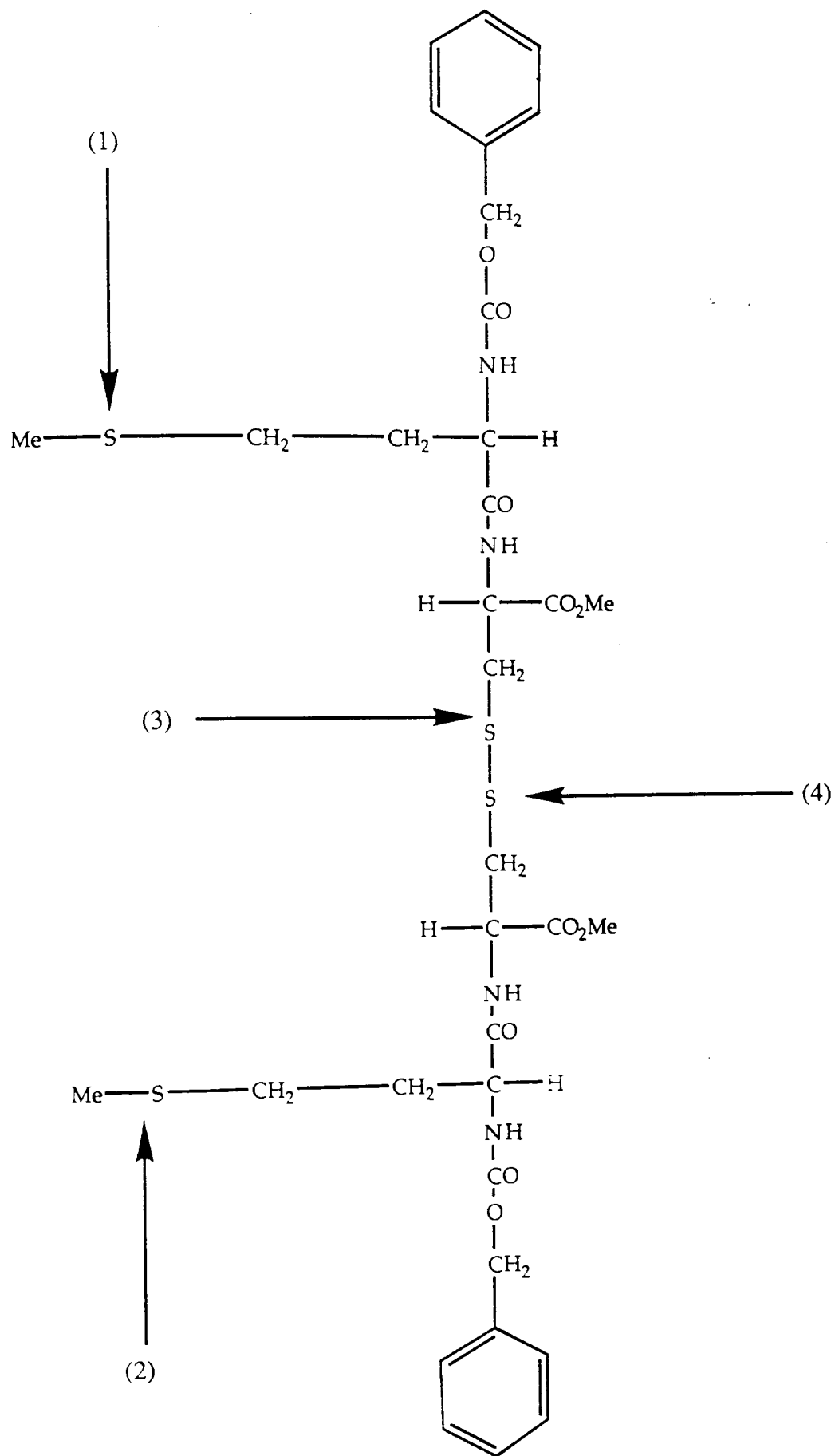
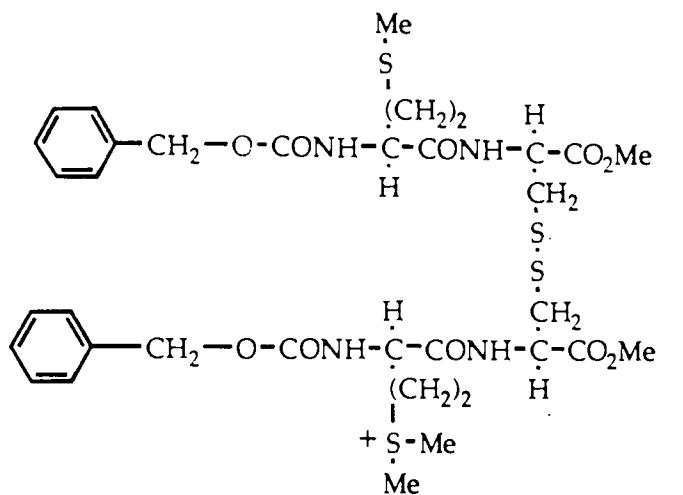
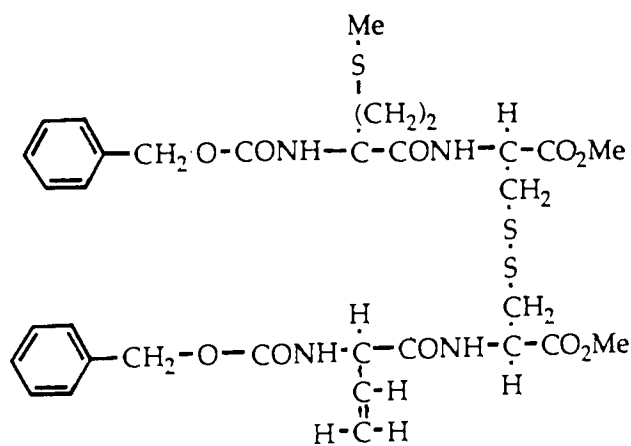
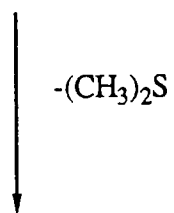


Figure 34 Sites in bis(N-CBZ-methionyl)cystine dimethyl ester (22) that are potential sites for alkylation



[M+H] 816



754

Figure 35 Proposed Fragmentation pattern of N-CBZ-S-methylmethionyl-N-CBZ-methionylcystine dimethyl ester (30) as suggested by the FAB mass spectrum

### 7.5.3 Alkylation of N-[2,2-dimethyl-3-formyl-4-thiazolidinecarboxy]methionine methyl ester (19)

N-[2,2-dimethyl-3-formyl-4-thiazolidinecarboxy]methionine methyl ester (19) possessed two possible sites for alkylation to form a sulphonium centre (Fig.36)

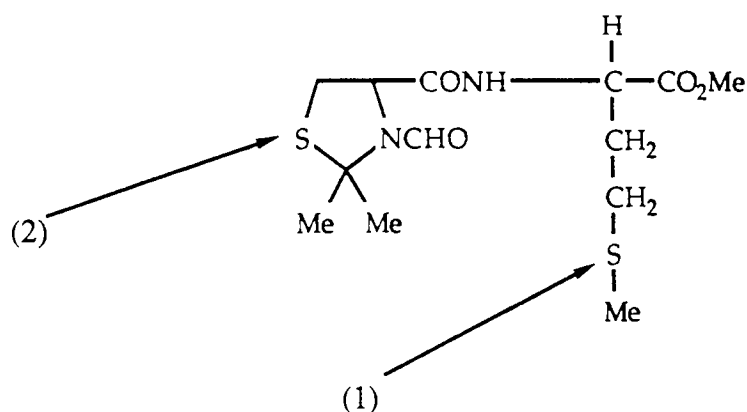


Figure 36 Sites in [N-2,2-dimethyl-3-formyl-4-thiazolidinecarboxy]methionine methyl ester (19) that are potential targets for alkylation

Site 1 would be expected to be most readily alkylated from a steric viewpoint, but site 2 is also a potential site for alkylation. Alkylation using methyl iodide failed to produce any sulphonium compound. A number of the other methylating agents (see section 7.4.2) were used in attempts to produce a sulphonium salt of N-[2,2-dimethyl-3-formyl-4-thiazolidinecarboxy]methionine methyl ester but all proved unsuccessful.

#### 7.5.4 Alkylation of [N-Fmoc-S-(Trt)cysteinamido]-4-thioanisole (25)

Using methyl iodide in methanol no alkylation of [N-Fmoc-S-(Trt)cysteinamido]-4-thioanisole (25) was observed.

Aryl sulphides show less nucleophilic character than aliphatic sulphides and are therefore less readily alkylated than their aliphatic counterparts, thus catalysis of the reaction is required (Fig.37). Silver perchlorate is an effective catalyst for the alkylation of aromatic sulphides.<sup>89</sup>

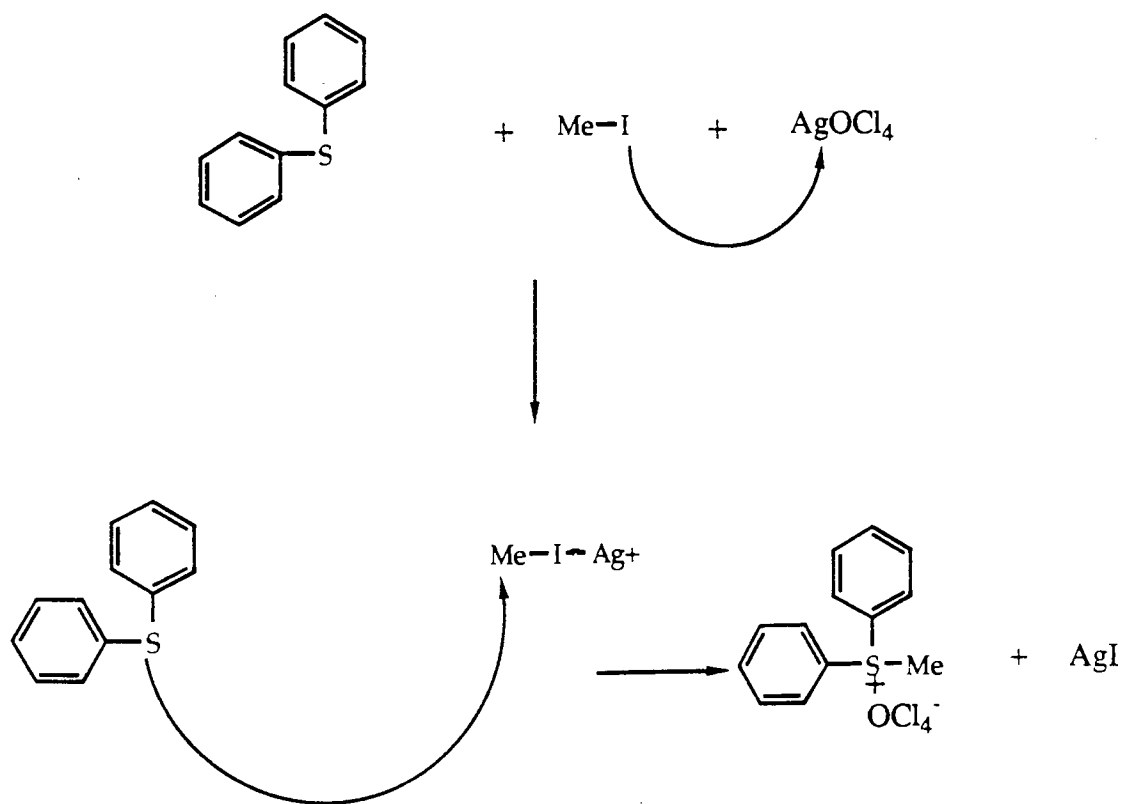


Figure 37 Mechanism of alkylation of aromatic sulphides by methyl iodide

The mechanism of catalysis involves the silver ion forming a complex with the alkyl halide resulting in a greater polarity being introduced to the carbon-halogen bond, an  $S_N2$  displacement occurs causing silver iodide to precipitate out of the reaction mixture, forming a sulphonium perchlorate salt. In the literature a reported 1.1 equivalents of methyl iodide and catalyst were used to alkylate aromatic sulphides.<sup>89</sup> Under these conditions the sulphonium salt of [N-Fmoc-S-(Trt)cysteinamido]-4-thioanisole was produced in 10% yield. On monitoring the reaction by TLC (acetic acid/butanol/water [12:5:3]) the reaction had not reached completion, starting material was still present. The addition of further equivalents of methyl iodide and silver perchlorate failed to result in an improved yield. The desired product [N-Fmoc-S-(Trt)cysteinamido]-S,S-dimethylphenylsulphonium perchlorate (**31**), was isolated by column chromatography to give a clear oil.

## 7.6 Deprotection of selected sulphonium-thiol molecules

The ultimate target for these syntheses was the generation of deprotected sulphonium-thiol molecules. N-CBZ-S-methylmethionyl-S-(Trt)cysteine methyl ester sulphonium iodide (**29**) was the first compound on which deprotection was attempted. Under acidic conditions cleavage of the trityl function usually occurs in peptides within a few hours. This time is increased if any steric hindrance exists. Sulphonium-thiol molecules offer a more challenging synthetic problem.

### 7.6.1 Cleavage of the trityl group from N-CBZ-S-methylmethionyl-S-(Trt)cysteine methyl ester sulphonium iodide (**29**).

As introduced earlier N-CBZ-S-methylmethionyl-S-(Trt)cysteine methyl ester sulphonium iodide (**29**) may possibly undergo an inter and/or intramolecular reaction on removal of the

trityl group. Earlier investigations (section 7.4.1) suggested that the likelihood of intermolecular reaction was small, but intramolecular reaction remains a possibility.

Therefore the deprotection reaction was monitored closely in order to determine:

- (i) Whether decomposition of the sulphonium centre occurred on cleavage of the trityl group.
- (ii) If decomposition occurred, the optimum incubation time required to achieve maximum cleavage with minimal decomposition.
- (iii) Whether the CBZ group remained intact under these cleavage conditions (the N-CBZ group is acid labile)

The reaction was monitored by  $^1\text{H-NMR}$ . An initial spectrum was recorded in  $\text{CD}_3\text{OD}$ . From the spectrum, the aromatic region shows both the CBZ and trityl groups to still be intact (Fig.38).

After addition of TFA and ethanethiol, the reaction was left for 15 mins before a spectrum was recorded. After this time trityl cleavage had occurred as shown by the collapse of the singlet at  $\delta 7.2\text{ppm}$ . The peaks at  $\delta 7.3$  and  $\delta 5.2\text{ppm}$  corresponding to the CBZ group appeared to be unchanged, suggesting no cleavage of the group had occurred. Following the reaction using  $^1\text{H-NMR}$ , the reaction appeared to have reached completion after 30 mins. The NMR study also confirmed that the CBZ protecting group was stable under the conditions employed. One further equivalent of trifluoroacetic acid (TFA) was added to the reaction mixture and after 30 mins a  $^1\text{H-NMR}$  spectrum was run (Fig.39). No change



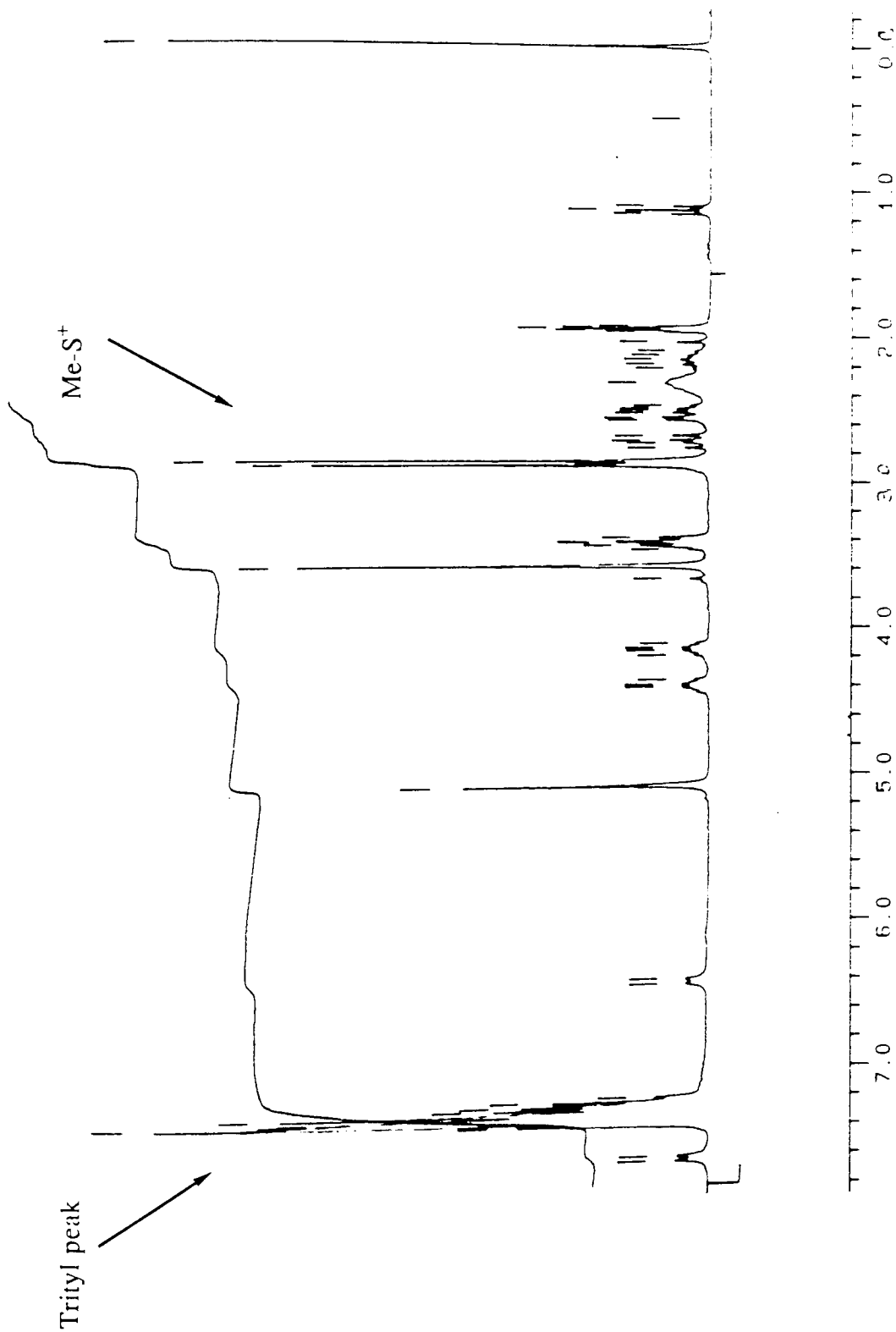


Figure 38  $^1\text{H}$  NMR spectrum of N-CBZ-S-methylmethionyl-S-(Trt)cysteine methyl ester sulphonium iodide (29)

was noted in the spectrum suggesting that complete cleavage of the trityl group had been achieved.

After standing overnight, analysis of the reaction mixture by NMR showed the disappearance of the sulphonium methyl group signals at  $\delta 2.9$  ppm suggesting that decomposition of the molecule had occurred.

The compound N-CBZ-S-methylmethionylcysteine methyl ester (**32**) was tested immediately following synthesis for biological activity to ensure no decomposition resulted.

#### 7.6.2 Cleavage of other sulphonium-thiol molecules

Cleavage of the trityl function of [N-Fmoc-S-(Trt)cysteinamido]-S,S-dimethylphenyl sulphonium perchlorate (**31**) was attempted. The reaction was carried out on a very small scale and the resulting NMR of the cleavage reaction proved inconclusive. This cleavage reaction needs to be studied further to determine whether the sulphonium-thiol molecule **31** can be isolated as a stable compound.

### 7.7 Biological testing

Both the simple sulphonium compounds and the sulphonium-thiol molecules were assessed for their ability to act as cytotoxic agents. The labelled compound S-[ $^3\text{H}$ -methyl]methionine sulphonium iodide was assessed for its ability to specifically alkylate the C5 position of cytosine. The results from these studies are discussed in Chapter 9.

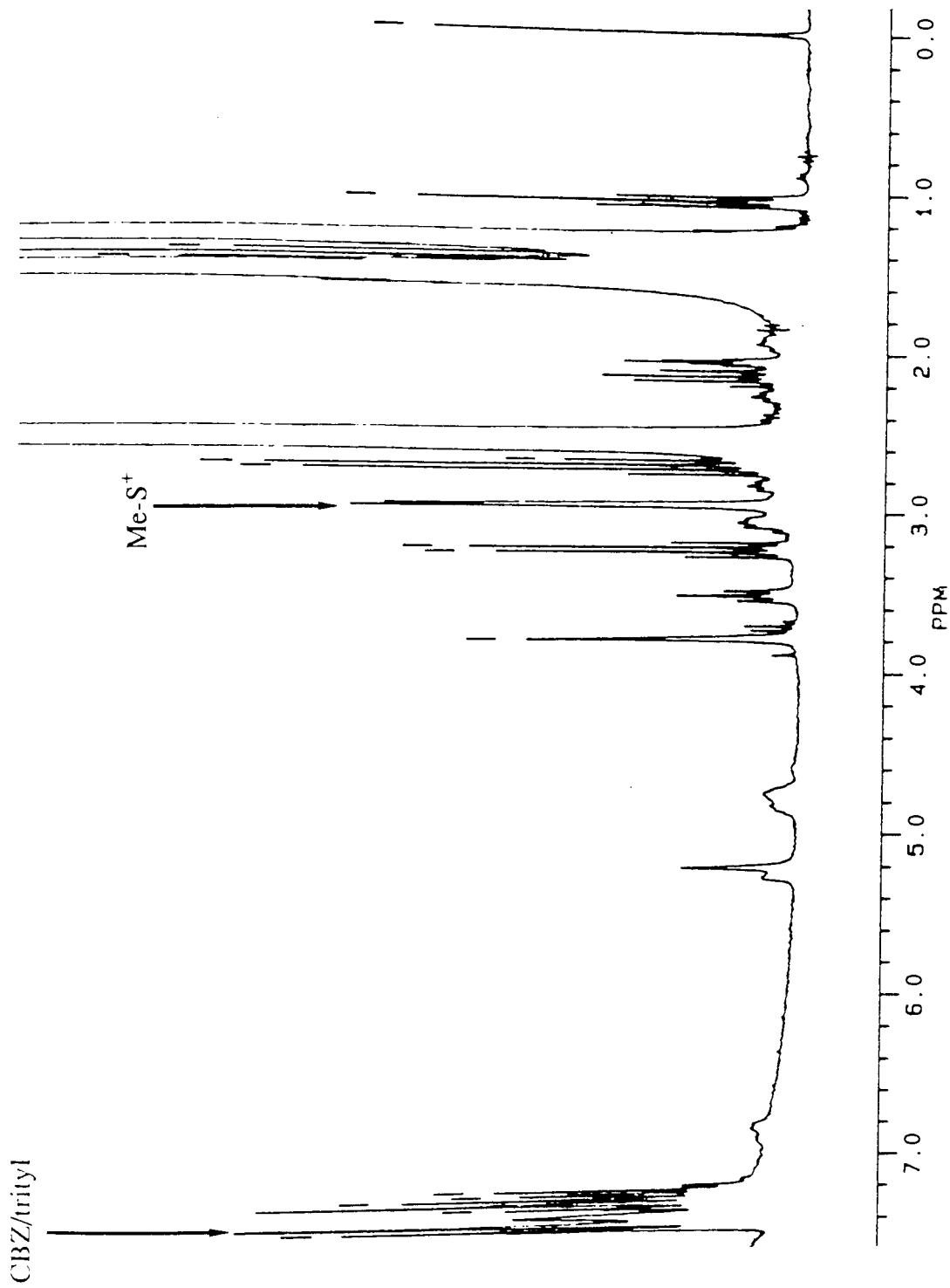


Figure 39 <sup>1</sup>H NMR of N-CBZ-S-methylmethionyl-S-(Trt)cysteine methyl ester

sulphonium iodide (29) after incubation for 30 mins with 95% TFA and anisole

## CHAPTER 8

### Synthesis of peptides to activate the 5 position of cytosine

#### 8.1 Introduction

A proposed hypothesis for the onset of cancer suggests that a deficiency of the enzyme DNA MTASE in malignant cells results in cells possessing hypomethylated oncogenes which are responsible for the malignant phenotype.<sup>91</sup> If the deficiency of DNA MTASE could be overcome by replacing it with another molecule capable of catalysing cytosine C5 methylation, malignant cell lines might revert to a normal phenotype. That a simple short peptide may be able to provide the catalytic mechanism for this activation and hence promote C5 alkylation is investigated in this chapter.

Restriction enzyme studies using S-methylmethionine sulphonium iodide (Chapter 5) suggested that the alkylation of the C5 position of cytosine by simple sulphonium compounds may be catalysed by DNA MTASE. By using a simple sulphonium compound and a molecule capable of replacing DNA MTASE, it may be possible to replace the natural methylating system. The use of a simple thiol as the activating species was considered in Chapter 7, and efforts to produce a more complex system to perform this function were also investigated and are discussed in Chapter 8.

At present very little is known about the eukaryotic DNA MTASE. The enzyme thymidylate synthetase is thought to have a similar mechanism and active site structure.<sup>11</sup> Thymidylate synthetase therefore was chosen as a starting point to direct syntheses of a possible activating molecules. The crystal structure of thymidylate synthetase from *E. coli* was published but to only poor resolution and therefore provided little three dimensional detail about the enzyme's active site.<sup>92</sup> However in a report by Demple *et al.*, which presented the sequence of the enzyme O<sup>6</sup>-guanine DNA alkyltransferase, it was noticed that

the active site sequence possessed a very similar sequence to that of the active site of thymidylate synthetase.<sup>93</sup> Demple and colleagues suggested that the sequence Pro-Cys-His with hydrophobic residues (eg. Phe, Ile) on the amino terminal side may provide a chemical mechanism for the activation of cysteine (Table 12).

Table 12 Prokaryotic methylene transfer enzymes.

ENZYME	SOURCE	ACTIVE SITE SEQUENCE
O <sup>6</sup> -MeG-DNA MTASE	<i>E.coli</i>	Ile- <u>Pro-Cys-His</u> -Arg
Thymidylate Synthetase	<i>E.coli</i>	Ala- <u>Pro-Cys-His</u> -Ala
Thymidylate Synthetase	T4 phage	Pro- <u>Pro-Cys-His</u> -Met
Thymidylate Synthetase	<i>L.casei</i>	Pro- <u>Pro-Cys-His</u> -Thr
Thymidylate Synthetase	Yeast	Pro- <u>Pro-Cys-His</u> -Ile

The histidine residue may play a role by acting as a proton acceptor as seen in the enzyme papain.<sup>94</sup> A number of prokaryotic DNA MTASE's have now been sequenced,<sup>95</sup> and recently it has been suggested they may share a common evolutionary origin.<sup>93,94</sup> Wilson identified that in DNA MTASE enzymes Pro-Cys-Gly was a highly conserved sequence and thus this tripeptide was identified as a synthetic target. Wilson suggested that the conserved sequence Pro-Cys-Gly must be in very close proximity to the DNA MTASE's active site.<sup>63</sup> Pro-Cys-His and Pro-Cys-Gly were therefore both potential synthetic target molecules. However the Pro-Cys dipeptide was the one consistent feature of both types of alkyltransferase and therefore this dipeptide was the starting point on which investigation was concentrated.

## 8.2 Synthesis of target molecules

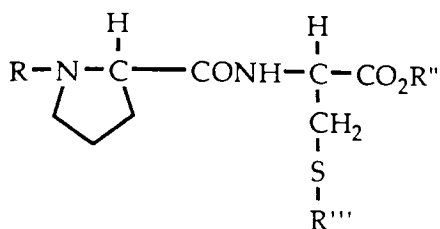
### 8.2.1 Synthesis of Pro-Cys derivatives

The dipeptide Pro-Cys was identified as an initial synthetic target. A number of protected derivatives were synthesised (Table 13).

Table 13 Conditions for the synthesis of prolycysteine derivatives

Compound		Coupling agent	Yield(%)
N-CBZ-prolyl-S-(Bzl)cysteine methyl ester	(5)	Ethyl chloroformate	12
N-CBZ-prolyl-S-(Bzl)cysteine ethyl ester	(33)	DCC/HOBT	51.4
N-Boc-prolyl-S-(Bzl)cysteine methyl ester	(34)	DCC/HOBT	34.9
Bis(N-Fmoc-prolyl)cystine dimethyl ester	(35)	EDCI	45
Bis(N-Boc-prolyl)cystine dimethyl ester	(36)	DCC/HOBT	35.4

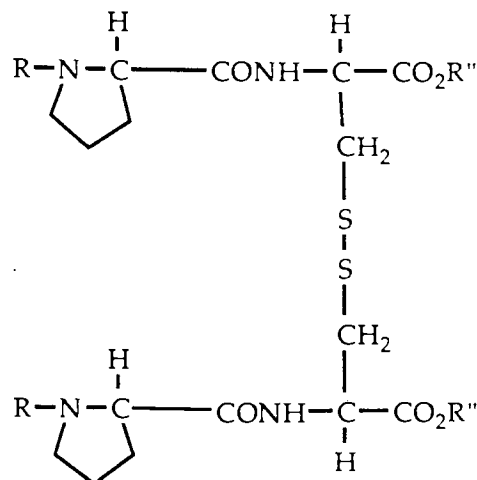
For structures see (Fig.40)



(5) R=CBZ R''=methyl R'''=Bzl

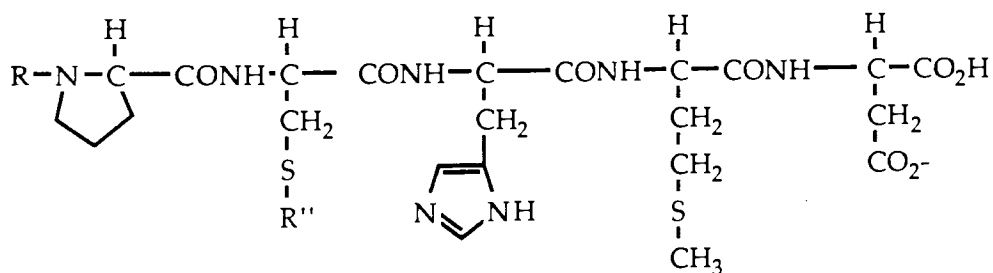
(33) R=CBZ R''=ethyl R'''=Bzl

(34) R=Boc R''=methyl R'''=Bzl



(35) R= Fmoc R''= methyl

(36) R= Boc R''= methyl



(38) R=Fmoc R''= trityl

(37) R=H R''= H

Figure 40 Structures of protected derivatives of prolinylcysteine peptides synthesised

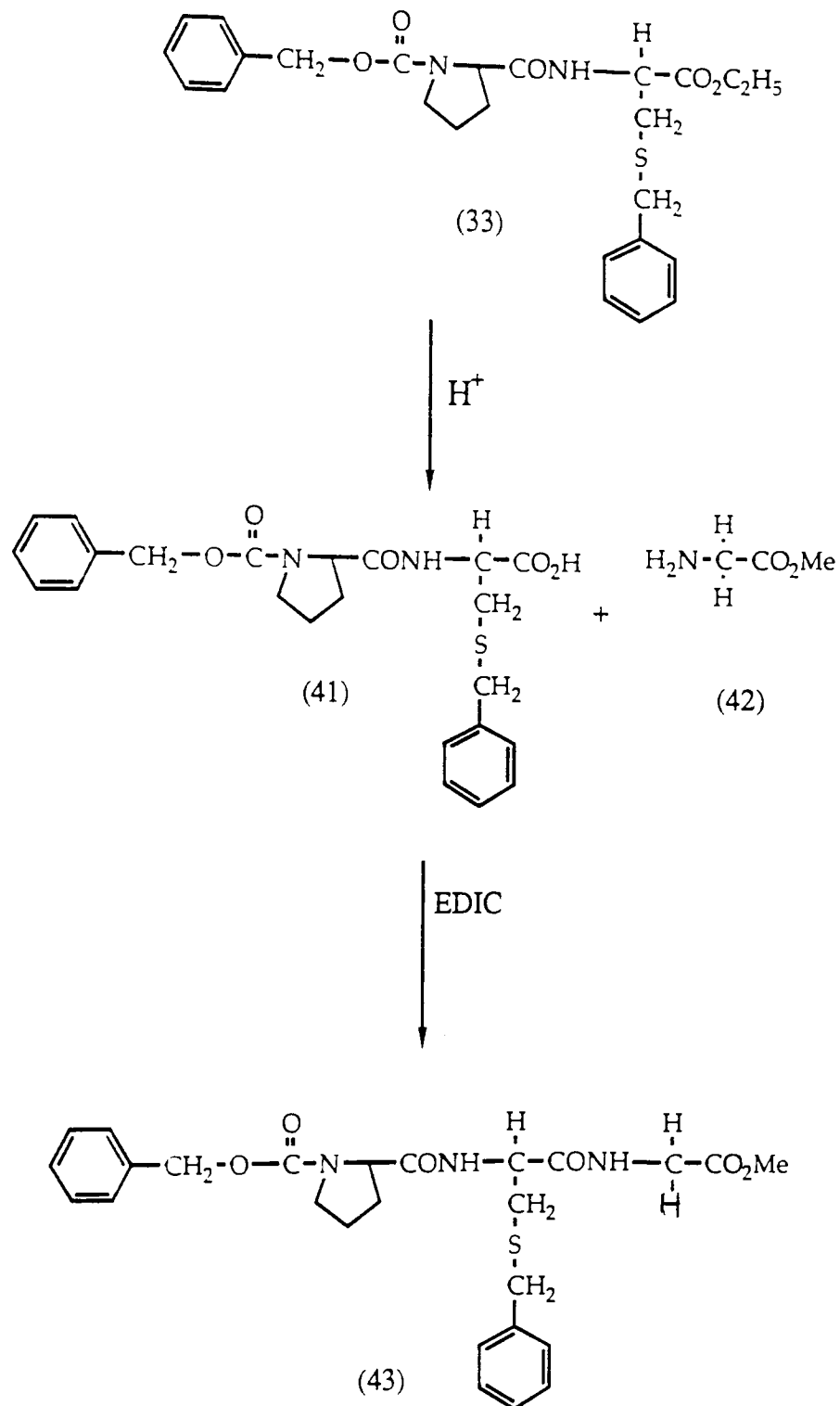
Initially the CBZ group was used to protect the amino function of proline.<sup>96</sup> The thiol function of cysteine was protected as the benzyl derivative with the carboxyl function protected as the methyl ester. Coupling was attempted using ethyl chloroformate, purification by flash chromatography gave a 12% yield of N-CBZ-prolyl-S-(Bzl)cysteine methyl ester (**5**). Use of DCC as the coupling agent proved more successful. The ethyl ester of N-CBZ-prolyl-S-(Bzl)cysteine (**33**) and the methyl ester of N-Boc-prolyl-S-(Bzl)cysteine (**34**) were isolated following recrystallisation in 51 and 35% yield respectively. **34** was synthesised using DCC coupling in the presence of the catalyst HOBT.<sup>97</sup> Purification by flash chromatography resulted in a pale yellow oil. The S-Bzl group of (**34**) was removed with sodium in liquid ammonia resulting in a 60% yield of deprotected product, N-Boc-prolylcysteine methyl ester (**39**).

#### 8.2.2 Synthesis of N-Boc-prolyl-(S-N-methylcarbamoyl(Nmc))cysteinylglycine ethyl ester (**40**)

A number of routes were attempted to synthesise the tri-peptide Pro-Cys-Gly. The first route involved the hydrolysis of the carboxy ester of N-CBZ-prolyl-S-(Bzl)cysteine ethyl ester (**33**) to give N-CBZ-prolyl-(S-Bzl)cysteine (**41**), followed by coupling to glycine methyl ester (**42**) to give N-CBZ-prolyl-(S-Bzl)cysteinylglycine methyl ester (**43**) (see scheme 6).

Initially hydrolysis of **33** was attempted under acidic conditions using 2M H<sub>2</sub>SO<sub>4</sub> in tetrahydrofuran (THF). The reaction reached 30% completion as assessed by TLC. Hydrolysis of the same compound was attempted using the enzyme  $\alpha$ -chymotrypsin in 0.05M ammonium acetate/ acetic acid adjusted to pH =7.8. The product precipitated out as a white solid immediately which was collected by filtration. The reaction however went to only 10% completion as assessed by TLC.



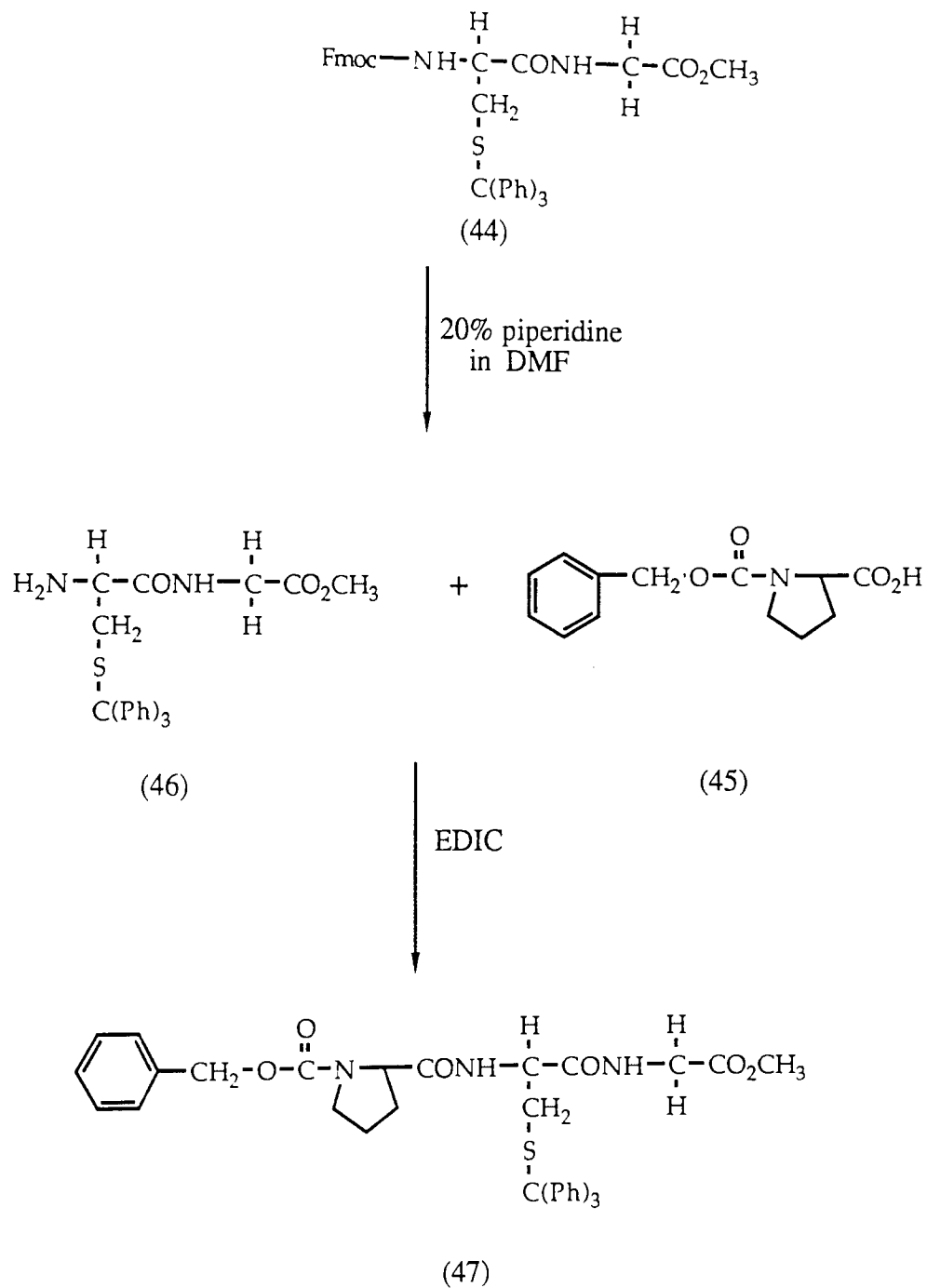


Scheme 6 Synthesis of N-CBZ-prolyl-(S-Bzl)cysteinylglycine methyl ester (43)

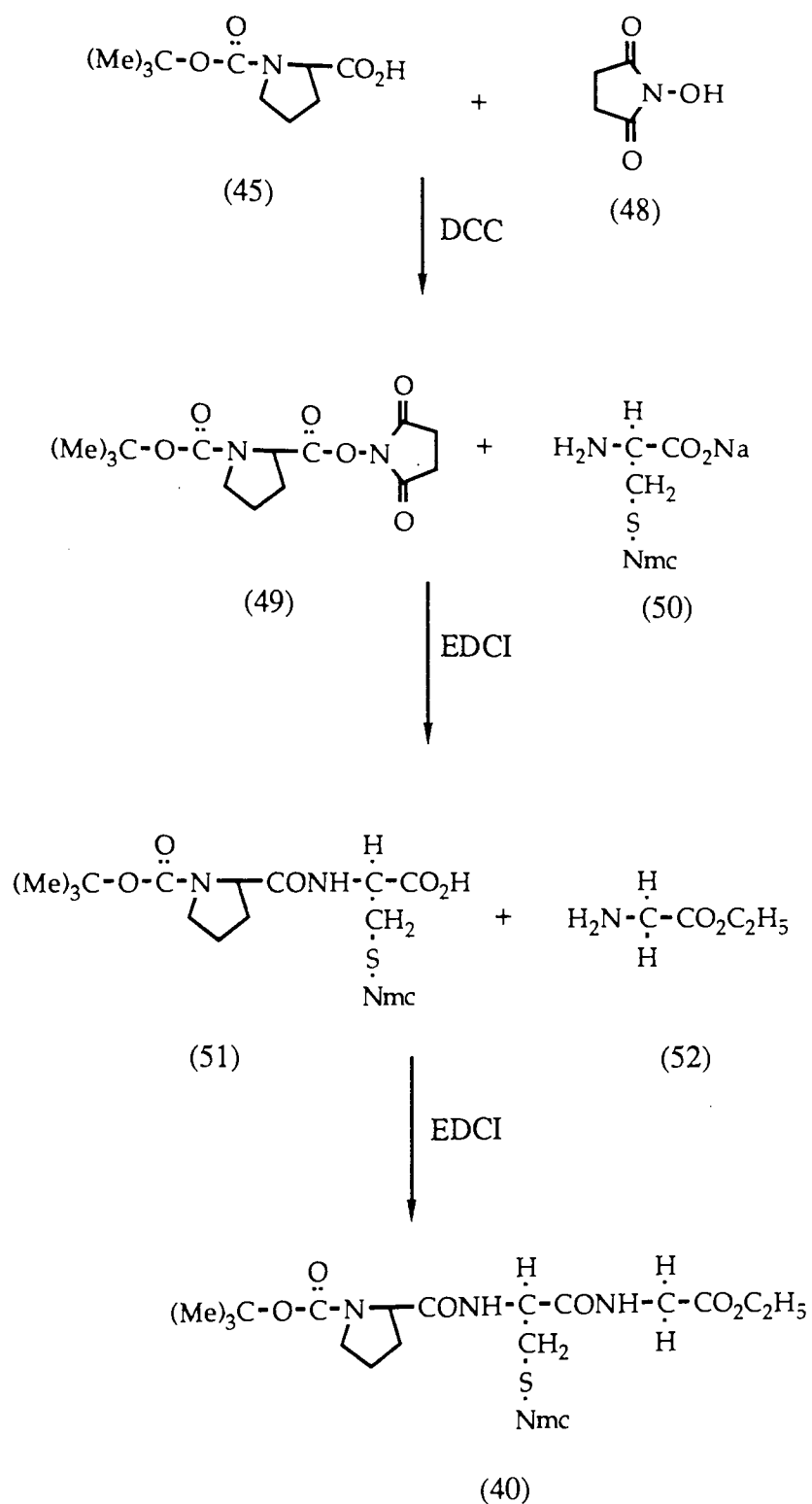
A second route to the product (scheme 7) included the synthesis of the dipeptide Cys-Gly followed by amino deprotection and subsequent coupling to amino-protected proline. The Fmoc amino protecting group was employed to protect the amino function of S-(Trt)cysteinyglycine ethyl ester. N-Fmoc-S-(Trt)cysteinyglycine ethyl ester (**44**) was synthesised using EDIC in a reaction catalysed by HOBT.

The Fmoc protecting group can be cleaved by a number of bases, commonly 20% piperidine in dimethylformamide (DMF) is used.<sup>99</sup> This method has found wide application in solid phase peptide synthesis. The dipeptide derivative **44** was allowed to stand in a solution of 20% piperidine in DMF for 2 hrs. S-(Trt)cysteinyglycine ethyl ester (**46**) was isolated in approximately 30% yield. This yield was considered poor, in light of the further reaction in the synthetic scheme.

As a result of the problems of deprotection reactions in these syntheses, an alternative approach was explored. A third possible synthetic route involved performing the first coupling under aqueous conditions (scheme 8). N-Boc-proline (**45**) was initially converted to the N-hydroxysuccinimide ester (**49**) using DCC. The active ester **49** was then coupled with S-(N-methylcarbamoyl(Nmc)) cysteine using EDIC in 5% NaHCO<sub>3</sub>, to give N-Boc-prolyl-S-(Nmc)cysteine (**51**) in 75% yield. The S-trityl derivative was replaced by the S-(Nmc) derivative in this synthesis due to the poor solubility of the former in 5% NaHCO<sub>3</sub>. Coupling of **51** with glycine ethyl ester (**52**) was achieved using EDIC. The tripeptide N-Boc-prolyl-S-(Nmc)cysteinyglycine ethyl ester (**40**) was isolated from a solution of ethyl acetate in 30% yield.



Scheme 7 Synthesis of N-CBZ-prolyl-S-(Trt)cysteinylglycine methyl ester (47)



Scheme 8 Synthesis of N-Boc-prolyl-(S-Nmc)cysteinylglycine ethyl ester (40)

Attempts at crystallisation of this compound for structural studies proved initially unsuccessful. Future work should concentrate on determining the crystal structure of this compound.

### 8.3 Design and synthesis of complex peptides with potential C5 activating ability

The biological results reported in Chapter 9 suggested that the short peptides described in section 8.2 were incapable of activating the C5 position of cytosine toward alkylation by simple sulphonium compounds. In addition molecular modelling studies (see Chapter 6) also suggested that short peptides (eg.2-5 amino acids long) based around the Pro-Cys motif would not exhibit a consistent 3 dimensional structure and would therefore be unlikely to mimic the active site of DNA MTASE. In 1988 some more information became available about a eukaryotic DNA MTASE. Bestor *et al.* determined the complementary DNA sequence of mouse DNA MTASE, from which it was possible to determine the amino acid sequence of the enzyme.<sup>69</sup> The enzyme sequence was found to contain only 3 proline residues that were adjacent to cysteine residues in the 175Kda enzyme. At the same time, Lauster *et al.* compared the sequence of the mammalian enzyme with the sequences of 11 bacterial enzymes. The major difference between the bacterial and mammalian enzyme was in size. The mammalian enzyme is approximately 9 times larger than its bacterial counterparts. Therefore it is likely that a large part of the mammalian enzyme is associated with recognition and DNA binding. The bacterial enzymes possessed 4 large regions which exhibited good homology.<sup>16</sup> Lauster suggested regions in the mammalian MTASE that he considered were equivalent to the four homologous regions identified in the bacterial enzymes. Studies by Wu and Santi identified cysteine 81 (see Table 14)

Table 14 Active site regions in DNA MTASEs

MTASE	70	(81)	94
Mouse	LQLPHLLQVQK(L)PSVSFL		SYC
Mouse	RASACVCA	P(C)LPAERCGDDKKFVS	
11 <sub>s</sub>	PEFDLLBGGSP(C)QSFSVAGYRKGFE		
Spr	PEFDLLBGGSP(C)QSFSVAGHRKGFE		
3T	PYFDLLTSGFP(C)PTFSVAGGRDGME		
BspRI	PSANLVIGGFP(C)PGFSEAGPRLV D		
EcoRI	PDHDVLAGFPP(C)QPFSLARAHGFE		
DdeI	PDHDILCAGFP(C)QAFSISGKQKGF		
MspI	PQHDILCAGFP(C)QPF SHIGKREGFE		

in the bacterial enzymes as being directly involved in the activating mechanism for methyl group transfer.<sup>100</sup> The block of residues between positions 78 and 88 contains 6 residues that are identical in all prokaryotic MTASEs. This Pro-Cys conserved motif is not found in N<sup>4</sup>-cytosine MTASEs or N<sup>6</sup>-adenine MTASEs, which suggests the Pro-Cys dipeptide is more important in carbon MTASEs. The amino acid sequence of the active site region in the mouse (eukaryotic) enzyme that was proposed by Lauster *et al.*, did not compare well with the amino acid sequences of the active sites of the bacterial enzymes. It would be slightly premature to surmise that the Pro-Cys region suggested by Lauster *et al.* was the actual Pro-Cys motif located in the active site of the mammalian enzyme. Therefore all three Pro-Cys sequences within the mammalian enzyme provided potential synthetic targets. Syntheses of peptides of up to 10 amino acids in length were undertaken. It was

hoped that peptides of this length might adopt more consistent three dimensional conformations. Four main target sequences were identified (Table 15):

Table 15 Target peptide sequences

- 
- a) Cys-Ala-Pro-Cys-Leu-Pro-Ala-Gly (53)
  - b) Gly-Phe-Pro-Cys-Pro-His-Phe-Ser (54)
  - c) Ser-Pro-Cys-Met-Arg-Leu-Ile-His-Leu (55)
  - d) Pro-Cys-His-Met-Asp(37)
- 

Peptides 53 and 55 encompassed Pro-Cys motifs present in the mammalian enzyme. Peptide 54 was taken from the active site sequence of the bacterial enzyme BspRI. Peptide 37 was an experimental sequence.

### 8.3.1 Methodology

Syntheses of the target peptides 53-55 and 37 were carried out on an Applied Biosystems 430A automated solid phase peptide synthesiser. The solid support was a polystyrene resin 1% cross-linked with divinyl benzene which used the acid labile benzyloxybenzyl alcohol as the linker molecule (Fig.41).

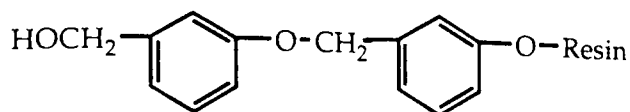


Figure 41 The benzyloxybenzylalcohol linker molecule, an acid cleavable linker molecule for solid phase peptide synthesis

The initial amino acid can be supplied already linked to the resin or can be coupled manually using the coupling agent diisobutylcarbodiimide (DIC). Pre-loaded resin was preferred in the synthesis of peptides **53** to **55**. For the synthesis of the smaller peptide Pro-Cys-His-Met-Asp (**37**) the proline residue was coupled to the resin manually.

In solid phase synthesis, it is vital to ensure that each coupling reaction proceeds in yields approaching 100%, this was achieved by using two coupling cycles (Fig.42). Symmetrical anhydrides react more efficiently with a free amino group than an activated ester, therefore the first cycle involved the reaction of two equivalents of Fmoc amino acid in the presence of DIC to form a symmetrical anhydride (Pathway I, Fig.43). The anhydride was then reacted with the free amino group of the resin bound amino acid. In order to ensure the degree of coupling was maximised a third equivalent of Fmoc-amino acid was then reacted with the resin bound peptide, under standard DIC coupling conditions catalysed by HOBT (pathway II, Fig.43).

It was not possible to monitor the degree of coupling directly. However the progress of the synthetic reaction could be followed by monitoring the removal of the Fmoc group. Practically this was achieved by using an on-line system which collected the deprotection wash and then passed it through a flow cell of a UV spectrophotometer. The degree of cleavage could be estimated from the UV absorbance due to the chromophore of the Fmoc group ( $\lambda_{\text{max}}$  254nm). Removal of the Fmoc group is seen as a peak in the UV spectrum



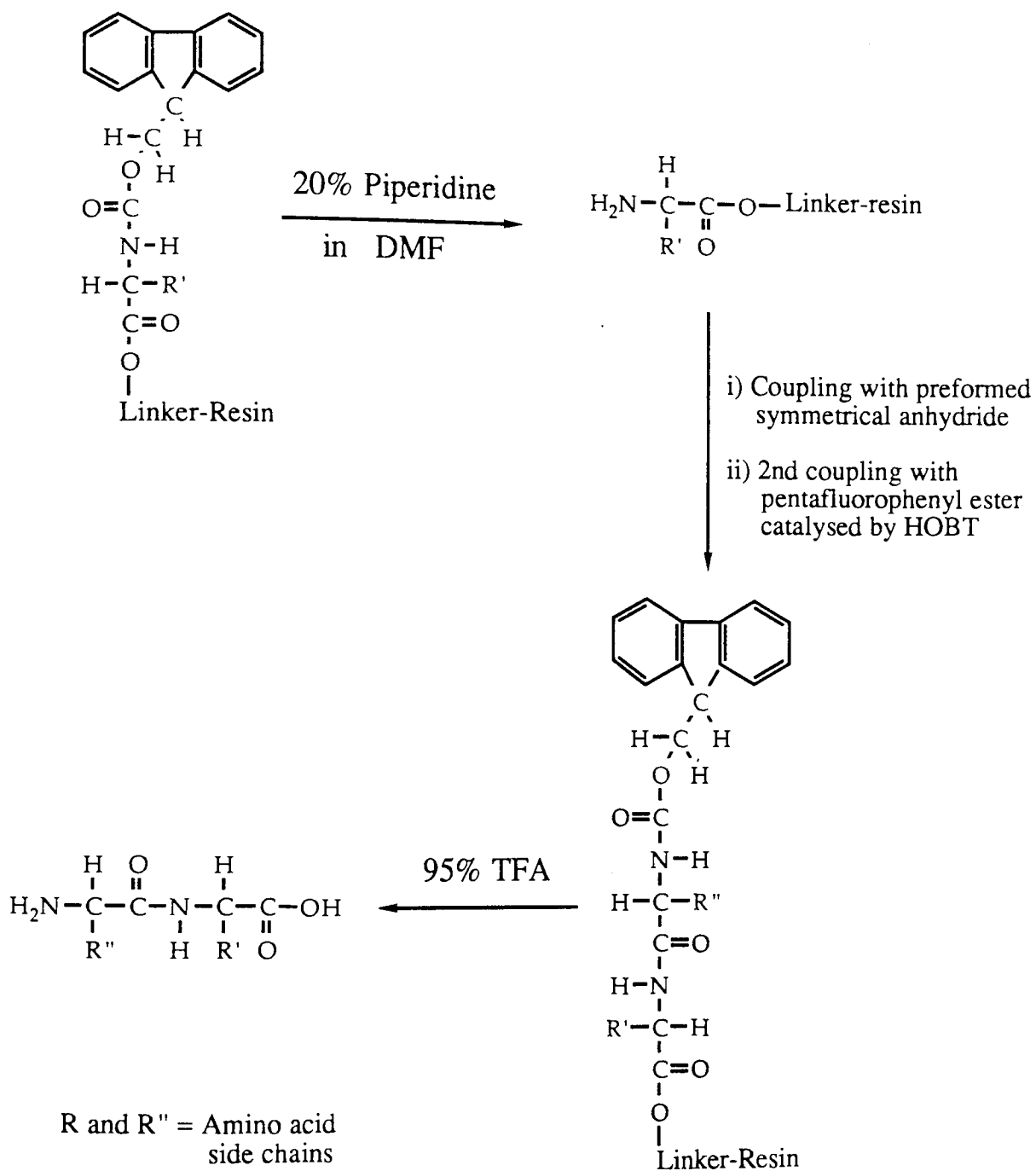


Figure 42 Method flow diagram for solid phase synthesis of peptides 53-55 and 37

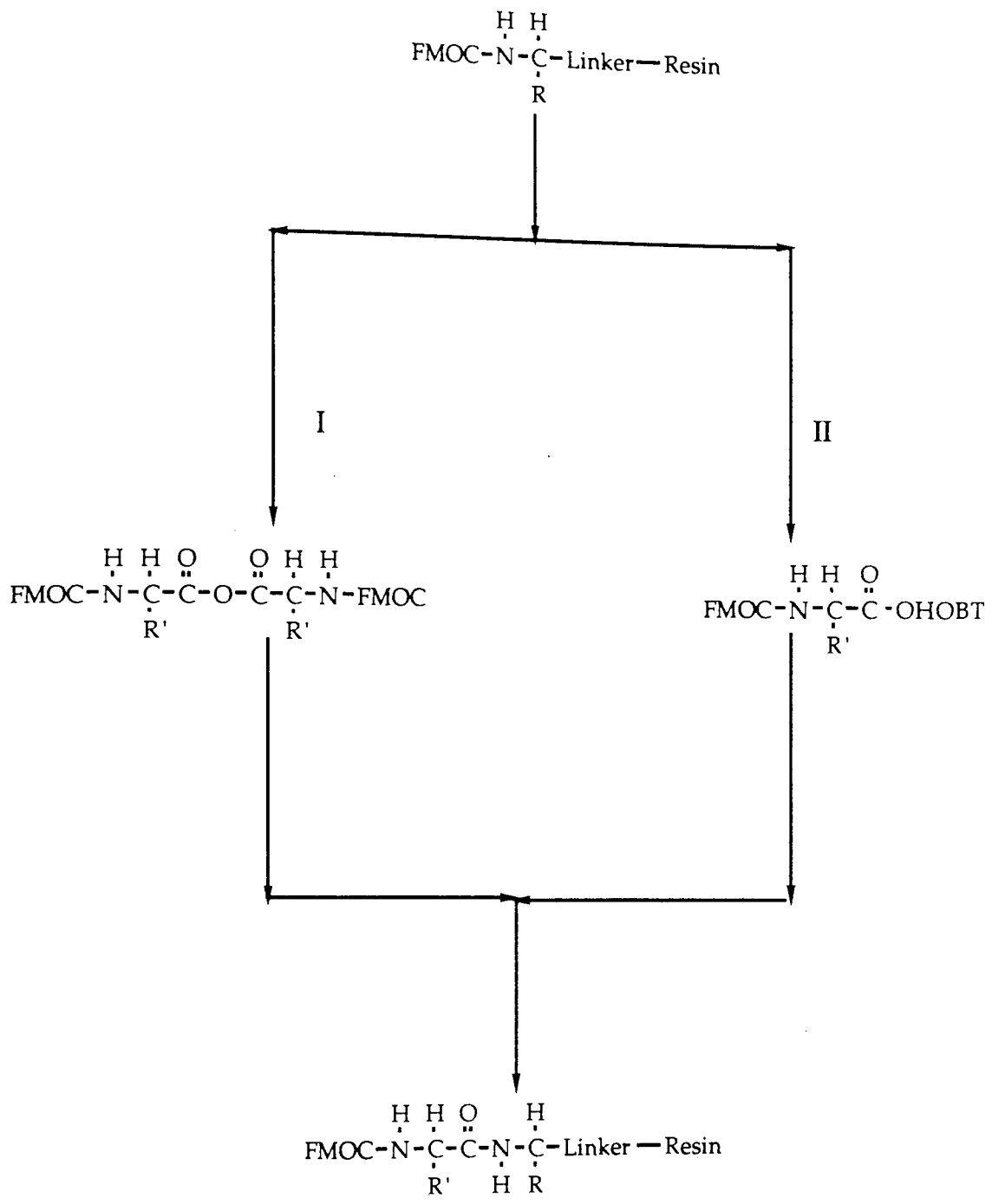


Figure 43 Synthetic pathways for the solid phase synthesis of peptides **53-55** and **37** on an Applied Biosystems 430A peptide synthesiser.

of the cleavage reaction. Deprotection was carried out using 20% piperidine in DMF and was repeated twice to ensure complete removal of the Fmoc group.

A possible problem with solid phase synthesis is that if reaction is incomplete, any remaining free amino groups are available for reaction at the next coupling. The possibility of a small amount of peptide being formed containing 1 less amino acid than the majority of product would provide great practical difficulties at the purification stage. In order to prevent this after every coupling stage, the resin bound peptide is incubated with acetic anhydride to ensure any free amino groups are acetylated and hence rendered unreactive.

### 8.3.2 Synthesis of target peptides

For the synthesis of peptide **53**, inspection of the UV profile of removal of the Fmoc protecting group (Fig.44) showed that all cleavages were successful. Deprotection of the amino function did become slightly slower as the synthesis progressed but the difference was insignificant.

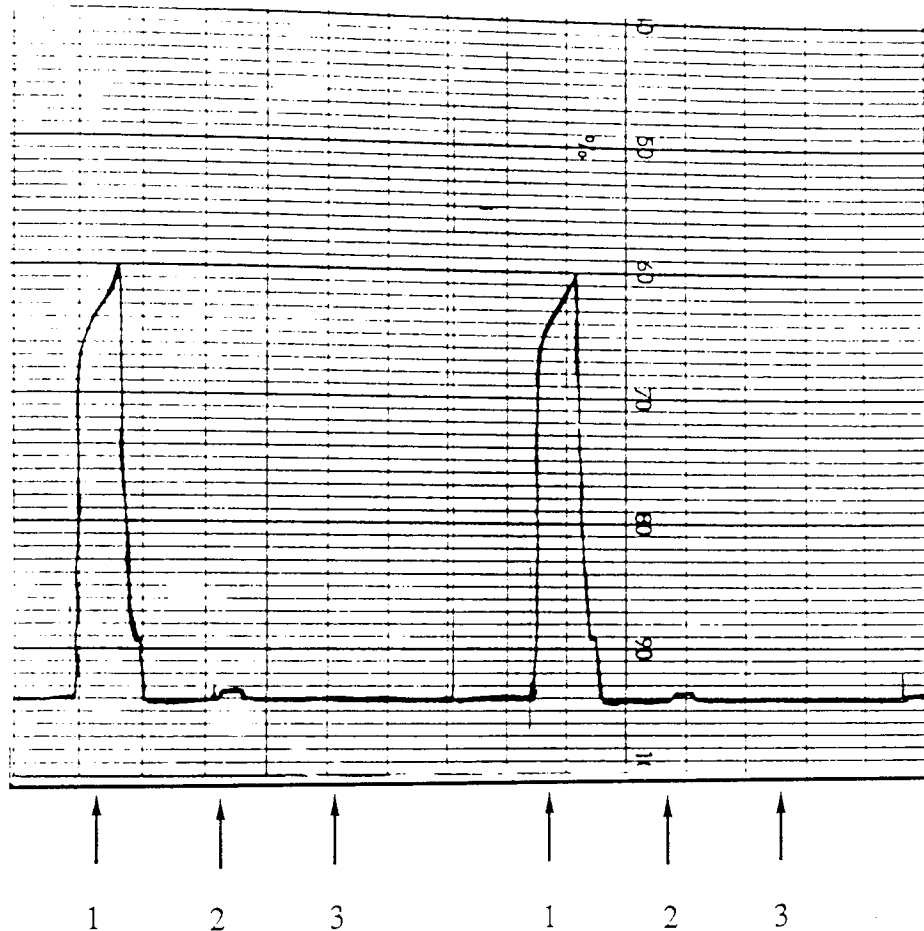
In the synthesis of peptide **54**, the fifth deprotection which was undertaken after the coupling of a trityl protected cysteine residue appeared to be slightly hindered. The UV deprotection profile showed that the Fmoc protecting group was not completely removed after 2 washes. However the small hindrance of the deprotection was unlikely to significantly affect the efficiency of the synthesis.

The synthesis of peptide **55** proceeded smoothly, no difficult couplings were apparent from the UV deprotection profile.

UV absorbance  
at 254nm

2.0

0



Deprotection wash number

Figure 44 UV deprotection profile, monitoring the removal of the Fmoc group after peptide coupling during the solid phase synthesis of peptide 53

The final peptide synthesised by solid phase methods was peptide 37. In the synthesis of this peptide the initial amino acid, aspartic acid was coupled to the resin manually. The acidic side chain of aspartate was protected as the t-butyl ester (tBu).

Manual coupling of the resin linker molecule benzyloxybenzylalcohol to N-Fmoc-O-(tBu)aspartate (**56**) was achieved by initially forming a symmetrical anhydride of **56** using DIC. The anhydride was then reacted with resin-bound benzyloxybenzylalcohol in DMF in the presence of the catalyst dimethylaminopyridine (DMAP). The resulting product was isolated by filtration. The loading of the resin (ie. the amount of resin bound amino acid actually present) was assessed by comparing a UV scan of a solution of Fmoc-OH of known concentration to that shown by the resin bound amino acid. The absorbance produced by the resin bound Fmoc amino acid was related to the amount of resin bound amino acid that was actually present. A loading of 84% was achieved. Any free benzyloxybenzylalcohol remaining was rendered unreactive by reacting the crude resin with benzyl chloride in a reaction catalysed by pyridine. After washing with both dichloromethane and ether, the resin was thoroughly dried under vacuum overnight. The loaded resin was then ready for automated coupling.

The cysteine and histidine residues in this peptide were protected as the S-(trityl) and N-(trityl) derivatives respectively. The coupling reactions were carried out as for the other peptides.

On cleavage the acetylated biproducts are easily separated from the free peptide. After the synthesis was complete the resin was washed with DMF, chloroform and ether. The peptide was then ready for deprotection.

## 8.4 Cleavage and deprotection of peptides 53-55 and 37

Removal of the peptide from the resin and the subsequent removal of the protecting groups can be achieved by a number of routes depending on the methodology employed. It was desired in the syntheses of target peptides to isolate free peptide after the cleavage reaction and therefore the methodology and protecting group strategy chosen reflected this fact. For all the peptides synthesised, cleavage from the resin and deprotection of the peptide was to be carried out in a single step. The benzyloxybenzyl alcohol linker molecule is readily cleaved by 95% TFA in under 1 hour.<sup>101</sup> Complete cleavage and deprotection of peptides **53-55** and **37** was achieved using 95% TFA.

### 8.4.1 Cleavage and deprotection of Cys-Ala-Pro-Cys-Leu-Pro-Ala-Gly (**53**)

Peptide **53** contained two cysteine residues which required protection prior to reaction. The methodology employed required cleavage from the resin and removal of protecting groups to occur under acidic conditions. The trityl group appeared most suitable protecting group to be employed.<sup>102</sup> A possibility exists that on removal of the trityl groups from the two cysteine residues an intramolecular reaction may occur to form a disulphide bond. Therefore a reducing agent, ethyl methyl sulphide was required in the deprotection mixture. The cleavage reaction is reversible,<sup>101</sup> attributable to the fact that the trityl group is capable of forming a highly stabilised cation. The problem was overcome by using the scavenger molecule anisole. The cleavage/deprotection mixture consisted of 95% TFA, anisole, and ethylmethyl sulphide (91:4.5:4.5).

HPLC was used to monitor the deprotection reaction. After 1 hour a small peak was observed at retention time ( $R_t$ ) = 15 minutes (mins) (20%), with a much larger peak being present at  $R_t$  = 18mins (75%). After 4 hrs the peak at  $R_t$  = 15mins had grown to become

the only significant peak (80%) (Fig.45). Further monitoring of the reaction by HPLC showed no further increase in this peak.

Therefore the optimum time for complete deprotection was considered to be 4 hrs. When deprotection was complete, the mixture was filtered to remove the insoluble polymeric resin. The resin was washed firstly with DMF, then chloroform. The solvent was removed *in vacuo* to yield an oily residue which on trituration with ether resulted in a white fluffy precipitate. At this stage the peptide was impure and required purification. Purification of peptide **53** proved to be difficult. The initial purification step commonly employed in peptide purification is gel filtration. An aqueous solution of the crude peptide was passed down a sephadex G-25 column at 4<sup>0</sup>C with a mobile phase of 30% acetic acid (flow rate 0.5ml/min). The resulting UV and polarimetry profiles showed 3 major broad peaks. Only one of these peaks showed any optical activity. HPLC analyses of the three peaks revealed that each was a complex mixture of compounds indicating decomposition had occurred, and so this purification step was by-passed. Peptide **53** possessed no charged residues and therefore the use of ion-exchange chromatography, another commonly used peptide purification technique was not possible. Purification by HPLC seemed the most likely method to achieve rapid and efficient purification of peptide **53**, with a minimum amount of decomposition. Preparative HPLC was therefore evaluated as a purification technique. Separations were attempted employing both C18 and C8 RP columns using a mobile phase of CH<sub>3</sub>CN (0.01% TFA)/water (0.01% TFA) employing a linear gradient. The RP C8 column was slightly more efficient at peak resolution and therefore was adopted for use in all subsequent HPLC separations. Preparative HPLC of the crude peptide **53** revealed two significant peaks and a number of minor peaks. The first major peak at  $R_t = 12.23$ mins was thought to be the most likely to be produced by free peptide. The second major peak possessed a  $R_t = 15.0$ mins and was most likely to be

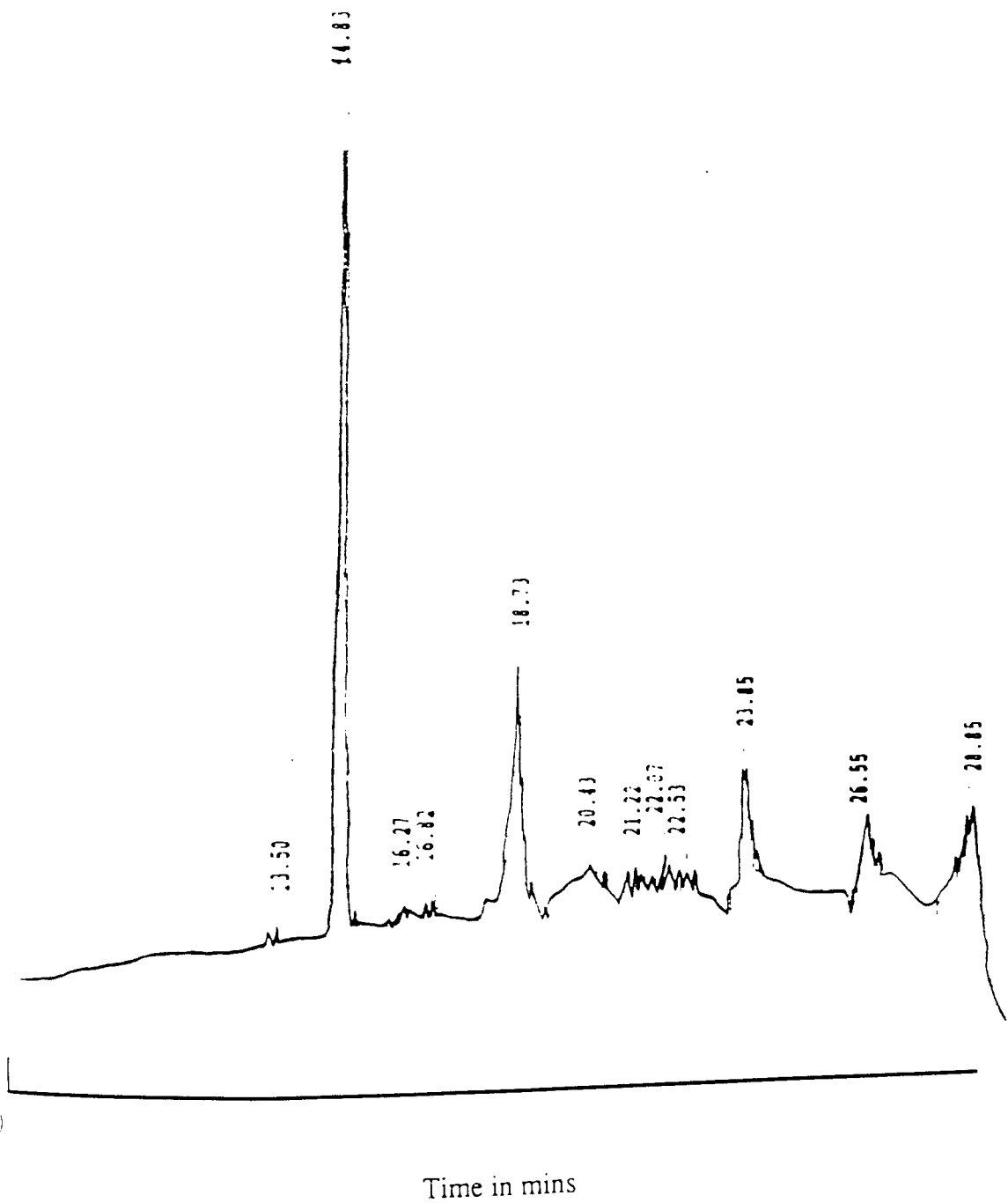


Figure 45 HPLC profile of crude peptide **53** after 4 hrs incubation of the resin with 95% TFA:anisole: ethylmethyl sulphide (91:4.5:4.5)



produced by peptide still possessing some protection or possibly a dimer of peptide **53** produced by the formation of a disulphide bridge. The peak at  $R_t = 12.3$ mins was collected and reduced in volume *in vacuo* and then lyophilised to dryness. The peptide was rechromatographed on an analytical scale using the same conditions. The chromatogram showed the sample consisted of more than one component, with significant peaks being present at short retention times. It appeared that the free peptide was unstable. The decomposition may have been a result of oxidation. Further cleavage reactions performed under nitrogen, followed by isolations on a preparative scale failed to result in pure peptide. This suggested that peptide may be unstable in a free form. Analysis of the peak  $R_t = 12.3$ mins from the preparative HPLC column by  $^1\text{H}$  NMR and by FAB mass spectrometry suggested that free peptide was present in the isolate. The peak at  $R_t = 15$ mins was isolated and was identified as the trityl protected derivative of peptide **53** (**57**).

#### 8.4.2 Cleavage and deprotection of Gly-Phe-Pro-Cys-Pro-His-Phe-Ser (**54**)

Peptide **54** possessed a histidine and a cysteine protected residue prior to cleavage. Both were protected as the trityl derivative. Deprotection was carried out as for **53** using 95% TFA, anisole and ethylmethyl sulphide. HPLC was used to monitor the deprotection and after one hour revealed a minor peak  $R_t = 14.4$ mins, and a large peak at  $R_t = 15$ mins. After 3 hrs, the peak  $R_t = 14.4$ mins had become the major peak. After 4 hrs the peak at  $R_t = 15$ mins was relatively insignificant compared to the large peak at  $R_t = 14.4$ mins. From these results it was assumed that the peak at  $R_t = 15$ mins was produced by protected peptide, whereas the peak at  $R_t = 14.4$ mins was produced by free peptide (Fig.46).

Due to the problems experienced with the isolation of peptide **53**, attempts to purify peptide **54** were limited to HPLC techniques. Initially method development was performed on an

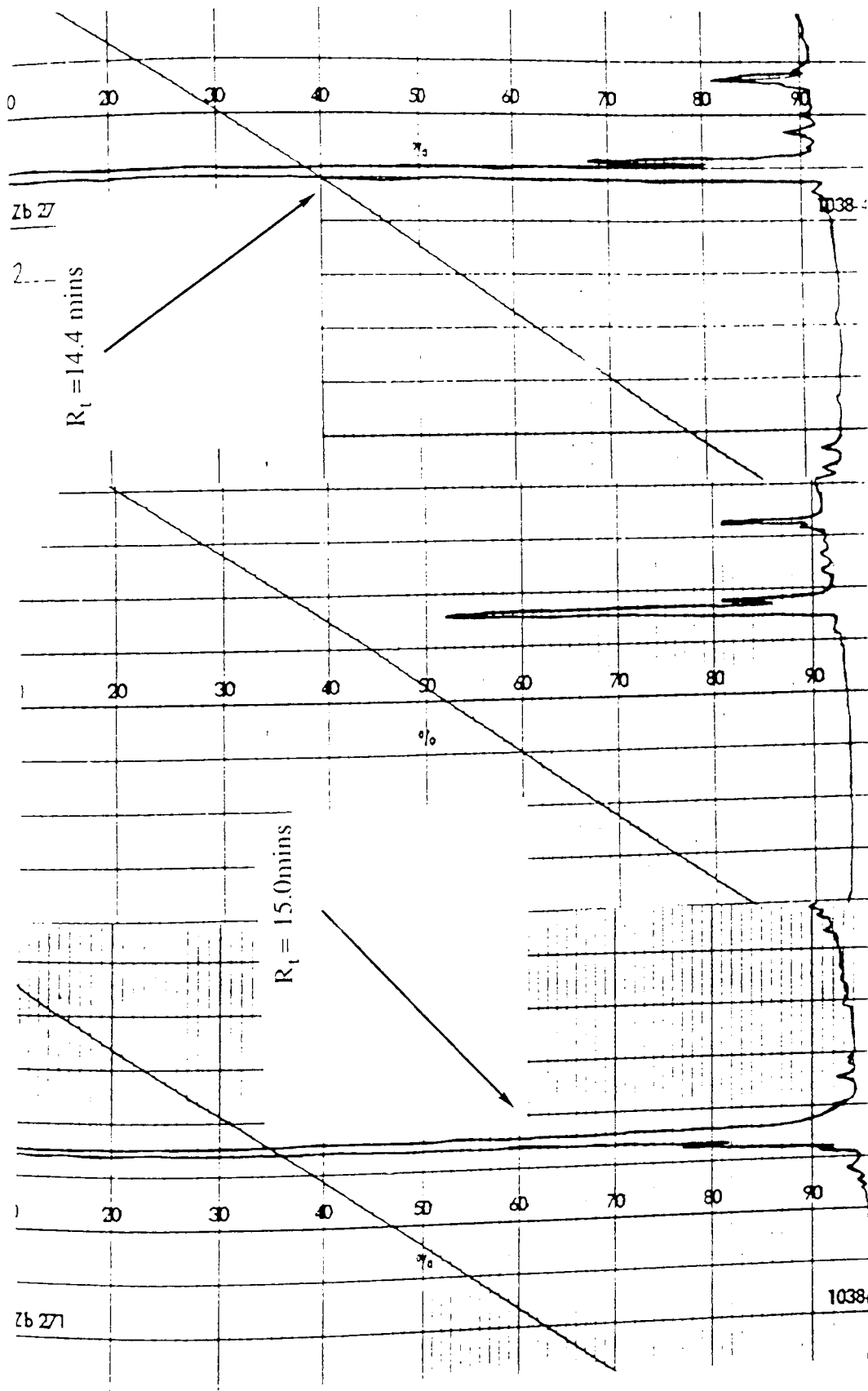


Figure 46 HPLC profiles of crude peptide 54 after 1, 3, and 4 hrs incubation with 95% TFA: anisole: ethylmethylsulphide (91: 4.5: 4.5)

RP C8 column using a mobile phase of CH<sub>3</sub>CN (0.01% TFA) : water (0.01% TFA) employing a linear gradient (90-10% over 30 mins). Chromatograms of the crude peptide showed three main peaks ( Fig.47 and Table 16).

Table 16 Details of peaks from the analytical HPLC purification of peptide 54

Peak number	R <sub>t</sub> (mins) <sup>a</sup>	Area (%)
1	13.33 (10.6)	17.24
2	15.45 (11.4)	21.02
3	18.77 (12.1)	13.64

a) Figures in brackets are corresponding retention times on preparative HPLC

Base line resolution was not possible on this scale. The peaks at R<sub>t</sub> =11.4mins and R<sub>t</sub> =12.1mins were identified as the peaks of interest, corresponding to peak 2 and peak 3 on an analytical scale. The compound at R<sub>t</sub> = 10.6 mins co-eluted with anisole and hence was not collected. Isolation was achieved using a linear gradient range from 90% to 10% water (0.01% TFA) over a time of 30mins. The desired peaks were collected and lyophilised to dryness, only peak 2 produced peptidic material in the form of a white fluffy solid. Purity of this material was confirmed using analytical HPLC which showed a single peak with a slight shoulder. The FAB mass spectrum of the sample showed a peak at 891

corresponding to the  $[M+H]$  of the free peptide, little fragmentation was observed in the spectrum.

#### 8.4.3 Cleavage and deprotection of Ser-Pro-Cys-Met-Arg-Leu-Ile-His-Leu (55)

Peptide **55** contained 4 residues which required protection; cysteine, histidine and arginine and serine. The cysteine and histidine residues were protected as the trityl derivatives as in the other peptides. The serine residue was protected as the tertiary butyl (tBu) ether, an acid labile protecting group used widely to protect alcohols.<sup>103</sup> The arginine residue was protected using the 2,2,5,7,8-pentamethylchroman(Pmc) protecting group (Fig.48).<sup>104</sup> This group was introduced in 1987 specifically to meet the demands of solid phase synthesis and is readily cleaved by TFA.

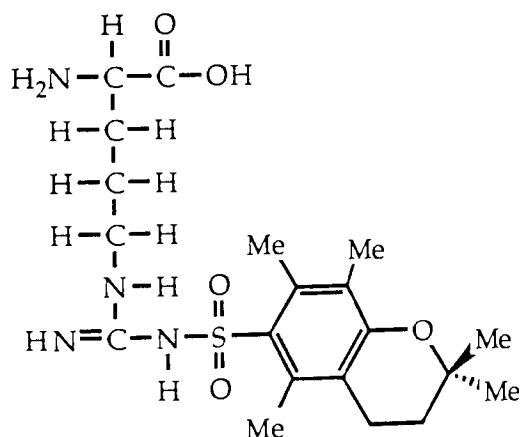


Figure 48 Structure of 2, 2, 5, 7, 8-pentamethylchromanarginine

Deprotection of peptide **55** was carried out using TFA with the scavengers, thioanisole and anisole (91:4.5 v/v). HPLC analysis after 1 hour showed 1 large peak at  $R_t = 15$ mins, and a number of smaller peaks at  $R_t = 15.5$ mins, 16.0mins, and 16.5mins. After 2 hrs the

peak at  $R_t = 15$  mins had diminished and the peak at  $R_t = 16$  mins had increased. After 5 hrs the peak at  $R_t = 16$  mins was the only significant peak. Therefore 5 hrs was chosen as the optimum time for deprotection. Following deprotection, analysis of the crude product by analytical HPLC revealed a number of peaks (Fig.49 and Table 17).

Table 17 Major peaks present in the HPLC chromatogram of crude peptide 55

Peak number	$R_t$ (mins) <sup>a</sup>	Relative area %
1	13.18	8.65
2	15.48 (9.65)	17.48
3	16.32 (10.67)	17.36
4	17.15 (12.63)	9.98
5	18.07	14.7

a) Figures in brackets are corresponding retention times on preparative HPLC

The peak at  $R_t = 13.18$  mins was found to co-chromatograph with the scavenger molecule, anisole. The peaks at  $R_t = 15.48$  mins, 16.32 mins, and 18.07 mins were thought to be potential product peaks which on a preparative scale corresponded to peaks of  $R_t = 9.65$  mins, 10.67 mins (**58**) and 12.63 mins (**59**), respectively. These peaks were collected, lyophilised and each yielded white fluffy material, highly characteristic of peptides. The three samples were analysed by FAB mass spectroscopy. The peak at

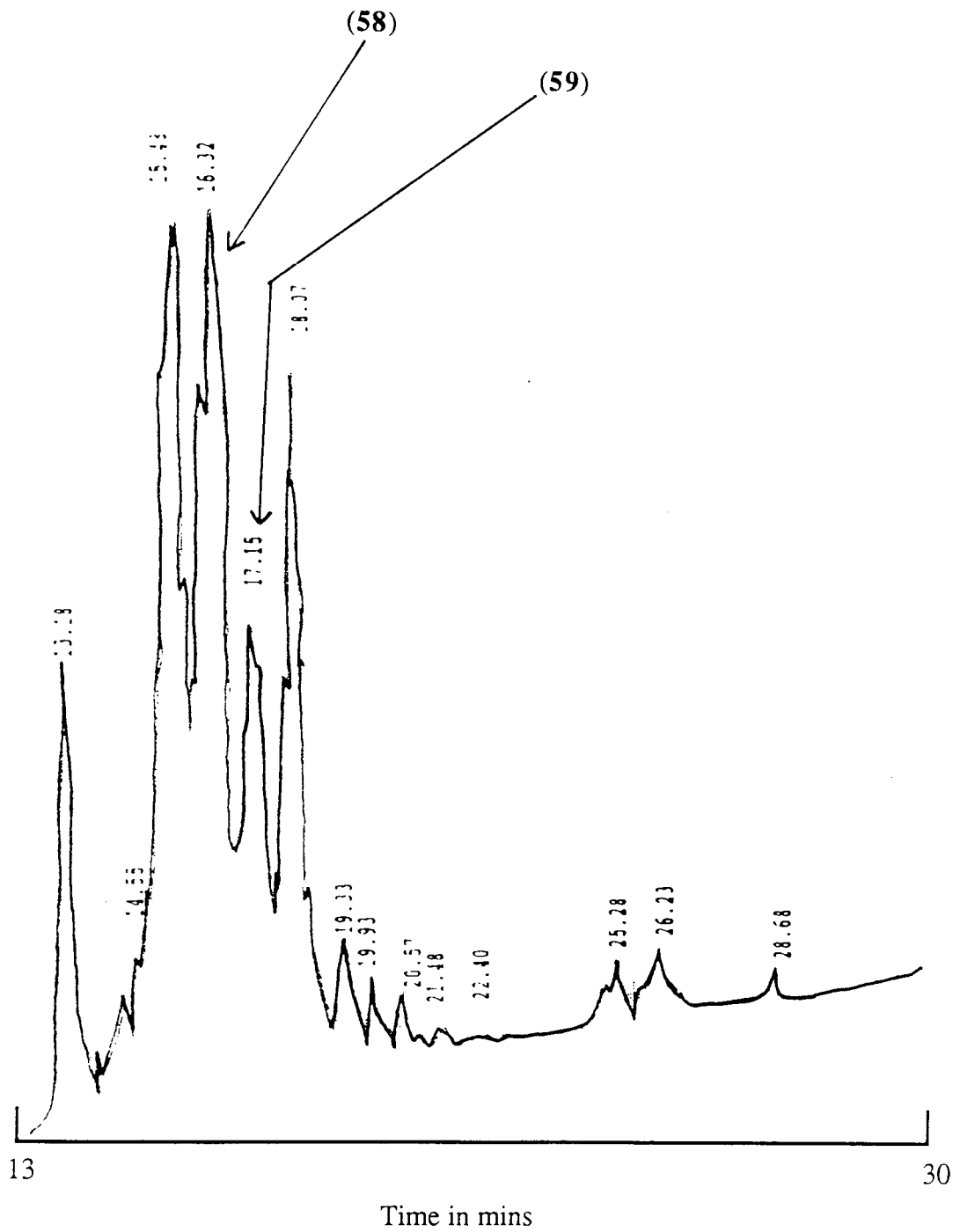


Figure 49 HPLC profile of crude peptide **55** after complete cleavage from the resin and deprotection by 95% TFA: anisole: thioanisole (91:4.5:4.5)

$R_t = 9.65$  mins corresponded to the desired product showing an  $[M+H]$  peak at 1069. The other samples both exhibited peaks at 1178, the identity of which was initially unclear. Examination of the mass spectra for **58** and **59** showed no peak at 1069, the expected  $[M+H]$  peak. A number of possible fragmentation pathways were considered. The higher molecular weight compound present may have formed as a result of reaction between a thiophenol anion (generated from acidic cleavage of thioanisole) and the thiol group of the cysteine residue in the peptide. A possible fragmentation pattern is suggested in Fig.50.

The high resolution FAB mass spectrum recorded did not monitor peaks below 200, therefore the existence of a fragment corresponding to thiophenol was not confirmed. A peak was seen at 212 which corresponds to a S-thiophenol cysteine fragment.

#### 8.4.4 Cleavage and deprotection of Pro-Cys-His-Met-Asp (**37**)

A deprotection study on peptide **37** was carried out using TFA as the cleaving agent, in the presence of anisole and ethylmethyl sulphide as scavengers. After a 1 hour incubation, the presence of 2 main compounds were indicated by HPLC [Conditions: linear gradient from 90% to 10% water (0.01% TFA)/CH<sub>3</sub>CN (0.01% TFA)], a minor peak at  $R_t = 16.5$  mins and a major peak at  $R_t = 19.4$  mins. Following a further hour incubation, the two peaks were approximately equivalent in height, and a new peak appeared at  $R_t = 15.2$  mins. After 4 hrs, the peaks at  $R_t = 15.2$  mins (peak 1) and at  $R_t = 16.4$  mins (peak 2) were the most significant and were targeted for collection on a preparative scale. Lyophilisation of both isolates produced white material. On reanalysis of the isolated peaks using analytical HPLC, peak 1 appeared to have broken down. Peak 2 still exhibited a single peak (**38**). Analysis of peak 2 by NMR suggested that the trityl protection was still intact (Fig.51).

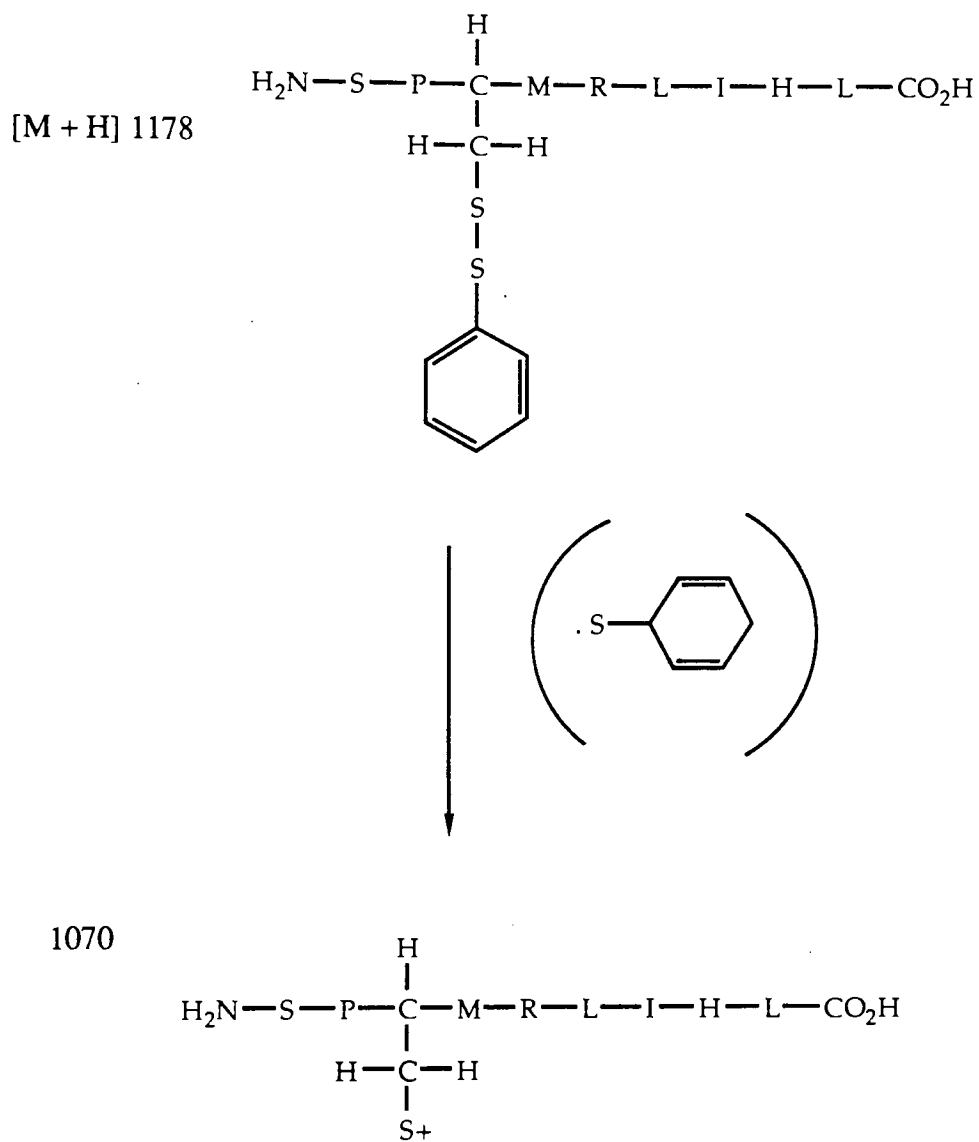


Figure 50 Proposed fragmentation pathway of the peptidic compound (58) isolated by HPLC possessing a  $R_t=10.67$ mins from the cleavage reaction of peptide (55)



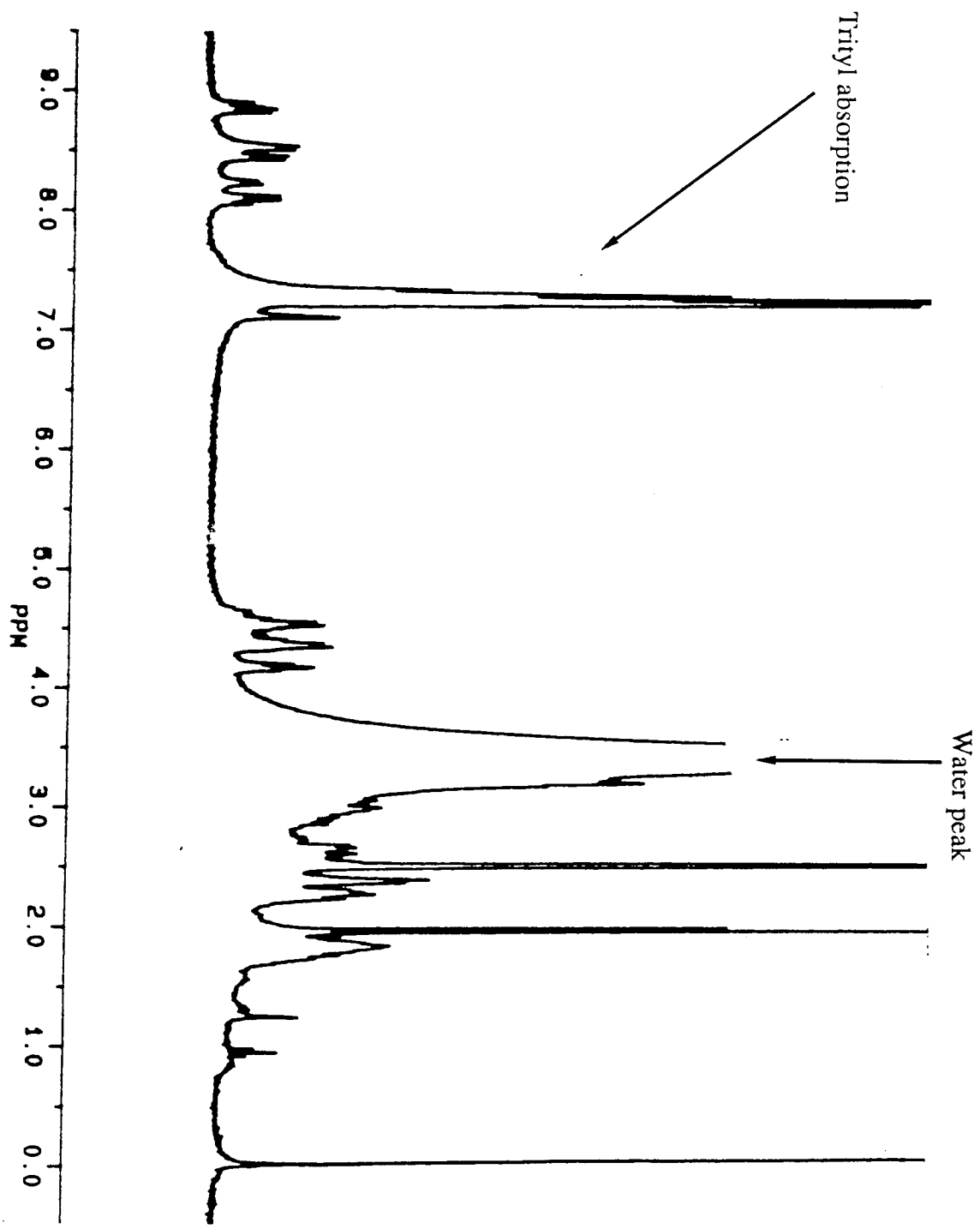


Figure 51  $^1\text{H}$  NMR spectrum of prolyl-S-(Trt)cysteinylnhistidylmethionylaspartate (38)

The absorption at  $\delta 7.2$ ppm was due to aromatic protons of the trityl group. Compound **38** was identified as prolylcysteinyl(S-Trt)histidinylmethionylaspartate.

**38** was reincubated with 95% TFA in the presence of anisole and ethyl methyl sulphide overnight. The peptide was again isolated from the deprotection medium by collection following HPLC and lyophilisation. Analysis by HPLC under the same conditions as employed in the purification stage resulted in a material showing at least 4 major peaks and a number of minor peaks. It was concluded that the close proximity of the charged residues in this peptide was acting to make the peptide unstable when all protecting groups were removed.

The peptides **53-55** and **37** were tested to see whether they were capable of directing the methylation of the sulphonium compound S-methylmethionine sulphonium iodide to cytosine position C5 (see Chapter 9). For the peptides which appeared unstable with a free thiol group, a crude mixture of protected and unprotected compound was used.

#### 8.4.5 Purity of peptides for biological evaluation

The purity of the peptides synthesised by solid phase methods is summarised in Table 18.

Table 18 Purity of peptides for biological evaluation

Peptide Number	% Purity <sup>a</sup>
53	55
57	100
54	100
55	100
37	60
38	100

a) Purity as assessed by HPLC

## CHAPTER 9

### Development of an assay system for C5 methylating agents

#### 9.1 Requirement for a new assay system

The compounds synthesised in Chapters 7 and 8 needed to be assessed for their ability to act as potential C5 methylators or as a part of a C5 methylating system.

Before the compounds could be screened it was important to devise an effective system by which to assay the compounds. Restriction enzyme analysis has been used by a large number of investigators to detect levels of 5-MC in DNA,<sup>14,23</sup> and was initially used to test the ability of simple sulphonium compounds to alkylate cytosine at position C5 in calf thymus DNA (Chapter 5). The results from the restriction enzyme assay suggested the possibility that S-methylmethionine sulphonium iodide could in conjunction with the enzyme DNA MTASE cause methylation at the C5 position of cytosine. However this evidence could not be accepted as conclusive due to a number of factors. Firstly the reproducibility of the results presented above was not very good, this could be explained by the fact that high concentrations of sulphonium salt would affect the pH of the methylating buffer. It was also feasible that the observed changes may be a result of alkylation at sites other than cytosine C5 within the recognition sequence of the restriction enzyme. This proposition appears reasonable as Green *et al.* have shown that restriction enzyme cutting is affected by other alkylations. When using the endonuclease PstI they found cleavage was inhibited by oligonucleotides containing O<sup>6</sup>-methylguanine.<sup>85</sup>

In order to be an effective substrate for DNA MTASE it would be hoped that S-methylmethionine sulphonium iodide would be active at much lower concentrations than those seen in the restriction enzyme assay.

The problems associated with the restriction enzyme assay therefore identified the need for a different assay approach to determine conclusively that :

- i) simple sulphonium salts eg. S-methylmethionine sulphonium iodide can act to alkylate DNA at cytosine position C5.
- ii) whether sulphonium salts other than SAM may act as methyl donors for DNA MTASE.

Using [<sup>3</sup>H]-SAM and S-[<sup>3</sup>H-methyl]methionine sulphonium iodide both in the presence and absence of DNA MTASE it was possible to determine the ability of both compounds to act as a substrate for DNA MTASE and to act independently as cytosine C5 methylators. The assay involved incubation of radiolabelled material with DNA, if the compound showed ability to alkylate cytosine, it would transfer a labelled methyl group to the DNA. Isolation of the DNA by filtration, followed by hydrolysis would result in the degradation of the DNA to its constituent bases. Separation of these bases by HPLC and radioactive counting would determine the specific sites and degree of methyl transfer.

Using this method it would be possible to prove conclusively whether any methyl transfer to cytosine C5 had occurred. This method could also be adapted to investigate whether any of the peptides synthesised were capable of replacing DNA MTASE as facilitators of cytosine methylation. Using [<sup>3</sup>H]-SAM as the methyl donor, experiments could be carried

out as above, replacing DNA MTASE with potential activating peptides. The results of these experiments are presented in the following section.

## **9.2 Experiments to determine the methylation sites of both S-methylmethionine sulphonium iodide and S-adenosylmethionine**

### **9.2.1 Experiments to determine whether S-methylmethionine sulphonium iodide is capable of acting as the methyl donor for DNA MTASE**

In order to determine whether S-methylmethionine sulphonium iodide was actually working as a weak substrate for DNA MTASE, a competition assay was carried out between SAM and S-methylmethionine sulphonium iodide. The assay was used to determine whether inhibition of methylation of *M.Lysodeikticus* DNA by [<sup>3</sup>H]-SAM and DNA MTASE occurs in the presence of unlabelled S-methylmethionine sulphonium iodide. If S-methylmethionine sulphonium iodide was acting as a substrate for the enzyme, inhibition of methylation would be seen as a drop in the number of labelled methyl groups transferred to DNA.

No inhibition was observed at varying concentrations of S-methylmethionine sulphonium iodide, in fact a slight increase in methylation was observed at a concentration at 10mmol. The reason for this is unclear, but a possible explanation may lie in the fact that if the DNA becomes more fragmented then there may well be an increase in DNA substrate concentration.

A more conclusive experiment was performed using labelled SAM and S-methylmethionine sulphonium iodide in the assay system described in section 9.1.

[<sup>3</sup>H]-SAM and S-[<sup>3</sup>H-methyl]methionine sulphonium iodide were incubated with *M.Lysodeikticus* DNA for 4 hrs at varying radioactive concentrations in the presence and absence of DNA MTASE. As can be seen from the results (Fig.52) the transfer of a radiolabelled methyl group from S-[<sup>3</sup>H-methyl]methionine sulphonium iodide was not significantly different in the presence of DNA MTASE when compared to the results seen in the absence of DNA MTASE. HPLC analysis showed N<sup>7</sup>-methylguanine present as the only alkylated base. Repeat of these experiments with greater concentrations of radiolabelled compound may produce some more interesting results.

#### 9.2.2 Experiments to determine whether S-methylmethionine sulphonium iodide is capable of independently alkylating cytosine at position C5

One of the principal targets of this thesis was to determine whether a simple sulphonium compound may be capable of alkylating the C5 position of cytosine in the absence of DNA MTASE. An experiment to determine whether S-methylmethionine sulphonium iodide may perform this function would involve using a labelled sulphonium compound and HPLC analysis as introduced in section 9.1.

Incubation of S-[<sup>3</sup>H-methyl]methionine sulphonium iodide with poly d[G-C].poly d[C-G] was carried out to determine possible sites of alkylation both in the presence and absence of DNA MTASE. The synthetic oligonucleotide poly d[G-C].poly d[C-G] was chosen as the substrate for the reaction in order to reduce the number of possible sites susceptible to alkylation when compared to DNA. Following incubation with S-[<sup>3</sup>H-methyl]methionine sulphonium iodide, the polydeoxynucleotide was hydrolysed using formic acid. The resultant hydrosylates were then analysed by HPLC using radiochemical detection methods employing the solvent system described in Chapter 7. In

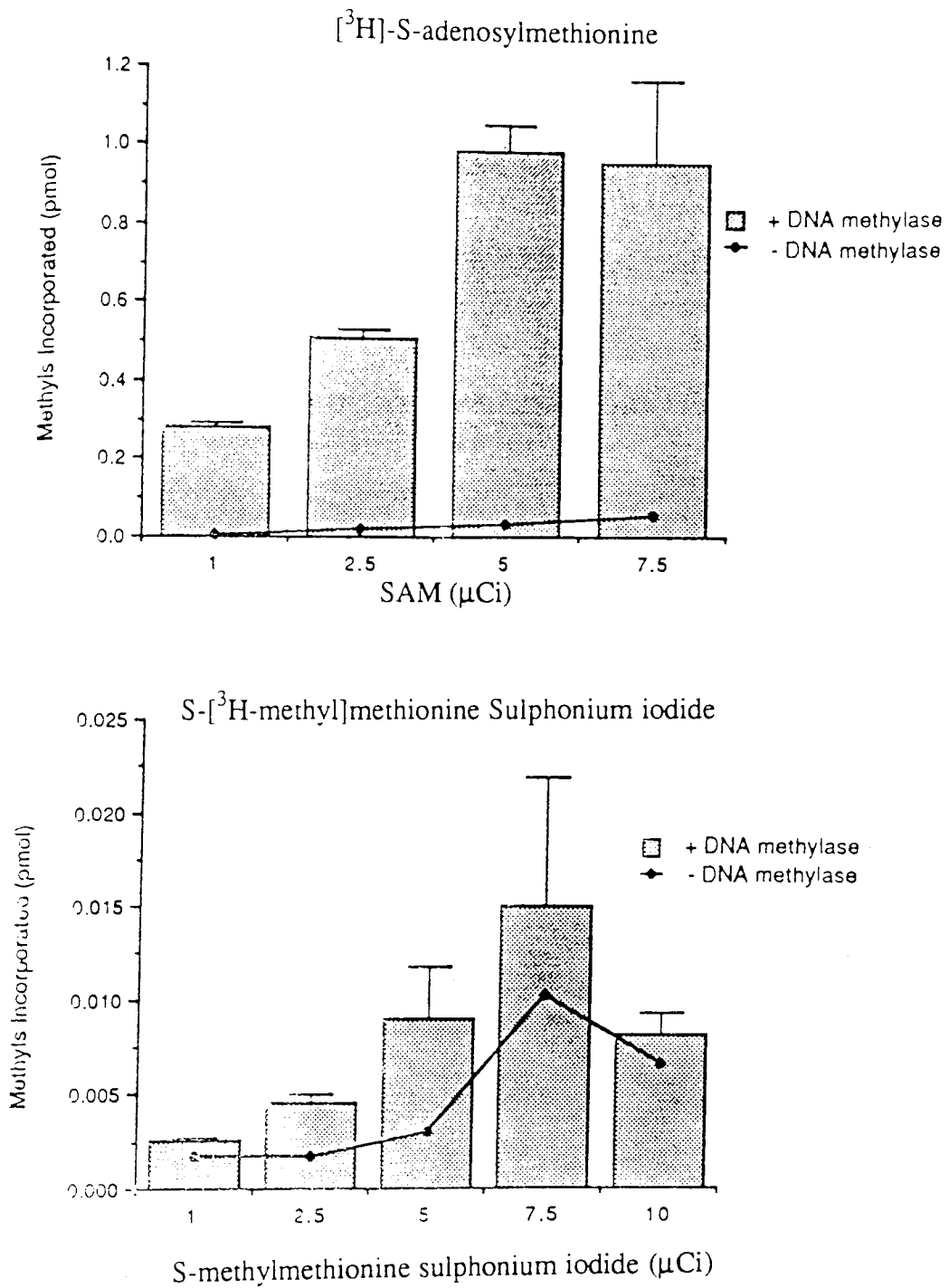


Figure 52 The effect of S-methylmethionine sulphonium iodide and SAM on the methylation of *M. lysodeikticus* DNA in the presence and absence of DNA MTASE



an experiment using 50 $\mu$ Ci (equivalent to 649pmol) of S-[<sup>3</sup>H-methyl]methionine sulphonium iodide, the hydrolysed DNA contained no labelled material corresponding to 5-MC. A very small peak was detected that co-eluted with the N<sup>7</sup>-methylguanine equivalent to 0.0018pmol. In a parallel experiment run with 50 $\mu$ Ci of [<sup>3</sup>H]-labelled SAM, N<sup>7</sup>-methylguanine was again the only detectable peak. It should be noted that under the hydrolysis conditions described O<sup>6</sup>-methylguanine is broken down and therefore would be undetectable.

Similar experiments were carried out by Barrows and Magee investigating the ability of SAM to alkylate DNA.<sup>105</sup> In their experiments they used 200 $\mu$ Ci of labelled SAM in an assay and revealed the presence of the methylated bases N<sup>7</sup>-methylguanine, N<sup>3</sup>-methyladenine and O<sup>6</sup>-methylguanine. These findings were mirrored in a similar investigation by Rydberg and Lindahl,<sup>106</sup> using 100 $\mu$ Ci of [<sup>3</sup>H]-SAM. Both groups of workers used a similar method for hydrolysis of DNA which only acted to free purine residues. Their method using 0.1M HCl at 70<sup>0</sup>C for 40 mins to hydrolyse DNA left all pyrimidine residues still intact. Therefore neither group were able to assess whether SAM was capable of C5 methylation of cytosine independently of the enzyme.

From the results presented it can be concluded that both SAM and S-methylmethionine sulphonium iodide are unable to alkylate the C5 position of cytosine in the absence of DNA MTASE.

Comparing the conditions employed with cellular conditions, the SAM concentration in the reaction was equivalent to  $5.88 \times 10^{-7}$  mol which is about 70 times lower than the average intracellular concentration.<sup>107</sup> SAM would not be expected to alkylate cytosine at position C5 independently at physiological concentrations, therefore experiments using

much higher concentrations of SAM might provide more interesting results. However the high cost of radiolabelled materials prevented this work from being carried out.

These experiments did manage to prove definitely that neither SAM nor S-methylmethionine sulphonium iodide are potent C5 methylating agents. Therefore their use as C5 methylating agents for DNA may be limited due to their ability to alkylate DNA at other sites, in particular at the O<sup>6</sup>-position of guanine which is a mutagenic lesion. Any gene silence that may result from high concentrations of these compounds would be overshadowed by their potential as weak carcinogens.

### **9.3 Experiments to determine the ability of synthetic peptides to facilitate the alkylation of cytosine position 5**

A number of the peptides synthesised were assayed for their ability to catalyse the methylation of DNA substrates in the presence of [<sup>3</sup>H]-SAM i.e act in place of DNA MTASE. From the results (Table 19), none of the potential activating molecules showed any significant increase in the amount of radioactivity transferred to *M.lysodeikticus* DNA, above controls. In order to ensure that any potential effect was not being masked by a general non-specific methylation, the assay was repeated using a poly d[G-C].[C-G] sequence. Again these results, showed no significant rise in the amount of radioactivity incorporated into the oligonucleotide sequence. From the results it was concluded that the sequences synthesised were not of sufficient length for any enzyme-like structure to be adopted and hence were unable to adopt a MTASE-like active site conformation. The fact that the methyl donor was not binding to any specific site as it would in DNA MTASE would mean that a much higher concentration of methyl donor would need to be present in the medium. The likelihood that methyl donor, oligonucleotide and activating sequence

Table 19 Assay of synthetic peptides to determine their ability to aid alkylation of DNA in the presence of [<sup>3</sup>H]-S-adenosylmethionine

Peptide	CPMA after incubation <sup>a</sup>	
	A <sup>b</sup>	B <sup>c</sup>
Cys-Ala-Pro-Cys-Leu-Pro-Ala-Gly (53)	892±67	639±124
bis (N-Boc-Pro)cystine dimethyl ester (36)	638±41	683±92
Gly-Phe-Pro-Cys-Pro-His-Phe-Ser (54)	NR*	1163±197
Ser-Pro-Cys-Met-Arg-Leu-Ile-His-Leu (55)	NR*	815±87
Pro-Cys-His-Met-Asp (37)	NR*	1611±85
bis (N-CBZ-Met)cystine dimethyl ester (22)	1152±151	521±64
N-Boc-prolylcysteine methyl ester (39)	638±41	655±76
Glutathione	609±92	703±91
Untreated	935±67	NA**
Control (-DNA)	928±48	766±149

a incubation with *M. Lysodeikticus* DNA for 0.5hrs

\* NR = no reaction

b Concentration 1.0µmol

\*\* NA = not applicable

c Concentration 0.1mmol

will be in the correct position for C5 methylation to occur would be of very low probability in the assay system adopted unless the reaction medium possessed nearly saturated levels of labelled methyl donor. This point highlights the difficulty of the task undertaken.

#### 9.4 Cytotoxicity assays

Assessment of the compound's cytotoxicity was necessary in order to determine the extent to which compounds are toxic to a particular cell type. Any successful chemotherapeutic agent would be required to show a reduced toxicity towards normal cells when compared to malignant cells under similar conditions. Whether any potential C5 methylating molecule would be expected to show any cytotoxic activity is unclear although SAM at high concentrations has been shown to produce alkylations at N<sup>7</sup>-guanine, N<sup>3</sup>-adenine and O<sup>6</sup>-guanine (a potential cytotoxic lesion).<sup>59</sup>

In order to highlight particularly interesting target compounds, it was decided to test a number of the compounds synthesised for their ability to act as cytotoxic agents. The more interesting compounds discovered could then be investigated further to discover whether their mode of action involved any methylation at cytosine position C5 using the assay system introduced earlier. Initially the three commercially available sulphonium compounds S-methylmethionine sulphonium iodide (4), dimethylpropargylsulphonium bromide (2) and carboxydimethyl sulphonium bromide (3) were assessed for their toxicity against the GM892 cell line. The cells were exposed to a single dose of sulphonium compound at varying concentrations and were then left for four days before cell numbers were counted. Very high concentrations of drug (over 200mM) were required before any significant effect on cell number resulted. Therefore it appeared that simple sulphonium compounds had low cytotoxicity.

A number of the compounds synthesised were then tested against the A549 human derived lung adenocarcinoma cell line (Table 20). The methionine sulphonium salts, S-methyl (**4**) and S-ethylmethionine sulphonium iodide (**6**) both showed little activity against this cell line, no inhibition of growth was seen below 200 $\mu$ mol. This confirmed the results seen with the GM892 cell line, that simple methyl sulphonium salts show little cytotoxic activity. The simple sulphonium compound tributylsulphonium perchlorate (**7**) showed some very slight cytotoxic activity, with inhibition of growth being seen at a concentration of 200 $\mu$ mol, reaffirming that simple sulphonium salts appear to be poor cytotoxic agents, but suggesting the possibility that the larger the alkyl moiety attached to the sulphonium centre, the greater the potency of the sulphonium compound. It is difficult to interpret from these results whether the difference in cytotoxicity is significant. The effect of groups attached to **4** appeared to show a more significant effect. The compound N-CBZ-S-methylmethionine nitrophenyl ester sulphonium iodide (**8**) showed a greater potency than the simple amino acid sulphonium, with an  $IC_{50} = 50\mu$ mol. The possibility that the increased potency may be due to an electronic effect from the electron rich benzene ring is unlikely due to the fact that the sulphonium centre is 3 bonds away from the aromatic ring. The most likely explanation is that the removal of the zwitterionic character of the amino acid portion of the molecule results in the sulphonium centre losing a slight stabilising influence and more readily donating a methyl group. It must be stressed that this is only a hypothesis, as the mechanism of action of the cytotoxic activity of these sulphonium compounds is unknown. The compound N-CBZ-(S-Bzl)cysteinyl-S-methylmethionine methyl ester sulphonium iodide (**28**) showed a similar spectrum of activity to the compound **8**. The compound N-CBZ-cysteinyl-S-methylmethionine methyl ester sulphonium iodide (**32**), a specific target compound showed some potency against the A549 cell line. Comparing this result with the result seen for **28**, it appears that the free thiol group of cysteine appears to alter the cytotoxic effect of the dipeptide sulphonium salt.

Table 20 Assay for inhibition of A549 cell growth by sulphonium compounds

Compound	Activity
S-methylmethionine sulphonium iodide (4)	No inhibition below 200 $\mu$ mol
S-ethylmethionine sulphonium iodide (6)	No inhibition below 200 $\mu$ mol
[N-Fmoc-S-(Trt)cysteinamido]-S,S-dimethyl phenyl sulphonium perchlorate (31)	100% inhibition at 50 $\mu$ mol
N-CBZ-S-methylmethionine-4- nitrophenyl ester sulphonium p-toluene sulphonate (8)	100% inhibition at 200 $\mu$ mol
N-CBZ-S-(Bzl)cysteinyl-S-methylmethionine methyl ester Sulphonium iodide (28)	100% inhibition at 200 $\mu$ mol
N-CBZ,S-methylmethionylcysteine methyl ester sulphonium iodide (32)	100% inhibition at 75 $\mu$ mol

An interesting result was seen with the compound [N-Fmoc-S-(Trt)cysteinamido] S,S-dimethylphenyl sulphonium perchlorate (31). This compound was the most potent compound tested against the GM892 cell line with an  $IC_{50} = 35\mu\text{mol}$ . This result suggests that aromatic sulphonium compounds may show a greater cytotoxic effect than aliphatic sulphonium compounds. From these results a number of interesting directions for future work have been identified.

Generalised genome wide methylation would be expected to produce highly cytotoxic compounds. Therefore the initial results seen with these compounds are encouraging, suggesting that if the compounds resultant cytotoxicities are as a result of methylating DNA, this methylation is not widespread and indiscriminate. They may certainly have some selective action in the sites they methylate but from experimentation it is unlikely any methylation occurs at cytosine position C5.

## CHAPTER 10

### Conclusions and further work

#### 10.1 Development of small molecules as C5 methylating agents

The initial hypothesis proposed in this thesis was to test the possibility that a small molecule may be capable of selectively alkylating the C5 position of cytosine. The work presented has shown that simple sulphonium compounds show no specific ability to alkylate cytosine. In particular it highlights the requirement for an activation mechanism to facilitate alkylation at the C5 position of cytosine. In the natural system this role is performed by the enzyme DNA MTASE. The investigations using restriction enzyme analysis and radiolabelled sulphonium compounds have shown that SAM is the only sulphonium salt capable of being used by the enzyme as a methyl donor. The experiments carried out also seem to suggest that an activating mechanism for methylation will not be found in small peptidic molecules.

Considering the information above, there are two possible ways to proceed from the position reached. The first way would be to attempt to synthesise larger peptides containing the conserved Pro-Cys motif and testing their ability to work as activating molecules. This avenue was pursued in some form by synthesising longer peptides taken from areas in the eukaryotic enzyme thought to possibly be present in the active site. These initial experiments failed to produce any active molecules. Further pursuit of this route of investigation would be very time consuming.

A second possible method would involve taking the much smaller prokaryotic DNA MTASE (molecular weight 20Kda), and breaking it down using restriction enzyme digestion and testing each fragment produced for activity. Using this method repeatedly



it would be possible to discover the smallest fragment capable of retaining the MTASE's activity. The fragment would then prove very useful for possible three-dimensional structural studies and also, dependent on the eventual size of the fragment, a possible synthetic target. Support for this strategy can be found in work by Adams,<sup>108</sup> who, using proteolytic digestion, degraded mouse ascites DNA MTASE of molecular weight 185Kda, to give polypeptides of 170Kda, 100Kda, and 50Kda all possessing MTASE activity.

## **10.2 Identification of the active site sequence of eukaryotic DNA MTASE**

The peptides synthesised, although possessing no activity as facilitators of sulphonium methylation, may still be able to play a role in helping determine which of the three cysteine residues present in the eukaryotic DNA MTASE is vital for enzyme activity. A prospective route of investigation would depend on whether antibodies could be raised to the peptide sequences synthesised 34-37. If so each antibody raised could be incubated in turn with the active enzyme and then the enzyme could be assayed for activity. If activity is destroyed or hindered following one of the incubations then this would signify the particular residues that raised the antibody were involved in or highly adjacent to the active site.

## **10.3 Structural homology of peptides**

The molecular modelling studies highlighted some useful information even though no actual information on DNA MTASE enzymes was readily available. The homology of the phi and psi angles about the Pro-Cys dipeptide present in a number of proteins extracted from the BIPED database (see Chapter 7) suggests the possibility that a peptide similar in sequence to the active site region of DNA MTASE may well adopt a similar structure. It

would certainly be interesting to solve the crystal structure from one of the larger peptides **53-55**. From this structure the phi and psi angles associated with the proline and cysteine residues could then be compared to those seen in the proteins extracted from the BIPED database to see whether any structural homology existed. This work would help to give an idea of the length of peptide required to show some similar three-dimensional structure characteristics to those seen in proteins. Work at the moment is concentrating on attempting to grow crystals of **55**, in an effort to achieve the above goals. Simply using molecular modelling techniques at the moment would be impractical due to the enormous amounts of computer time required to minimise molecules of such a size.

#### **10.4 Methylating ability of sulphonium compounds**

The cytotoxicity assays of the sulphonium compounds synthesised produced some interesting results. The low cytotoxicity of S-methylmethionine sulphonium iodide was to be expected as it is found as a natural methyl donor in the plant kingdom. The fact that the aromatic sulphonium compound showed a greater potency than any of the aliphatic sulphonium compounds was very interesting. Further work with a number of other aromatic sulphonium compounds would help to clarify if this effect is a general effect or just an individual case. The results seen with the two compounds based on the dipeptide methionylcysteine, seem to suggest that if the cysteine residue is unprotected, there is a slight enhancement of the cytotoxicity of the compound. The next logical step would be to determine whether this compound is capable of methylating DNA and if so at what specific sites the methylation occurs. This could be carried out by radiolabelling the methyl groups attached to the sulphonium centre as in the experiments to determine the sites of methylation of labelled S-methylmethionine sulphonium iodide.

The work in this thesis contributes to the basic background knowledge required to produce a cytosine C5 methylating molecule and suggests a slightly more complex molecule may be required to achieve this goal. It also proposes some of the physical characteristics that would be required by such a molecule. Some of the further work suggested will provide additional information which should make the realisation of the goals set out more achievable. The project set out was ambitious and the information discovered confirms this.

### **Section 3: Materials and Methods**

## CHAPTER 11

### Experimental

#### 11.1 Chemistry experimental

All melting points were measured on an electrothermal digital melting point apparatus and were uncorrected. 60MHz NMR spectra were recorded on a Varian EM360A NMR spectrometer, 250MHz and 300MHz were recorded on Bruker AM250 and AM300 NMR spectrometers respectively. Infrared (IR) spectra were recorded on a Perkin-Elmer 1310 IR spectrophotometer on sodium chloride disks unless stated otherwise. Mass spectra were recorded on a V.G. micromass 12 instrument at 70eV and a source temperature of 300°C. Fast atom bombardment (FAB) mass spectra were recorded by the SERC mass spectrometry service centre using nitrobenzaldehyde (NOBA) as the matrix. Elemental analyses were within  $\pm 0.3\%$  of the theoretical values. The TLC systems employed Kieselgel 60F254 (0.25 mm) as the absorbant. Melting points (m.p.) are uncorrected.

##### 11.1.1 Synthesis of sulphonium salts

#### Preparation of S-methylmethionine sulphonium iodide (4)

A mixture of methionine (0.60g, 4mmol), 89% formic acid (6.5ml), acetic acid (2ml) and methyl iodide (1ml) was left to stand in the dark at room temperature for three days. The mixture was then distilled *in vacuo* to leave a syrup. Methanol (40ml) was added and the mixture left to stand overnight at 0°C. A pale yellow solid was filtered off and washed well with methanol.

Yield=68%; m.p. 148-150°C (lit.150°C)<sup>81</sup>; M<sup>+</sup> 163 gives C<sub>6</sub>H<sub>13</sub>NO<sub>2</sub>S;  
λ<sub>max</sub> (nujol) 3300 (b), 2850,1710,1550,1375, 1210 cm<sup>-1</sup>; δH (D<sub>2</sub>O) 4.3 (1H,t) 3.0  
(6H,s) 2.7 (2H,d) 2.3 (2H,t) ppm; υ max (water) 225nm.

#### **Preparation of S-methylmethionine sulphonium chloride (60)**

L-Methionine (1.49g,1mmol) was dissolved in methanol saturated with HCl (50ml) and the mixture was stirred at 60°C for 24 hrs. The methanol was then removed in *vacuo* to leave a syrup. On addition of ether the syrup solidified, recrystallisation from hot ethanol produced a white solid.

Yield=70%; m.p. 147-148°C (lit 150°C)<sup>81</sup>; analysis as for iodide.

#### **Preparation of S-methylmethionine sulphonium dimethanesulphonate (61).**

Methionine (0.5g, 0.33mmol) was added to a mixture of anisole (1ml), and methanesulphonic acid (3ml, 4.62mmol) and the reaction mixture was stirred at room temperature for 24 hrs, after this time excess ether was added. The ether was decanted from the syrup like mixture and on addition of ethanol (100ml), a solid precipitated out. Recrystallisation from hot ethanol resulted in a white crystalline solid.

Yield= 93% ; m.p. 182-185°C (ref 187-190°C)<sup>82</sup>; M<sup>+</sup> 163

λ<sub>max</sub> (KBr disc) 3140, 2100, 1710, 1620, 1540, 760 cm<sup>-1</sup>; δH(D<sub>2</sub>O) 4.2 (1H,t) 3.5  
(2H,dd) 2.95 (6H,s) 2.8 (6H,s) 2.4 (2H,m) ppm;

### Preparation of S-ethylmethioninesulphonium iodide (6)

L-Methionine (7.4g, 0.05mol) was dissolved in water (40ml) and to this was added ethyl iodide (8.6g) the reaction was heated in a water bath at 45°C for 6 hrs. Solvent was removed by rotary evaporation and the resulting oil was dissolved in methanol. On standing overnight at 0°C a solid was isolated which was recrystallised from hot ethanol to give a pale yellow solid.

Yield=51%; m.p. 160-163°C (lit. 161-162°C)<sup>82</sup>; M<sup>+</sup> 178 gives C<sub>7</sub>H<sub>16</sub>NO<sub>2</sub>S ;  
λ<sub>max</sub> (nujol); 3500 (b), 2800, 1550, 1375, 1040 cm<sup>-1</sup>; δH (D<sub>2</sub>O) 3.9 (1H,t) 3.4 (2H,q)  
3.3 (2H,d) 2.9 (3H,s) 2.5 (2H,d) 1.5 (3H,t) ppm.

### Preparation of S-tributylsulphonium perchlorate (7)

A solution of n-butyl sulphide (0.5ml, 5mmol) in butan-1-ol (15ml) with 70% perchloric acid (0.72g) was refluxed overnight in a dean-stark apparatus. After allowing the solution to cool, a large excess of ether was added and a white solid precipitated. On standing overnight fine white crystals were filtered off and after washing well with ether were dried *in vacuo*.

Yield= 24%; m.p. 100-102°C (lit. 104°C)<sup>83</sup>; M<sup>+</sup> 235 gives C<sub>12</sub>H<sub>27</sub>S  
λ<sub>max</sub> (nujol) 2950,1660, 1590 cm<sup>-1</sup>; δH( DMSO-d<sub>6</sub>) 1.3 (t,9H), 1.9 (q,6H) 1.2 (m,6H)  
2.4 (t,6H)

**Preparation of N-CBZ-S-methylmethionine 4- nitrophenyl ester  
sulphonium p-toluenesulphonate (8)**

To a solution of N-CBZ-methionine 4-nitrophenyl ester (2.0g, 5mmol) in ethyl acetate (10ml), methyl p-toluenesulphonate (50mmol) was added and the reaction mixture was kept at room temperature for four days. After this time a white precipitate was observed. The initial crop was filtered off and the filtrate was concentrated using a rotary evaporator. The residue on standing at 0°C overnight, yielded a white crystalline solid which was isolated by filtration.

Yield = 22%; m.p. 148-152°C; M<sup>+</sup> 591 gives C<sub>20</sub>H<sub>23</sub>N<sub>2</sub>O<sub>6</sub>S (C<sub>7</sub>H<sub>7</sub>O<sub>3</sub>S);

λ<sub>max</sub> (nujol) 3450,3250,3025,1760,1710,1610,1300,860 cm<sup>-1</sup>;

δH(DMSO-d<sub>6</sub>) 8.4,8.2 (4H,dd),7.4,7.1 (4H,dd), 7.3 (5H,s), 5.1 (2H,s), 4.6 (1H,m),

3.4 (2H,t), 2.9 (3H,s), 2.5 (6H,d), 2.2 (2H,s) ppm;

**Preparation of N-CBZ-S-methylmethionyl-S-(Trt)cysteine methyl ester  
sulphonium iodide (29)**

N-CBZ-methionyl-S-(Trt)cysteine methyl ester (0.57g, 0.9mmol) was dissolved in methanol (10ml). Methyl iodide (2ml) was added to the mixture and the reaction was kept in the dark overnight. Removal of the solvent by rotary evaporation *in vacuo* afforded a yellow orange oil. The oil was dissolved in the minimum amount of chloroform, on addition of ether (100ml) a yellow/brown solid precipitated out. The solid was collected by filtration and washed well with ether. Recrystallisation was afforded from chloroform/ether.



Yield=68% ; m.p.= 92-94<sup>0</sup>C; M<sup>+</sup> 657 gives C<sub>37</sub>H<sub>41</sub>N<sub>2</sub>O<sub>5</sub>S<sub>2</sub> ;  
λ<sub>max</sub> (KBr disc) 3315,3050,2900, 1740,1685, 1655, 1540, 1255cm<sup>-1</sup>; δH(CDC<sub>13</sub>) 7.9  
(1H,d), 7.3 (5H,s), 7.2 (18H,s), 6.2(1H,d), 5.0(2H,s), 4.6(1H,dt), 4.2(1H,q),  
3.85(2H,d), 3.6(3H,s), 3.05(3H,s), 2.95(3H,s), 2.8(2H,d), 2.6(2H,dt)ppm;  
(Found C, 67.84 ; H, 5.50 ; N, 4.12 ; Requires C, 67.56 ; H, 5.20 ; N, 4.26 )

### Alkylation of Bis(N-CBZ-methionyl)cystine dimethyl ester (22)

Bis(N-CBZ-methionyl)cystine dimethyl ester (0.80g, 0.1mmol) (22) was dissolved in dichloromethane (5ml). Silver perchlorate (0.023g, 0.11mmol) along with methyl iodide (0.016g,0.11mmol) were added to the mixture. The reaction was stirred in the dark for 2 hours at room temperature. After this time the crude reaction mixture was filtered through a 0.2μ PTFE filter. After removing the CH<sub>2</sub>Cl<sub>2</sub> by rotary evaporation, a clear viscous oil remained. Attempts to obtain crystals by trituration with diethyl ether failed. Analysis by TLC showed three spots (n-butanol: acetic acid: water 12:5:3)). Purification by column chromatography gave an oil (30).

Yield= 20%; R<sub>f</sub> (n-butanol:acetic acid:water, 12:5:3) 0.5.; [M+H] 941 (813) gives C<sub>35</sub>H<sub>49</sub>N<sub>4</sub>O<sub>10</sub>S<sub>4</sub>I; λ<sub>max</sub> 3350, 2900, 1740, 1690,1640,1520,1250 cm<sup>-1</sup>; δH(CD<sub>3</sub>OD) 7.3 (10H,d) 5.0 (4H,d) 4.3 ( 2H, t) 4.2 (2H,t) 3.6 (6H,s) 3.2 (4H,t) 3.1 (4H, t,t) 2.7 (6H,d) 2.0 (3H,s) ppm.

### Preparation of N-CBZ-S-(Bzl)cysteinyl-S-methylmethionine methyl ester sulphonium iodide (28)

A mixture of N-CBZ-S-(Bzl)cysteinylmethionine methyl ester (**18**) (1.9g, 4mmol), 89% formic acid (6.5ml), acetic acid (2ml) and methyl iodide (5ml) was left to stand at room temperature in the dark for three days. After this time, the mixture was evaporated *in vacuo*. A dark red oil resulted which was triturated with ether. Attempts at recrystallisation failed. Purification was achieved using flash chromatography, (mobile phase methanol/chloroform (50:50)).

Yield=35%; [M+H] 506 gives  $C_{25}H_{33}N_2O_5S_2(I)$ ;  $\lambda_{max}$  (nujol) 3250, 2900, 1735, 1685, 1650, 1530, 1375, 1240, 690  $cm^{-1}$ ;  $\delta H(CDCl_3)$  7.8 (6H,s) 7.3 (6H,s) 6.3 (1H,d) 5.1(2H,s) 4.7(1H,m) 4.5 (1H,t) 3.7 (3H,s) 3.1(6H,s) 3.0 (2H,d) 2.9(2H,d) 2.4 (2H,m) ppm.

### Preparation of 4-[N-Fmoc-S-(trityl)cysteinamido]-S,S-dimethylphenylsulphonium perchlorate (31)

4-[N-Fmoc-S-(Trt)cysteinyl]thioanisole (0.32g, 0.5mmol) was dissolved in dichloromethane (5ml). Silver perchlorate (0.11g, 0.55mmol) was added along with methyl iodide (0.078g, 0.55mmol). The reaction was left in the dark for 4 hrs. Analysis by TLC showed the presence of three spots. Purification by flash chromatography with a mobile phase of acetic acid/water/propan-2-ol yielded a clear pale lilac oil.

yield= 10%; [M+Na] 745 gives  $C_{45}H_{41}N_2O_3S_2$ ;  $\lambda_{max}$  3300, 3050, 2900, 1660, 1590, 770, 690  $cm^{-1}$ ;  $\delta H(CDCl_3)$  7.5 (d,2H) 7.4 (d,2H) 7.35 (m,8H) 7.25 (s,18H) 6.25 (d,1H) 5.3 (s,2H) 4.4 (t,1H) 3.1 (s,6H) 2.7 (d,2H) ppm.

### 11.1.2. Synthesis of amino acid derivatives

#### Preparation of S-(Trt)cysteine (15)

L-cysteine hydrochloride (15.7g, 0.1mol) was dissolved in DMF (60ml) along with trityl chloride (42g, 0.15mol). The reaction was stirred at room temperature for two days. After this time 10% sodium acetate solution (500ml) was added to the mixture and a white solid precipitated. The solid was filtered off and dissolved in acetone (400ml). The solution was then refluxed for 15 mins. After standing and allowing to cool an off-white solid precipitated out, this was filtered off and washed well with acetone.

Yield=76.2%; m.p. 195-197°C (lit. 200-205°C)<sup>102</sup>;

M<sup>+</sup> 363 gives C<sub>22</sub>H<sub>21</sub>NO<sub>2</sub>S; λ<sub>max</sub> (nujol) 3300, 1650, 1590, 750, 690 cm<sup>-1</sup>

δ (CD<sub>3</sub>COOD) 7.3 (18H,s) 5.3 (2H,s) 4.2 (1H,t) 2.8 (2H,d) ppm.

#### Preparation of S-(Trt)cysteine methyl ester oxalate (16)

S-(Trt)cysteine acetate (4.11g, 10mmol) was dissolved in methanol (50ml), to this solution was added p-toluene sulphonic acid monohydrate (3.8g, 20mmol). The mixture was refluxed for 24 hrs. After removing the solvent *in vacuo*, the resulting residue failed to crystallise on trituration with ether. The product was dissolved in ethyl acetate and oxalic acid (5.4g, 0.04mol) was added with stirring. The mixture was then left to stand at 4°C overnight. A white solid precipitated out, which was filtered and washed well with ethyl acetate.

$\lambda_{\text{max}}$  (nujol) 3400, 3000 (broad), 1740, 1440, 1240, 710, 690  $\text{cm}^{-1}$ ;  $\delta_{\text{H}}$ (DMSO- $d_6$ ) 7.3 (18H,s), 6.2 (2H,s), 4.1 (1H,t), 3.7 (3H,s), 2.7 (2H,d) ppm.

### 11.1.3. Synthesis of peptidic compounds

#### General method A for peptide coupling<sup>110</sup>

A solution of the amine protected amino acid (1 equivalent) in dichloromethane (100ml) was cooled in a card-ice/ethylene glycol bath to  $-10^{\circ}\text{C}$ . Triethylamine (1 equivalent) and ethyl chloroformate (1 equivalent) were added to the solution and the mixture was stirred for 20 mins. Then a precooled solution of the carboxyl protected amino acid (1 equivalent) in dichloromethane (100ml) was added to the above mixture and the reaction was left to stir for 3 hours at  $0^{\circ}\text{C}$  and then overnight at room temperature. The reaction mixture was then transferred to a separating funnel and washed with water (100ml), 5%  $\text{NaHCO}_3$  (100ml), 1M HCl (100ml), and water (100ml) again. The organic phase was then dried over sodium sulphate, the mixture was filtered and the solvent was removed by rotary evaporation *in vacuo*.

#### General method B for peptide coupling<sup>98</sup>

The amino protected amino acid (1 equivalent) and the carboxyl protected amino acid (1 equivalent) were dissolved in dry dichloromethane (100ml), the mixture was cooled in an ice-water bath at  $0^{\circ}\text{C}$  and DCC (1 equivalent) was added. The solution was stirred at  $0^{\circ}\text{C}$  for 1 hour and then at room temperature for 5 hrs. The reaction mixture was then left to stand at  $0^{\circ}\text{C}$  overnight. The dicyclohexylurea which precipitated out was filtered off and the filtrate was washed with 5%  $\text{NaHCO}_3$  (100ml), water (100ml), 1M HCl

(100ml), and water (100ml) again. The solution was then dried over sodium sulphate and the dichloromethane was removed by rotary evaporation *in vacuo*.

### General method C for peptide coupling<sup>87</sup>

The amine protected amino acid (5mmol) was dissolved in DMF (30ml). To the solution was added the carboxyl protected component (5mmol) and triethylamine (5mmol), followed by EDIC (5.5mmol). The reaction was left to stir at room temperature overnight. The solvent was then removed from the reaction mixture *in vacuo*. The resulting oily residue was redissolved in CH<sub>3</sub>Cl (50ml) and the solution was washed with 5% NaHCO<sub>3</sub> (50ml), water (50ml), 1M HCl (50ml), water (50ml). After drying over sodium sulphate, the chloroform was removed from the mixture by rotary evaporation. The resulting product was then recrystallised.

### Preparation of N-CBZ-S-(Bzl)cysteinylmethionine methyl ester (18)

N-CBZ-S-(Bzl)cysteine (6.96g, 0.021mol) was coupled with methionine methyl ester hydrochloride (4.5g, 0.021mol) using method A. On rotary evaporation, a clear viscous oil remained. The liquid was triturated with pet-ether (60-80). The resultant white solid was recrystallised from chloroform/pet-ether (60-80) (50:50).

Yield= 74%; m.p.94-95°C; M<sup>+</sup> 490 gives C<sub>24</sub>H<sub>30</sub>N<sub>2</sub>O<sub>5</sub>S<sub>2</sub>; λ<sub>max</sub> (KBr disc) 3300,2850,1740,1690,1650,1540,1375,1240,700 cm<sup>-1</sup>; δH(CDCl<sub>3</sub>) 7.3 (5H,s) 7.2 (5H,s) 5.1 (2H,s) 5.0 (2H,s) 4.7 (1H,q) 4.6 (1H,q) 3.7 (3H,s) 2.5 (2H,dd) 2.4 (2H,t) 2.2 (2H,t) 2.1 (3H,s) ppm;

### Preparation of N-CBZ-methionyl-S-(Trt)cysteine methyl ester (10)

N-CBZ-methionine (1.42g, 5mmol) was dissolved in DMF (30ml), along with S-(Trt)cysteine methyl ester (1.88g, 5mmol), and triethylamine (0.5g, 5mmol). EDIC (1.4g, 5.5mmol) was then added and the reaction was stirred at room temperature overnight. After this time the solvent was removed by rotary evaporation *in vacuo*. The pale yellow oily residue remaining was dissolved in  $\text{CHCl}_3$  (50ml) and the solution was washed separately with water (50ml), 5%  $\text{NaHCO}_3$  (50ml), 1M HCl (50ml) and then water (50ml) again. After drying over sodium sulphate, the solvent was removed by rotary evaporation to leave a pale yellow oil. The oil was washed with pet-ether (60-80) and then allowed to dry *in vacuo*. TLC in chloroform gave 1 spot  $R_f=0.8$ .

yield=71%; m.p. 95-97°C;  $[M+H]$  643 gives  $\text{C}_{36}\text{H}_{38}\text{N}_2\text{O}_5\text{S}_2$ ;

$\lambda_{\text{max}}$  gives 3300, 3050, 2950, 1760, 1670, 1520, 1490, 1375, 1240, 740, 690  $\text{cm}^{-1}$ ;

$\delta\text{H}(\text{CDCl}_3)$  7.3 (5H,s) 7.2 (18H,s) 6.6 (1H,d) 5.6 (1H,d) 5.0 (2H,s) 4.4 (1H,q) 4.3 (1H,q) 3.5 (3H,s) 2.5 (2H,dd) 2.4 (2H,dd) 2.35 (2H,t) 2.0 (3H,s) ppm.

### Preparation of Bis(N-CBZ-methionyl)cystine methyl ester (22)

N-CBZ-methionine (2.84g, 10mmol) was reacted with cystine dimethyl ester dihydrochloride (1.70g, 5mmol) according to method B, using N-hydroxysuccinimide (1.20g, 10.5mmol) as a catalyst. The reaction yielded a white powdery solid, which was recrystallised from chloroform/ diethyl ether (50:50).

Yield= 31% ; m.p. 159-161°C (lit. 157-159°C<sup>82</sup>);  $[M+H]$  799 gives  $\text{C}_{34}\text{H}_{46}\text{N}_4\text{O}_{10}\text{S}_4$

$\lambda_{\text{max}}$  (KBr) 3300, 1735, 1680, 1645, 1250, 720, 680  $\text{cm}^{-1}$ ;  $\delta\text{H}(\text{CDCl}_3)$  7.3 (10H,s) 5.0 (4H,s) 4.6 (4H,q) 3.7 (6H,s) 3.0 (4H,dd) 2.6 (4H,m) 2.4 (4H,m) 2.1 (6H,s) ppm.

### Preparation of N-Fmoc-S-(Trt)cysteinyl glycine ethyl ester (30)

N-Fmoc-S-(Trt)-cysteine (0.25g,0.5mmol) was reacted with glycine ethyl ester hydrochloride (0.07g,0.5mmol) in DMF (30ml) using the catalyst HOBT (0.07g,0.5mmol) according to method C. A white powder resulted which was recrystallised from chloroform/petroleum-ether (60-80) to give a white crystalline solid.

Yield=56%; m.p. 205-207<sup>o</sup>C; [M+H] 671 gives C<sub>41</sub>H<sub>38</sub>N<sub>2</sub>O<sub>5</sub>S;

$\lambda_{\max}$  (KBr disc) 3450,3050,1710,1650,1510,1210,770,690cm<sup>-1</sup>;  $\delta$ H(CDCl<sub>3</sub>)

9.3 (1H,s),7.3 ( 8H,m) 7.2 (15H,m) 5.7 (2H,d), 4.7 (1H,t), 2.7(2H,d) ppm.

### Preparation of [N-Fmoc-S-(Trt)-cysteinamido]-4-thioanisole (25)

p-Aminothioanisole (0.42g,3mmol) was reacted with N-Fmoc-S-(Trt)cysteine (1.74g, 3mmol) as in method B with HOBT (0.42g, 3mmol) as catalyst. After rotary evaporation an oil was left. Analysis by TLC in chloroform showed 3 spots. The spot at R<sub>f</sub> =0.9 was identified as the desired compound. Pure product was isolated by flash chromatography with ethyl acetate as the mobile phase. Rotary evaporation yielded a white powdery solid.

Yield =25.7%; m.p. 130-132<sup>o</sup>C; [M+H] = 707 gives C<sub>44</sub>H<sub>38</sub>N<sub>2</sub>O<sub>3</sub>S<sub>2</sub>;

$\lambda_{\max}$  (nujol) 3280,1660,1580,1260,740,690cm<sup>-1</sup>;  $\delta$ H(CDCl<sub>3</sub>) (300MHz)

7.7-7.5 (4H,dd) J=3Hz, 7.3 (8H,m), 7.2(18H,m), 5.3(2H,dd), 4.4(1H,t) 4.35 (1H,s) 4.2 (1H,t), 2.7(2H,d), 2.4 (3H,s)ppm; (Found C,74.8; H,5.79; N,3.98; S,9.20; Requires C, 74.9; H, 5.4; N, 3.95; S, 9.06.)

### Preparation of N-2,2-dimethylthiazolidinecarboxylic acid hydrochloride (20)

Cysteine hydrochloride (10g, 0.063mol) was stirred in acetone (2L) in a water bath at 60°C for 6 hrs. The hydrochloride salt dissolved slowly and at the same time crystalline flakes began to separate from the solution. The mixture was reduced in volume to approximately 200ml by rotary evaporation *in vacuo* and then cooled overnight at 0°C. The resulting crystalline material was collected by filtration.

Yield=76% ; m.p. 160-165°C (lit. 165-168°C<sup>132</sup>);

M+161 gives C<sub>6</sub>H<sub>11</sub>NO<sub>2</sub>S (HCl); λ<sub>max</sub> (nujol) 3400, 3000 (b), 2500 (b), 1730,

1550 cm<sup>-1</sup>; δH(DMSO-d<sub>6</sub>) 10.3 (1H,bs), 4.8 (1H,t), 3.5(2H,d), 3.2 (1H,bs), 1.9 (6H,s) ppm.

### Preparation of N-2,2-dimethyl-3-formyl-thiazolidinecarboxylic acid (21)

Acetic anhydride (7ml) was added dropwise to a stirring solution of 98% formic acid (21ml), containing N-2,2-dimethylthiazolidine carboxylic acid hydrochloride (2.5g, 0.015mol) and sodium formate (1.0g, 0.0125mol) at a temperature between 0°C. After stirring at room temperature for 6 hrs, ice-water (20ml) was then added to the mixture and a white solid precipitated out. The product was isolated by filtration.

Yield= 85%; m.p. 223-225°C. (lit. 221-225.5°C<sup>132</sup>);

M<sup>+</sup> 189 gives C<sub>7</sub>H<sub>11</sub>NO<sub>3</sub>S; λ<sub>max</sub> (nujol) 3450, 1720, 1610, 1240 cm<sup>-1</sup> ; δH(DMSO-d<sub>6</sub>) 8.3 (1H,s), 8.15 (1H,d), 4.7 (1H,t), 3.2 (2H,d), 1.7 (1H,t) ppm.



### Preparation of [N-2,2-dimethyl-3-formylthiazolidinecarboxy]methionine methyl ester (19)

N-2,2-dimethyl-3-formyl-2,2-thiazolidinecarboxylic acid (0.95g, 5mmol) and methionine methyl ester (0.99g, 5mmol) were coupled as in method C. Recrystallisation from chloroform /hexane resulted in colourless needle shaped crystals.

Yield=45%, m.p. 152-155<sup>0</sup>C; M<sup>+</sup> 335 gives C<sub>13</sub>H<sub>23</sub>N<sub>2</sub>O<sub>4</sub>S<sub>2</sub>;

$\lambda_{\max}$  (nujol) 3300,2950,1740, 1660,1640,1530,1240,690 cm<sup>-1</sup>;  $\delta_{\text{H}}(\text{CDCl}_3)$  8.4 (1H,d), 7.5 (1H,d), 5.2 (1H,q), 4.6 (1H,t), 3.8(3H,s), 3.3 (2H,t), 3.2 (2H,t), 2.5 (2H,t), 2.1 (3H,s), 1.9 (6H,d) ppm.

( Found C, 46.42; H, 6.65; N, 8.01; Requires C,46.54; H, 6.91; N, 8.35)

### Preparation of N-CBZ-methionine-4-nitrophenyl ester (62)

N-CBZ-methionine (4.37g,15mmol) and p-nitrophenol (2.5g, 18mmol) was dissolved in ethyl acetate (50ml) and the mixture was cooled in an ice-bath. DCC (3.1g, 15mmol) was added and the mixture was stirred at 0<sup>0</sup>C for 30 mins, and then at room temperature for a further hour. In order to aid precipitation of dicyclohexylurea the mixture was left standing overnight at 0<sup>0</sup>C and then filtered. The resultant filtrate was evaporated to dryness *in vacuo* to leave a solid. Recrystallisation from hot ethanol afforded a white crystalline solid.

Yield=78%; m.p. 99-101<sup>0</sup>C; M<sup>+</sup> 388 gives C<sub>19</sub>H<sub>20</sub>N<sub>2</sub>O<sub>6</sub>S;  $\lambda_{\max}$  (nujol) 3300, 2250, 1750, 1650, 1350, 710, 690 cm<sup>-1</sup>;  $\delta_{\text{H}}(\text{CDCl}_3)$  8.2 (2H,d) 7.2 (2H,d) 7.3 (5H,s) 5.4 (1H,d) 5.1 (2H,s) 4.3 (1H,t) 3.2 (2H,m) 2.6 (2H,d) 2.1 (3H,s) ppm.

### **N-CBZ-prolyl-S-(Bzl)cysteine ethyl ester (33)**

N-CBZ-proline (2.49g, 10mmol) was coupled to S-(Bzl)cysteine ethyl ester (2.25g, 10mmol) according to method B. Trituration with pet-ether (60-80) yielded a cream solid. Purification by flash chromatography produced a single spot  $R_f = 0.6$  in chloroform. Recrystallisation from dichloromethane/ether afforded a white crystalline solid.

yield = 51.4%; m.p. 82-83°C;  $M^+$  470 gives  $C_{25}H_{30}N_2O_5S$ ;

$\lambda_{max}$  (nujol) 3300, 1740, 1700, 1650, 1240, 770  $cm^{-1}$ ;  $\delta H(CDCl_3)$  7.3 (5H,s), 7.25 (5H,s), 5.25 (2H,s), 5.2 (1H,s), 5.15 (1H,s), 5.05 (2H,s), 4.2 (2H,q), 2.9 (2H,d), 2.1 (4H,m), 1.3 (3H,t) ppm

(Found C, 63.95; H, 6.50; N, 6.05; Requires C, 64.10; H, 6.41; N, 5.98)

### **Preparation of N-boc-prolinyl-S-(Bzl)cysteine methyl ester (34)**

N-boc-proline (5.38g, 0.025mol) was coupled to S-(Bzl)cysteine methyl ester (6.54g, 0.025mol) under conditions employed in method B. The reaction produced an oil which by T.L.C. in chloroform showed 3 spots..

Purification by flash chromatography isolated spot 2  $R_f = 0.5$  as a yellow oil.

Yield = 34.9%,  $M^+$  427 gives  $C_{21}H_{30}N_2O_5S$

$\lambda_{max}$  3300, 2950, 1760, 1650, 1500, 770, 690  $cm^{-1}$ ;  $\delta H(CDCl_3)$  7.2 (5H,s), 5.2 (2H,s), 4.7 (1H,bs), 4.15 (1H,q), 3.6 (2H,d), 3.7 (3H,s), 3.4 (2H,t), 2.85 (2H,dd), 2.7 (2H,dd), 1.4 (9H,S) ppm.

### **Preparation of bis( N-Fmoc-prolyl)cystine dimethyl ester (35)**

N-Fmoc-proline (1.69g, 5mmol) was coupled to cystine dimethyl ester (1.70g, 5mmol) using method C. The reaction produced a clear oil which showed 2 spots by TLC in

dichloromethane. The spot,  $R_f = 0.7$  was isolated by flash chromatography to give a white crystalline solid.

Yield= 45% ; m.p. 73-75<sup>0</sup>C; [M+H] 907 gives  $C_{48}H_{50}O_{10}N_4S_2$ ;

$\lambda_{max}$  (KBr disc) 3300, 3050, 2950, 1730, 1660, 1510, 1400, 730,770  $cm^{-1}$ ;

$\delta H$ (DMF- $d_7$ ) 7.9 (2H,d) 7.7 (2H,d) 7.4 (4H,m), 5.6 (1H,d) 4.2 (3H,s) 2.1 (4H,m) ppm.

### Preparation of bis(N-boc-propyl)cystine dimethyl ester (36)

N-boc-proline (4.30g, 2mmol) was coupled with cystine dimethyl ester (3.41g,1mmol) as in method B in the presence of the catalyst Nhs (0.063g,0.56mmol). The solvent dichloromethane was replaced by DMF (15ml). After the reaction was complete the DMF was removed by rotary evaporation *in vacuo*. The oily residue was redissolved in ethyl acetate (50ml). After leaving to stand at 0<sup>0</sup>c overnight, the mixture was filtered. The mixture was washed consecutively with 10% citric acid (50ml), water (50ml), 1M  $NaHCO_3$  (50ml) and water (50ml). After drying over sodium sulphate, the organic layer was evaporated down to dryness. The product was isolated by flash chromatography using ethyl acetate as the mobile phase to yield a fluffy white solid.

Yield= 35.4%, m.p. 65-68<sup>0</sup>C;  $M^+$  662 gives  $C_{28}H_{46}N_4O_{10}S_2$ ;

$\lambda_{max}$  (nujol) 3300,1690,1650,1550,1375 $cm^{-1}$ ;  $\delta H$  ( $CDCl_3$ ) 7.2 (2H,s) 4.8 (2H,t),

4.2(2H,s), 3.7 (4H,s), 3.4 (2H,t), 3.1 (4H,d) 2.0 (4H,mt), 1.4 (18H,s) ppm.

(requires C: 50.75; H: 7.0; N: 8.46; S: 9.69; found C: 50.67; H: 7.29; N: 8.63; S: 9.62)

### Preparation of N-boc-prolylcysteine methyl ester (39)

(Bis-N-boc-prolyl)cystine dimethyl ester (0.66g, 1mmol) was dissolved in conc HCl (5ml). Zinc dust was added over 10 mins whilst cooling in an ice/water bath. The mixture was stirred for a further five mins. After this time, the mixture was filtered and *p*-toluene sulphonic acid hydrate (0.5g) was added in water (5ml). On cooling white solid crystallised out, which was collected by filtration.

Yield=35%; m.p. 67-70°C;  $M^+$  320 gives  $C_{13}H_{24}N_2O_5S$

$\lambda_{max}$  (KBr) 3300, 3050, 2900, 1690, 1650, 1540, 1375  $cm^{-1}$ ;  $\delta_H$  (DMSO- $d_6$ ) 7.0 (1H,s) 4.7 (1H,q) 4.2 (1H,t) 3.7 (3H,s) 3.3 (2H,t) 2.8 (1H,t) 2.4 (2H,d) 2.2 (4H,m) 1.4 (9H,s) ppm

### Preparation of N-boc-prolyl-S-(Nmc)cysteinylglycine ethyl ester (40)

#### 1) Preparation of N-boc-prolyl-N-hydroxysuccinimide ester (49)

A solution of N-boc-proline (2.15g, 10mmol) and Nhs (1.15g, 10mmol) in dry THF (20ml) was cooled in an ice-water bath and DCC (2.06g, 10mmol) was added with stirring. After stirring the mixture for 30 mins, the reaction was left to stand overnight at 0°C. The N,N-dicyclohexylurea which separated out was filtered off and the solvent removed *in vacuo*, to yield an oil. On addition of propan-2-ol, a white solid precipitated out slowly. The crude product was then recrystallised from propan-2-ol.

Yield= 84.3%; m.p 145-147°C;  $M^+$  288 gives  $C_{12}H_{20}N_2O_6$ ;

$\lambda_{max}$  (KBr disc) 2924, 2854, 1818, 1788, 1745, 1701, 1458, 1373, 1205  $cm^{-1}$ ;

$\delta_H$  (CDCl<sub>3</sub>) [60MHz] 4.2 (1H,t) 3.3 (2H,t) 3.1 (4H,t) 2.9 (2H,m) 1.4 (9H,s) ppm.

## 2) Preparation of N-boc-prolyl-S(Nmc)cysteine (51)

N-boc-proline-N-hydroxysuccinimide ester (3.12g, 10mmol) was dissolved in absolute ethanol (35ml) and the solution is added to a solution of S-(Nmc)cysteine (1.64g, 10mmol) and NaHCO<sub>3</sub> (1.68g, 20mmol) in water (25ml). The mixture was stirred at room temperature overnight. The mixture was then concentrated *in vacuo* and acidified to pH=2 with conc HCl. An oil separated which solidified on cooling in an ice-water bath. The solid was filtered off and washed with a small volume of cold water and then dried over P<sub>2</sub>O<sub>5</sub>.

Yield= 53% ; M<sup>+</sup> 322 gives C<sub>14</sub>H<sub>24</sub>N<sub>3</sub>O<sub>5</sub>S;

λ<sub>max</sub> (nujol) 3320, 3000, 1740, 1700, 1660, 1550 cm<sup>-1</sup>; δH(D<sub>2</sub>O) [250MHz] 9.7 (1H, bs) 7.6 (1H, bs) NH 4.7 (1H, dd) Pro α-H 4.4 (1H, dd) Cys α-H 3.4 (4H, m) Pro-CH<sub>2</sub> 3.0 (1H, d) and 2.8 (1H, d) Cys-β CH<sub>2</sub> 2.3 (3H, s) N-CH<sub>3</sub> 1.7 (2H, m) Pro-γCH<sub>2</sub> 1.4 (9H, s) C(CH<sub>3</sub>) ppm.

## 3) Preparation of N-boc-prolyl-S-(Nmc)cysteinylglycine ethyl ester (40)

N-boc-prolyl-S-(Nmc)cysteine (0.96g, 3mmol) was reacted with glycine ethyl ester hydrochloride (0.42g, 3mmol) as in method C. After rotary evaporation, a clear oily residue. Trituration with ethyl acetate yielded peptidic material. Further purification was achieved by recrystallisation from chloroform/ diethyl ether to give a white fluffy solid.

Yield=38%; m.p. 94-96°C; M<sup>+</sup> 460 gives C<sub>19</sub>H<sub>32</sub>N<sub>4</sub>O<sub>7</sub>S;

λ<sub>max</sub> (nujol) 3320, 1745, 1690, 1660, 1540, 1240 cm<sup>-1</sup>; δ( D<sub>2</sub>O) 7.4 (1H, d) NH 6.3 (1H, broad) NH 4.7 (1H, dd) Pro α-H 4.4 (1H, dd) Cys α-H 4.0 (2H, d) Gly-CH<sub>2</sub> 3.6 (2H, q) O-CH<sub>2</sub> 3.4-3.2 (4H, m) Pro-CH<sub>2</sub> 2.9 (2H, m) Cys-CH<sub>2</sub> 2.3 (3H, s) N-CH<sub>3</sub> 1.9 (2H, m) Pro-CH<sub>2</sub> 1.4 (9H, s) C(CH<sub>3</sub>) 1.2 (3H, s) r-CH<sub>3</sub>

## 11.2 Stability study on sulphonium centre in the presence of thiol by 300 MHz $^1\text{H}$ NMR

Cysteine (1.21g, 10mmol) and S-methylmethionine sulphonium iodide (2.90g, 10mmol) were dissolved in  $\text{D}_2\text{O}$  (5ml) in an NMR tube. A spectrum was recorded immediately after mixing. The tube was then incubated in a waterbath at  $37^\circ\text{C}$ . Spectra were initially recorded at 30 min intervals up until 3 hrs, then the reaction was left until 24 hrs elapsed before a final spectrum was recorded. Also reference spectra of cysteine and S-methylmethionine sulphonium iodide were recorded separately.

reaction spectrum

$\delta$  ( $\text{D}_2\text{O}$ ) 4.65 (HOD), 3.90 (1H,t), 4.10 (1H,t), 3.50 (2H,m), 3.10 (2H,q), 3.0 (3H,s),  
2.40 (2H,dd)

reference spectra

S-methylmethionine sulphonium iodide

$\delta$  ( $\text{D}_2\text{O}$ ) 4.65 (HOD), 3.90 (1H, t), 3.50 (2H,m), 3.0 (3H,s) 2.40 ( 2H,dd)

cysteine

$\delta$  ( $\text{D}_2\text{O}$ ) 4.10 (t), 3.10 (q)

## 11.3 Study on methylation reaction to produce S-methylmethionine sulphonium iodide

Methionine (0.15g, 1mmol) and methyl iodide (1ml) were dissolved in  $\text{CD}_3\text{COOD}$  (0.5ml) in an NMR. tube. Using a Bruker 300MHz NMR spectrometer. A  $^1\text{H}$ -spectrum was recorded initially and then at the following times later

1) 30 mins 2) 1 hour 3) 2 hrs and then at hourly intervals up to 24 hrs.

## 11.4 Solid phase peptide synthesis

Solid phase peptide syntheses were carried out on 1% cross-linked divinylbenzene resin with an ABI 430A automated peptide synthesiser. The loaded resin and all protected amino acids were purchased from Novabiochem Ltd. and were used directly with no further purification. All solvents and other reagents were purchased from ABI and used directly with no further purification.

### General method for solid phase synthesis

The resin was initially swollen in DMF for approximately 30 mins.

The Reaction cycle:

The swollen resin with the initial amino acid already attached (1equiv/mmol) was loaded into the reaction vessel. The amino acid was deprotected by washing with 20% piperidine in DMF (20mlx3). The next amino acid (2mmol) was introduced automatically into a second reaction vessel along with DIC (1mmol). The reaction produced a symmetrical anhydride which was then transferred into the main reaction vessel and was left to react with the deprotected resin bound amino acid for 2 hrs. The resin was rinsed with DMF and was then reacted with a further equivalent of Fmoc-amino acid, in the presence of HOBT (1mmol) and DIC (1mmol). The mixture was again left to stand for 2 hrs. Finally the resin was incubated with acetic anhydride (1mmol) in DMF for 1 hour to ensure any remaining free amino groups are rendered unreactive. After washing the resin with DMF (2x50ml), the cycle was complete. In order to add further amino acids the process was repeated. After the final amino acid was added, the resin was washed with DMF and

removed from the reaction vessel. The resin was then washed consecutively with DMF, chloroform and finally ether. After air drying the resin was ready for deprotection.

### **Cleavage of Cys-Ala-Pro-Cys-Leu-Pro-Ala-Gly (53) and Gly-Phe-Pro-Cys-Pro-His-Phe-Ser (54)**

The synthesised peptides were removed from the resin and deprotected with TFA 95% (spectrophotometric grade) (25ml) containing anisole (1.5ml) and ethyl methyl sulphide (1.0ml). The mixture was incubated for between 4 and 5 hrs under nitrogen. The TFA was then removed by rotary evaporation under high vacuum at 30°C. The resulting oily residue was triturated with ether (10ml). The precipitated peptide was taken up in water (100ml) and washed with ether (3x100ml). The aqueous fraction was then taken and lyophilised. The crude free peptides were then purified by HPLC on a RP-C8 column (Aquapore internal diameter 10mm, 7-10µ, 300A) eluting with a linear gradient, 10-90% acetonitrile (0.001% TFA)/ water (0.001% TFA) over 30 mins at a flow rate of 5ml/min.

### **Cys-Ala-Pro-Cys-Leu-Pro-Ala-Gly (53)**

[M+H] 730 gives  $C_{30}H_{49}N_8O_9S_2$

$\lambda_{max}$  (nujol) 3350,3300,3250,1750, 1700, 1650,1550,1400, 1240 $cm^{-1}$

$\delta H$  (DMSO- $d_6$ ) [300MHz] 8.8 (1H,d) N-H 8.25 (2H,bs) N-H 8.0 (1H,d) N-H 7.7

(2H,d) N-H 7.1 Ar-H 4.6 (2H,t)  $\alpha$ -CH 4.5 (2H,q)  $\alpha$ -CH 4.2 (2H,dd)  $\alpha$ -CH 3.8 (1H,d)

$\alpha$ -CH 3.7 (2H,t) N-CH<sub>2</sub> 3.3 (HOD) 2.3 (2H,d) r-SH 1.8 (2H,m) r-CH<sub>2</sub>-r 1.7 (2H,m)

r-CH<sub>2</sub>-r 1.3 (6H,d)  $\alpha$ C-CH<sub>3</sub> 0.9 (2H,s) r-CH<sub>3</sub>

HPLC  $R_t$  = 14.8mins. [Conditions 90% CH<sub>3</sub>CN(0.001% TFA): 10% Water (0.001% TFA)]



### Gly-Phe-Pro-Cys-Pro-His-Phe-Ser (54)

[M+H] 891 gives  $C_{42}H_{54}N_{10}O_{10}S$ .

$\delta$  ( $D_2O$ ) 8.65 (1H,s)  $\underline{CH}$ -im 7.3 (10H,s) Ar- $\underline{H}$  5.0 (2H,m)  $\alpha$ - $\underline{CH}$  4.7 (1H,q)  $\alpha$ - $\underline{CH}$  4.5 (2H,dd)  $\alpha$ - $\underline{CH}$  3.9 (2H,d)  $\alpha$ - $\underline{CH}$  3.8 (4H,d)  $\underline{CH}_2$ -Ar 3.25 (2H,d)  $\underline{CH}_2$ -S 3.0 (4H,dt)  $\underline{CH}_2$ - $\alpha$ C 2.85 (4H,m)  $\underline{CH}_2$ -N 2.0 (4H,toft) r- $\underline{CH}_2$ -r

HPLC  $R_t$  =15.45mins [Conditions 90%  $CH_3CN$ (0.001% TFA): 10% Water (0.001% TFA)]

### Cleavage of Ser-Pro-Cys-Met-Arg-Leu-Ile-His-Leu (55)

Cleavage of the free peptide from the resin was achieved with 95% TFA (spectrophotometric grade) (25ml) containing thioanisole (1.5ml), ethyl methyl sulphide (1.5ml) and anisole (1.5ml). Workup and purification was as above.

[M+H] 1069 gives  $C_{46}H_{80}N_{14}O_{11}S_2$ ;

$\lambda_{max}$  (KBr) 3400,3150,3050, 2950,1720,1700,1660,1640,1500,1375,1240 $cm^{-1}$ ;

$\delta$  (DMSO- $d_6$ ) [300MHz] 8.3 (1H,d)  $\underline{NH}$ -im, 8.1(1H,dd)  $\underline{NH}$ -C 7.4 (1H,dd)  $\underline{CH}$ -im 7.2 (1H,d) R- $\underline{OH}$  7.1 (2H,m) C- $\underline{NH}_2$  4.5 (1H,q)  $\alpha$ - $\underline{CH}$  4.4 (1H,m)  $\alpha$ - $\underline{CH}$  4.35 (1H,m)  $\alpha$ - $\underline{CH}$  4.25 (1H,d)  $\alpha$ - $\underline{CH}$  4.1 (2H,q)  $\alpha$ - $\underline{CH}$  3.5 (b) HOD 3.1 (2H,t)  $\underline{CH}_2$ - $\alpha$ C 3.0 (2H,d)  $\underline{CH}_2$ -N 2.2 (3H, s)  $\underline{CH}_3$ -S 1.9 (2H,m) r- $\underline{CH}$ -r 1.5 (4H,m) r- $\underline{CH}_2$ -r 0.9 (12H, d) r- $\underline{CH}_3$  0.8 (2H,dd) r- $\underline{CH}_2$ -r

HPLC  $R_t$  =10.60mins. [Conditions 90%  $CH_3CN$ (0.001% TFA): 10% Water (0.001% TFA)]

### **Coupling of Fmoc-Asp(O-tBu)-OH (56) to the p-benzyloxybenzylalcohol resin**

Fmoc-Asp(O-tBu) (3.9g, 9.48mmol) was dissolved in DMF (peptide synthesis grade) (30ml) and this was mixed with a solution of DIC in DMF (0.74ml, 4.74mmol). The benzyloxybenzyl resin (2.0g, 1.58mmol) was dissolved in DMF (15ml) and the mixture was sonicated for 30mins. Dimethylaminopyridine (30mg) was added to the mixture followed by the symmetrical anhydride formed earlier. The reaction was then left for 1 hour. After this time the resin was filtered off from the reaction mixture and after washing with DMF (3x20ml), chloroform (2x20ml) and ether (2x20ml). The resin was then dried *in vacuo*.

Yield=2.35g

Loading 84%

Substitution= 0.52mmol/g

### **Capping of free alcohol groups in the resin**

The resin bound peptide (2.35g) was added to CH<sub>2</sub>Cl<sub>2</sub> (30ml) and the suspension was cooled in an ice bath. Pyridine (0.75ml) was added, followed by benzoyl chloride (0.75mmol). After 1 hour the mixture was filtered washed with dichloromethane (2x50ml) and ether (2x50ml) and was then dried *in vacuo* in a dessicator.

### **Cleavage of Pro-Cys-His-Met-Asp (37)**

The peptide was cleaved with 95% TFA (10ml) in the presence of anisole (1ml) and ethyl methyl sulphide (1ml). Work up and purification were as for the other peptides

synthesised. Attempts at isolation of the free peptide failed. The compound was only able to be isolated as the S-trityl derivative.

[M+H]<sup>+</sup> 845 gives C<sub>42</sub>H<sub>48</sub>N<sub>7</sub>O<sub>8</sub>S<sub>2</sub>;

$\lambda_{\text{max}}$  (KBr) 3250,3050,2900,1710,1660,1640,1500,1375,1250 cm<sup>-1</sup>;  $\delta$ H (DMSO-d<sub>6</sub>)

[250MHz] 8.9 (1H,q J =10Hz) 8.5 (1H,dd J = 8Hz) 7.3 (18H,s) 4.65 (1H,q) 4.5 (1H,m)

4.35 (2H,m) 4.15 (1H,t) 3.4 (HOD) 2.6 (2H,m) 2.3 (2H,d) 2.2 (2H,d) 2.0 (3H,s) 1.9

(1H,bs)

## 11.5 Biological methods

### 11.5.1 DNA MTASE assay

Each assay contained 20 $\mu$ g of *M.Lysodeikticus* DNA (2.0mg/ml), 1.0 $\mu$ Ci of [<sup>3</sup>H]-SAM (83-85 Ci/mmol) and the peptide (1mmol) dissolved in 100 $\mu$ l of assay buffer. After incubation at 37<sup>o</sup>C for 4 hours in a shaking waterbath, the reaction was stopped by adding 200 $\mu$ l of 1.0M NaOH and 40 $\mu$ g of calf thymus DNA (2.0mg/ml) to act as a carrier. The mixtures were then heated to 65<sup>o</sup>C for 20 mins in a waterbath. After this time the reaction mixture was cooled in an ice/water bath and the DNA was precipitated out by adding 80 $\mu$ l of 5.0M HClO<sub>4</sub>. After 1 hour the DNA was filtered onto a Whatman GF/C glass microfibre filter and washed with cold 5% trichloroacetic acid and absolute ethanol. After air drying the filters overnight the radioactivity was determined by adding 8ml of the toluene/PPO/POPOP scintillant and counting on a Packard Tri-carb 2000CA liquid scintillation counter.

### 11.5.2 DNA hydrolysis

DNA was placed in a pyrex glass tube (150x18mm) to which was added 1-2ml of 90% formic acid and the tube was then sealed using an oxygen/natural gas flame torch. Hydrolysis of the DNA was then carried out at 180°C for 25mins in an oil bath. The seal was then broken and the formic acid evaporated off under a stream of nitrogen gas, whilst warming in a waterbath at 60°C. The residue was then redissolved in 0.1M HCl and the equivalent of 20µg of DNA was injected onto the HPLC column.

### 11.5.3 HPLC analysis

Bases were separated using a Whatman partisil 10-SCX column. The composition of the mobile phase changed from buffer A (8% methanol adjusted to pH 4.0 with formic acid) to buffer B (0.4M ammonium formate, 8% methanol, pH 4.0) over a period of 40 mins using waters gradient \* 9. The flow rate was 2.0ml/min.

### 11.5.4 Assay for agents methylating cytosine at position 5

#### 1) Methylation of lambda phage in vitro:

##### a) In the absence of DNA MTASE

Lambda phage DNA (5µg) was treated with the indicated concentrations of sulphonium compound for 30mins at 37°C in a total volume of 200µl. The solution was then saturated with sodium acetate by adding 0.1 volumes of 2.5M sodium acetate and the DNA was precipitated with 2.5 volumes of cold absolute ethanol. The DNA was redissolved in water and the precipitation process was

repeated a further two times, then the DNA was washed 3 times with absolute ethanol. The DNA samples were then air dried ready for restriction enzyme analysis.

b) in the presence of DNA MTASE

Lambda phage DNA (5µg) was treated with the indicated concentration of sulphonium compound dissolved in the MTASE assay buffer and the DNA MTASE preparation made up to 100µl with assay buffer. The mixture was incubated at 37°C for 30 mins, after which the reaction was terminated by adding 500µl sterile 10mM tris-HCl pH 7.5, 0.2M NaCl (500µl) and a phenol reagent (500µl). The upper aqueous layer was collected and the DNA was precipitated with cold isopropanol (1ml). The suspension was kept at -20°C for 30 mins. Traces of phenol were removed from the DNA, by mixing samples with a volume of ether saturated with an equal volume of water and allowing them to separate slowly over a period of 5 mins. The upper layer was removed and discarded and this process was repeated twice. Traces of ether were then removed by heating in a waterbath at 70°C for 10 mins. The DNA was redissolved in water and the precipitation process repeated by adding 0.1 volumes of 2.5M sodium acetate and 2.5 volumes of cold absolute ethanol. The DNA was then washed 3 times with absolute ethanol and then air dried ready for restriction enzyme analysis.

2) Restriction enzyme assay

Lambda phage DNA (5µg), treated with compound as described above. The mixture was then digested with 10 units of either MspI or HpaII in the appropriate restriction buffer as recommended by the suppliers, in a total volume of 20µl for 4 hours at 37°C. The reaction was then stopped by adding 4µl of (6xconcentration) loading buffer and then the mixture

was heated in a waterbath at 65°C for 10 mins. Samples were then analysed on a 1.2% agarose gel made up in buffer containing 0.5µg/ml ethidium bromide and run at 5 volts per cm on an AE-610 submerged agarose electrophoresis kit. When the tracking dye was 1-2 cm from the end of the gel, electrophoresis was stopped and the DNA was revealed using a 302 nm transilluminator and then photographed with a polaroid MP4 camera and type 665 black and white film.

## 11.6 Cytotoxicity assays

### 11.6.1 Cells

The compounds were tested for their ability to inhibit the growth of A549 cells. A549 cells are a human-derived lung adenocarcinoma, said to originate from a type II alveolar cell line.

### 11.6.2 Methods

A549 cells were cultured in medium under an atmosphere of 10% CO<sub>2</sub> in air and 100% humidity at 37°C.

Cells were seeded at a density of 5 x 10<sup>4</sup> cells per well (2ml) and were exposed to drug. The final concentration of solvent in the medium was 0.5% v/v and had no effect on the growth of cells. The cells were exposed to drug for 72 hrs and then the cell population was counted on a coulter counter.

The IC<sub>50</sub> is the concentration of test compound at which the growth of treated cells is inhibited by 50% compared to control cells.

The cytotoxicity tests on A549 cells were carried out by H.Hussey in the Cancer research campaign laboratories at the University of Aston.

## **Bibliography**

- 1) The quantitative separation of purines, pyrimidines and nucleosides by paper chromatography  
Hotchkiss R.D.  
J. Biol. Chem. , **1948**, 175, 315
- 2) A new theory of Carcinogenesis  
Holliday F.  
Br. J. Cancer **1979**, 40, 513
- 3) Enzymatic synthesis of deoxyribonucleic acid. Influence of bacteriophage T2 on the synthetic pathway in host cells .  
Kornberg A., Zimmerman S., Kornberg S.R., Josse J.  
Proc. Natl. Acad. Sci. USA **1959**, 45, 772
- 4) Adams R.L.P., Burdon J.  
Molecular biology of DNA methylation **1983** Wiley N.Y.
- 5) CpG rich islands and the function of DNA methylation  
Bird A. P.  
Nature **1986**, 321, 209
- 6) DNA methylation : site specific methylations cause gene activation  
Doerfler W.  
Angew. Chem. Int. Ed. **1984** , 23, 919
- 7) Separation of calf thymus DNA into fractions of different composition  
Chargaff E., Crampton C., Lipshitz R.  
Nature **1953**, 172, 289
- 8) 5-MC localised in mammalian constitutive heterochromatin  
Miller O.J., Schnedel W., Allen J., Erlanger B.  
Nature **1974**, 251, 636
- 9) Msp I, an isochizomer of Hpa II which cleaves both methylated and unmethylated Hpa II sites.  
Waalwijk J., Flavell R.A.  
Nucleic Acids Res. **1978**, 5, 3231
- 10) Cellular differentiation , Cytidine analogs and DNA methylation  
Jones P.A., Taylor S.M  
Cell **1980**, 20, 85



- 11) Catalytic mechanism of thymidylate synthetase  
Piglotti A., Santi D.  
Bioorgan. Chem. 1 277 Acad. press Ed. Von Tamelen
- 12) DNA Methylation  
Hattman S.  
The Enzymes **1981**, XIV, 517 Acad. press Ed. Boyer
- 13) Enzymatic DNA methylation in higher eukaryotes  
Drahovsky D, Boehm T.  
Int. J. Biochem. **1980**, 12, 523
- 14) Use of restriction enzymes to study Eukaryotic DNA methylation  
Bird A.P., Southern E.M.  
J. Mol. Biol. **1978**, 118, 27
- 15) Mouse DNA methylase: methylation of native DNA  
Adams R., McKay E., Craig L., Burdon R.  
Biochimica Biophysica acta **1979**, 561, 345
- 16) Cytosine specific Type II DNA methylases. A conserved Enzyme core with a variable target recognising domain.  
Lauster R., Trautner T.A., Noyer-Weidner M  
J. Mol. Biol. **1980**, 206, 305
- 17) DNA methylation . The effect of minor bases on DNA-protein interactions.  
Adams R.  
Biochem. J. **1990**, 265, 309
- 18) Thin layer separation methods for nucleic acid derivatives  
Randerath K., Randerath E.  
Methods in Enzymology **1968**, 12, 323
- 19) Analysis of 5-methylcytosine in DNA II gas chromatography  
Razin A., Sedat J.  
Anal. Biochem. **1977**, 77, 370
- 20) Distribution of 5-MC in chromatin  
Razin A., Cedar H.  
Proc. Natl. Acad. Sci. USA **1977**, 74, 2725
- 21) 5-MC content of rat hepatoma DNA substituted with bromodeoxyuridine.  
Singer J.  
J. Biol. Chem. **1977**, 252, 5509

- 22) Identification of 5-MC, in DNA fragments immobilised on nitrocellulose paper.  
Sono H., Royer H.D., Sager R.  
Proc. Natl. Acad. Sci. USA **1980**, 77, 3581
- 23) Detection of 5-methylcytosine in DNA sequences  
Ohmori M., Tomizawa J., Maxami M.  
Nucleic Acids Res. **1978**, 5, 1479
- 24) Borek E., Srinivisan R.  
Prog. Nucleic Acids Res. **1966**, 5, 157
- 25) DNA methylation and gene function  
Razin A., Riggs A.D.  
Science **1980**, 210, 604
- 26) Selective silencing of eukaryotic DNA  
Sager R., Kitchen R.  
Science **1975**, 189, 426
- 27) DNA modification mechanisms and gene activity during development  
Holliday R., Pugh J.E.  
Science **1975**, 187, 226
- 28) The effect of 5-Azacytidine on DNA methylation in Ehrlich's ascites tumour cells  
Tanaka M., Hibasami H., Nagai J., Ikeda T  
J. Exp. Biol. Med. Sci. **1980**, 58, 391
- 29) DNA methylation : organ specific variations in the methylation pattern within and around ovalbumin and other chicken genes  
Mandel H. , Chambon J.  
Nucleic Acids Res. **1981**, 7, 2081
- 30) DNA methylation in the human  $\gamma$  and  $\beta$  globin  
Van der Ploeg L., Flavell R.  
Cell **1980**, 19, 947
- 31) Reconstitution of deoxyribonuclease I sensitive structure in active genes  
Gazit B., Paret A., Cedar H.  
Proc. Natl. Acad. Sci. USA **1980**, 77, 1787
- 32) Infectivity and methylation of retroviral genomes is correlated with expression in the animal  
Stuhlman H., Jahrer D., Jaenisch R.  
Cell **1979**, 26, 221

- 33) Altering genotype and phenotype by DNA mediated gene transfer  
Pellicer A., Robins D., Wold B., Jackson J., Lowry I., Roberts J., Sin G., Silverstein S., Axel R.  
*Science* **1980**, 209, 1414
- 34) An oestrogen-dependant demethylation at the 5' end of the chicken vitellogenin gene is independent of DNA synthesis  
Wilks A., Seldran M., Jost J.P.  
*Nucleic Acids Res.* **1984**, 12, 1163
- 35) Multiple new phenotypes induced in 10 T<sub>1/2</sub> and 3T3 cells treated with 5-azacytidine.  
Taylor R., Jones P.  
*Cell* **1979**, 17, 771
- 36) Effect of 5-AzaC on differentiation and DNA methylation in human promyelocytic leukaemia cells (HL-60)  
Christman J.K.  
*Cancer Res.* **1983**, 43, 763
- 37) DNA methylation in friend erythroleukemia cells : The effects of chemically induced differentiation and of treatment with inhibitors of DNA methylation  
Christman J.K.  
*Current Topics in Microbiology and Immunology* **1984**, 108, 49 Ed. T. Trautner  
Springer-Verlag
- 38) Unusual properties of DNA from Xanthomonas phage xp12 in which 5MC completely replaces cytosine.  
Ehrlich M., Ehrlich K., Mayo J.  
*Biochimica Biophysica Acta* **1975**, 395, 109
- 39) Hypomethylation of the 5'-region of the calcitonin gene is a property of human lymphoid and acute myeloid malignancies  
Baylin S., Fearon E., Vogelstein B., De bustros A., Sharkis S., Burke P., Staal S., Nelkin B.  
*Blood* **1987**, 70, 412
- 40) Hypomethylation distinguishes genes of some human cancers from their normal counterparts  
Feinberg A., Vogelstein B.  
*Nature* **1983**, 301, 89

- 41) Mouse nuclear satellite DNA: 5-MC content, pyrimidine isoplith distribution and elctron microscope appearance.  
Salomen R., Kaye A.  
J. Mol. Biol. **1969**, 43, 581
- 42) The 5-methylcytosine content of DNA from human tumours  
Gama-Sosa M., Slagel V., Trewyn R., Oxenhandler R., Kuo K., Gehrke G., Ehrlich M.  
Nucleic Acids Res. **1983**, 11, 6883
- 43) Hypomethylation of ras oncogenes in primary human cancers  
Feinberg A., Vogelstein B.  
Biochim. Biophys. Res. Commun. **1983**, 111, 47
- 44) Variable 5-MC levels in human tumour cell lines and fresh paediatric tumour explants  
Flatau E., Bergmann E., Jones P.  
Cancer Res. **1983**, 43, 4901
- 45) The nucleotide sequence of the human int-1 mammary oncogene, evolutionary conservation of coding and non-coding sequences  
Van Ooyen A., Kuvée V., Nusse R.  
EMBO J. **1985**, 4, 2905
- 46) Hypomethylation of DNA in human cancer cells. A site specific change in the C-myc oncogene  
Cheah M., Wallace D., Hoffmann R.M.  
J. Natl. Cancer Inst. **1984**, 73, 1057
- 47) DNA methylation patterns of the calcitonin gene in human lung cancer and lymphomas  
Baylin S.B, Hoppener J., De Bustros A. Sheerlegh D.  
Cancer Res. **1986**, 46, 2917
- 48) Mutagenicity of 5-AzaC and related nucleosides in C3H/10T<sub>1/2</sub> clone 8 and V79 cells.  
Landolph J.R., Jones P.A.  
Cancer Res. **1982**, 42, 817
- 49) 5-AzaC induced uncoupling of differentiation and tumourigenicity in a murine cell line  
Walker C., Ramsey D.F., Stay J.  
J. Natl. Cancer Inst. **1984**, 73, 877
- 50) The tumorigenicity of 5-azacytidine in the male Fischer rat  
Carr B.I, Reilly J., Smith S., Winberg C., Riggs A.D  
Carcinogenesis **1984**, 5(12), 1583
- 51) L- ethionine as an inducer of differentiation in human promyelocytic leukaemia cells  
Mendelson M., Michel J., Gilbert H., Christman J.K  
Cancer Res. **1980**, 40, 3206

- 52) Mechanism of action of rat liver DNA methylase  
Drahovsky D., Morris N.R.  
J. Mol. Biol. **1971**, 57, 475
- 53) Inhibition of DNA methylation by alkylated or depurinated DNA  
Cox F.  
Biochem. Int. **1982**, 5, 787
- 54) Grunberger D., Weinstein I.B.  
Chemical carcinogens in DNA 2, 59 CRC press.
- 55) Somatic replication of DNA methylation  
Wigler M., Levy D., Perucho M.  
Cell **1981**, 24, 33
- 56) DNA methylation decreases in aging but not in immortal cells  
Wilson V.L, Jones P.A  
Science **1983**, 220, 1055
- 57) The inheritance of epigenetic effects  
Holliday R.  
Science **1987**, 238, 163
- 58) Methylation of the mouse Hprt gene differs on the active and inactive X chromosomes  
Lock L., Melton D., Thomas Caskey C., Martin G.  
Mol. Cell. Biol. **1988**, 6, 914
- 59) Synthesis and biological activity of analogues of adenosylhomocysteine as inhibitors of methyltransferases  
Borchardt R.  
Biochemistry of S-Adosylmethionine Ed.E.Borek *et al.* Columbia University press  
New york **1977**
- 60) SAM dependant macromolecule methyltransferases. Potential targets for the design of chemotherapeutic agents.  
Borchardt R.  
J. Med. Chem. **1980**, 23, 347
- 61) Structure activity relationship of synthetic SAH analogues effect on oncogenic virus-induced cell transformation and on methylation of macromolecules  
Legaverand M., Lederer E., Ibanez S., Enauf J., Lawrence F., Robert-gero M.  
Eur. J. Med. Chem. **1977**, 12, 105

- 62) Potential inhibitors of SAM dependant MTASEs. 2. Modification of the base portion of SAH  
Borchardt R., Huber J., Yu Y.  
J. Med. Chem. **1974**, 17, 869
- 63) Type II restriction modification systems  
Wilson G.G  
Trends in Genetics **1988**, 4, 314
- 64) SAM : a new intermediate formed enzymatically from L-methionine and adenosine triphosphate  
Cantoni G.L.  
J. Biol. Chem. **1953**, 204, 403
- 65) Transmethylation  
Olsen J., Shiong Y., Borchardt R., Schowen R. Ed. Usdin, Borchardt, Creveling  
Elsevier **1979**
- 66) SAM: the relation of configuration at the sulphonium centre to enzymic reactivity  
De la Haba G., Jamieson G.A, Mudd S.H, Richards H.H.  
J. Amer. Chem. Soc. **1959**, 81, 3975
- 67) Transmethylation  
Carnforth J., Carrell H., Glusker J., Tatalay P.  
ed. Usdin, Borchardt, Creveling. Elsevier **1979**
- 68) Biochemistry of S-Adenosylmethionine  
Coward J.K.  
Ed. E. Borek Columbia University press
- 69) Cloning and sequence of a cDNA encoding DNA methylase of Mouse cells  
Bestor T., Laudano A., Mattaliano R., Ingram V.  
J. Mol. Biol. **1988**, 203, 971
- 70) Conformational studies of S-adenosyl-l-homocysteine, a potential inhibitor of S-adenosylmethionine-dependant methyltransferases  
Ishida T., Tanaka A., Inoue M., Fujiwara T., Tomita K  
J. Amer. Chem. Soc. **1982**, 104, 7239
- 71) Chem-X  
Developed and distributed by Chemical Design Ltd, Oxford, England.
- 72) MOPAC 5.0  
QCPE Bulletin 9,(4),116, **1989**. Quantum Chemistry Program Exchange, Dept. of Chemistry, Indiana University, Bloomington, IN., USA.

73) QUANTA

Copyright Polygen corporation, Waltham, MA, USA

74) Peptide conformation

Ramachandran G.N

Peptides, Polypeptides, Protein Proc. Rehovot Symposium 1974 Ed. Blout, Elkan Rogers, Bovey, Goodman. Wiley N.Y.

75) Crystal structure of S-adenosylhomocysteine

Shieh H-S., Berman H.M.

Acta Cryst. B 1982, 38, 1513

76) Crystal structure of S-methylmethionine

Del Re G., Gavuzzo E., Giglio E., Lelj F., Mazza F., Zappia V.

Acta Cryst. B 1977, 33, 3289

77) Prediction of protein conformation. Consequences and aspirations

Fasman G.D., Chou P.Y.

Peptides, polypeptides and proteins, Proc. Rehovot symposium 1974

Ed. Blout, Elkan, Rogers, Bovey, Goodman Wiley. New York

78) Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins.

Garnier, Osguthorpe, Robson

J. Mol. Biol. 1978, 120 (i), 97

79) Structure of the DNA-EcorRI endonuclease recognition complex at 3A resolution

McClarín J.A., Frederick C., Wang B., Greene P., Boyer M., Rosenberg J., Grable J.

Science 1986, 234, 1526.

80) Chemistry of the Sulphonium Group part 1

Lowe P.A. 1981, 207 Ed. S.Patai. Wiley and Sons

81) Methionine studies VII. Sulphonium derivatives

Toennies G., Kolb J.

J. Amer. Chem. Soc. 1945, 67, 849

82) Sulphonium derivatives of D-L methionine and 5-(2-methylthioethyl)hydantoin

Atkinson R.O., Popplesdorf F.

J. Chem. Soc. 1951, 1378.

83) A direct synthesis of sulphonium perchlorates

Milligan T.W., Minor B.C.

J. Org. Chem. 1963, 28, 235

- 84) The practice of peptide synthesis  
Bodanszky M., Bodanszky A.  
Springer-Verlag **1984**
- 85) Construction and Characterisation of extrachromosomal probes for mutagenesis by carcinogens; site specific incorporation of O<sup>6</sup>-methylguanine into viral and plasmid genes.  
Green C.L., Loechler E.L., Fowler K.W., Essigmann J.M.  
Proc. Natl. Acad. Sci. USA **1984**, 81, 13
- 86) Synthesis of a nonapeptide with properties of bradykinine  
Boissonnas R.A., Guttmann S., Jaquenod P.A.  
Helv. Chim. Acta **1960**, 43, 1349
- 87) A convenient synthesis of water soluble carbodiimides  
Sheehan J.G, Cruickshank P., Borshart G.  
J. Org. Chem. **1961**, 26, 2525
- 88) Studies on the mechanisms of detoxification of Histamine by S-methylmethionine sulphonium chloride  
Suzue R., Endo N., Koga S.  
Bull. Inst. Chem. Res. Kyoto University **1967**, 201
- 89) Alkyltransferase model reactions : synthesis of sulphonium and ammonium compound containing neighbouring nucleophiles. Kinetic studies of the Intramolecular reaction of amino, hydroxy, phenoxy and mercapto onium salts.  
Miller R.J., Kuliopulos A., Coward J.K.  
J. Org. Chem. **1963**, 28, 235
- 90) Mechanisms of reaction of thioisulphinates (sulphenic anhydrides). The thioisulphinate-sulphinic acid reaction.  
Kice J.L., Venier C.G, Heasley L.  
J. Amer. Chem. Soc. **1967**, 89, 3557.
- 91) In vitro DNA cytosine methylation of cis-regulatory elements modulates c-Ha-ras promoter activity in vivo  
Rachal M.J, Yoo H., Beckker F.F., Lapeyre J.  
Nucleic acids Res. **1989**, 17, 5135
- 92) Atomic structure of Thymidylate synthetase: Target for rational drug design  
Hardy L., Finer-Moore J., Montfort W., Jones M., Santi D., Stroud M.  
Science **1987**, 235, 448



- 93) Active site and complete sequence of the suicidal methyltransferase that counters alkylation mutagenesis  
Demple B., Sedgwick, Robins P., Totty N., Waterfield M.D., Lindahl T.  
Proc. Natl. Acad. Sci. USA **1985**, 82, 2688
- 94) Structure of papain  
Denth J., Jansonius J., Koekoek R., Swan H., Wolthes B.G.  
Nature **1968**, 218, 929
- 95) Studies on peptides LXXIII : Examination of the methanesulphonic acid procedure for the synthesis of peptides containing methionine  
Fujino M., Shingawa S.  
Chem. Pharm. Bull. **1977**, 25, 2929
- 96) Über die aldehyd-verbindungen der aminosäuren  
Bergmann M., Zervas L.  
Ber. **1932**, 58, 1034
- 97) Synthese und Verwedung von tert. butyloxycarbonylfluorid und anderen Fluorkohlensäureestern zur dorstellung  
Schnabel E., Herzog H., Hoffmann P., Klauke E., Ugi J.  
Liebigs Ann. Chem. **1968**, 716, 175
- 98) A new method of forming peptide bonds  
Sheehan J., Hess G.  
J. Amer. Chem. Soc. **1955**, 77, 1067
- 99) Side chain protected N $\alpha$ -fluorenylmethoxycarbonylamino acids for solid phase peptide synthesis. N $\alpha$ -fluorenylmethoxycarbonyl N<sup>G</sup>-4 methoxy 2,3,6 trimethylbenzenesulphonyl arginine  
Atherton E., Sheppard R.C., Wade J.D.  
J. Chem. Soc. Chem. Commun. **1983**, 1860.
- 100) Kinetic and catalytic mechanism of HhaI methyltransferase  
Wu J., Santi D.  
J. Biol. Chem. **1987**, 262, 4778
- 101) p-Alkoxybenzylalcohol resin and p-alkoxybenzylcarbonylhydrazide resin for solid phase synthesis of protected peptide fragments  
Wang S.  
J. Amer. Chem. Soc. **1973**, 95, 1328
- 102) A convenient preparation of S-benzhydryl and S-trityl-l-cysteine  
Hiskey R.G., Adams J.  
J. Org. Chem. **1965**, 30, 1340

- 103) Zur darstellung von hydroxyaminosaure-tert butylathern  
Wunsch E., Jentsch J.  
Chem. Ber. **1964**, 97, 2490
- 104) N<sup>G</sup>-2,2,5,6,7,8-pentamethylchroman-6-sulphonyl-arginine. A new acid labile derivative for peptide synthesis  
Ramage R., Green J.  
Tet. Lett. **1987**, 2287
- 105) Non-enzymatic methylation of DNA by SAM in vitro.  
Barrows L.R., Magee P.N  
Carcinogenesis **1982**, 3, 349
- 106) Non- enzymatic methylation of DNA by the intracellular methyl group donor SAM is a potentially mutagenic reaction.  
Rydberg B., Lindahl T.  
EMBO J. **1982**, 1, 211
- 107) Quantitative analysis of SAM and SAH in animal tissues.  
Salvatore F., Utili R., Zappia V., Shapiro S.K.  
Anal. Biochem. **1971**, 41, 16
- 108) Alkyltransferase model reactions : Synthesis of sulphonium and ammonium compounds containing neighbouring nucleophiles. Kinetic studies of the intramolecular reaction of amino, hydroxy, phenoxy and mercapto onium salts.  
Miller R.J, Kuliopulos A., Coward J.K.  
J. Org. Chem. **1989**, 54, 3436
- 109) Procède de production de sels de methionine-methylsulfonium  
French patent No.903.994 **1965** to Nippon Kayaku Kabushiki Kaisha
- 110) Acylalkylcarbonates as acylating agents for the synthesis of peptides  
Vaughn J.R  
J. Amer. Chem. Soc. **1951**, 73, 3547

## APPENDIX 1

### Crystal structure of N-CBZ-prolinyl-S-(Bzl)cysteine methyl ester (5)

N-CBZ-prolinyl-S-(Bzl)cysteine methyl ester (5) crystallised as colourless plates from a saturated solution of the compound in ethyl acetate allowed to stand in an atmosphere saturated with petroleum ether [60-80<sup>o</sup>] for 1 week. The crystal had dimensions 0.90 x 0.55 x 0.25mm. The data were collected on an Enraf-Nonius CAD4 diffractometer with monochromated Mo-K $\alpha$  radiation  $\lambda = 0.71069\text{\AA}$ .

#### Crystal data

C<sub>24</sub>H<sub>28</sub>N<sub>2</sub>O<sub>5</sub>S, M=456.5, monoclinic, a=5.022(2), b=10.157 (4), c=22.824 (8) Å,  $\beta=93.23$  (3), V=1162 (3) Å<sup>3</sup>, Z=2, D<sub>m</sub>=1.29 (2)g cm<sup>-3</sup>, D<sub>x</sub>= 1.30g cm<sup>-3</sup>, F(000)=968,  $\mu(\text{Mo-k}\alpha)=2.71$  cm<sup>-1</sup>, space group P2<sub>1</sub>.

The 4651 intensity data were collected by the  $\omega$ -2 $\theta$  scan technique, and of these 3240 reflections were deemed observed with  $F_o > 1.5\sigma$ . Using SHELX-76, the non-hydrogen atoms were located from an E map. Hydrogen atoms were included in calculated positions and all atoms refined isotropically. Further refinement with anisotropic non-hydrogen atoms reduced the unweighted discrepancy index to R=0.053. In the latter stages of refinement, reflections were weighted according to  $\omega=K/[\sigma^2(F)+0.003F^2]$ . Finally a difference electron density map was calculated which showed no feature greater than 0.60 Å<sup>-3</sup>. Calculations were performed on the VAX 8650 cluster at Aston University.

Observed and calculated structure factors and atomic thermal parameters are given in the appended tables.

Structural data was collected and determined by Dr. P.C.Yates.

Table 1. Final atomic coordinates ( $\times 10^{**4}$ ) for X with estimated standard deviations (e.s.d.s) in parentheses.

Atom	x/a	y/b	z/c
C(1)	6213(9)	1706(5)	6873(2)
C(2)	4268(9)	2076(5)	6459(2)
C(3)	2606(9)	1156(5)	6185(2)
C(4)	2822(9)	-0140(5)	6340(2)
C(5)	4787(9)	-0535(4)	6765(2)
C(6)	6462(7)	0385(4)	7038(2)
C(7)	8538(9)	-0044(6)	7502(2)
S(8)	7991(2)	0625(0)	8224(0)
C(9)	4855(6)	-0126(4)	8398(2)
C(10)	5142(6)	-1124(3)	8903(2)
N(11)	6045(5)	-2400(3)	8711(1)
C(12)	4298(7)	-3314(4)	8523(2)
C(13)	5486(7)	-4590(4)	8308(2)
N(14)	3651(6)	-5216(3)	7866(1)
C(15)	2817(8)	-4510(4)	7391(2)
O(16)	0799(6)	-5119(3)	7080(1)
C(17)	-0492(10)	-4358(7)	6600(2)
C(18)	0682(8)	-4657(4)	6025(2)
C(19)	2691(10)	-3867(5)	5808(2)
C(20)	3699(10)	-4130(7)	5272(2)
C(21)	2716(11)	-5171(7)	4943(2)
C(22)	0785(11)	-5979(6)	5157(2)
C(23)	-0259(10)	-5699(5)	5688(2)
C(24)	8064(13)	1347(5)	9934(2)
O(25)	6310(6)	0624(3)	9528(1)
C(26)	6976(7)	-0600(4)	9396(2)
C(27)	5692(8)	-5636(4)	8795(2)
C(28)	3015(8)	-6351(4)	8739(2)
C(29)	2373(8)	-6402(4)	8079(2)
O(30)	8833(8)	-1169(3)	9634(1)

O(32)	3747(8)	-3461(3)	7261(2)
H(10)	7547	2434	7071
H(20)	4039	3104	6344
H(30)	1132	1460	5849
H(40)	1483	-0858	6136
H(50)	4992	-1563	6880
H(71)	8514	-1106	7529
H(72)	10468	0280	7374
H(91)	4031	-0627	8012
H(92)	3506	0642	8520
H(101)	3178	-1261	9063
H(111)	8154	-2613	8721
H(130)	7396	-4322	8148
H(171)	-2590	-4599	6567
H(172)	-0249	-3322	6695
H(190)	3447	-3041	6065
H(200)	5262	-3520	5112
H(210)	3449	-5360	4516
H(220)	0094	-6829	4908
H(230)	-1838	-6307	5842
H(241)	7268	2307	10028
H(242)	9917	1457	9721
H(243)	8418	0807	10338
H(271)	5987	-5179	9221
H(272)	7313	-6308	8728
H(281)	1505	-5808	8955
H(282)	3178	-7330	8922
H(291)	0247	-6378	7979
H(292)	3204	-7275	7890

Table 2. Interatomic distances (A) for X with e.s.d.s in parentheses.

Bond	Interatomic distance (A)
C(1)-C(2)	1.374(6)

C(2)-C(3)	1.379(7)
C(3)-C(4)	1.366(7)
C(4)-C(5)	1.402(7)
C(5)-C(6)	1.382(6)
C(6)-C(7)	1.509(6)
C(7)-S(8)	1.817(5)
S(8)-C(9)	1.814(3)
C(9)-C(10)	1.535(5)
C(10)-N(11)	1.450(4)
C(10)-C(26)	1.511(5)
N(11)-C(12)	1.331(4)
C(12)-C(13)	1.521(5)
C(12)-O(31)	1.243(4)
C(13)-N(14)	1.472(5)
C(13)-C(27)	1.536(6)
N(14)-C(15)	1.346(5)
N(14)-C(29)	1.461(5)
C(15)-O(16)	1.354(5)
C(15)-O(32)	1.208(5)
O(16)-C(17)	1.462(6)
C(17)-C(18)	1.500(6)
C(18)-C(19)	1.402(7)
C(18)-C(23)	1.377(7)
C(19)-C(20)	1.376(7)
C(20)-C(21)	1.373(8)
C(21)-C(22)	1.381(8)
C(22)-C(23)	1.375(8)
C(24)-O(25)	1.443(5)
O(25)-C(26)	1.326(5)
C(26)-O(30)	1.200(5)
C(27)-C(28)	1.527(6)
C(28)-C(29)	1.523(6)

Table 3. Interatomic angles (degrees) for X.

Atoms	Bond angle (degrees)
C(2)-C(1)-C(6)	119.8(4)
C(1)-C(2)-C(3)	121.0(4)
C(2)-C(3)-C(4)	119.9(4)
C(3)-C(4)-C(5)	119.9(4)
C(4)-C(5)-C(6)	120.3(4)
C(1)-C(6)-C(5)	119.0(4)
C(1)-C(6)-C(7)	121.0(4)
C(5)-C(6)-C(7)	120.0(4)
C(6)-C(7)-S(8)	113.3(3)
C(7)-S(8)-C(9)	102.6(2)
S(8)-C(9)-C(10)	113.3(2)
C(9)-C(10)-C(26)	110.8(3)
C(9)-C(10)-C(11)	112.4(3)
C(11)-C(10)-C(26)	110.6(3)
C(10)-N(11)-C(12)	120.6(3)
N(11)-C(12)-C(13)	115.8(3)
N(11)-C(12)-O(31)	123.2(3)
C(13)-C(12)-O(31)	121.0(3)
C(12)-C(13)-N(14)	110.3(3)
C(12)-C(13)-C(27)	111.6(3)
N(14)-C(13)-C(27)	102.2(3)
C(13)-N(14)-C(15)	118.4(3)
C(13)-N(14)-C(29)	113.4(3)
C(15)-N(14)-C(29)	125.8(3)
N(14)-C(15)-O(16)	111.6(3)
N(14)-C(15)-O(32)	124.1(4)
O(16)-C(15)-O(32)	124.2(4)
C(15)-O(16)-C(17)	116.0(4)
O(16)-C(17)-C(18)	111.7(4)
C(17)-C(18)-C(19)	121.3(5)



C(19)-C(18)-C(23)	118.3(4)
C(18)-C(19)-C(20)	120.8(5)
C(19)-C(20)-C(21)	119.8(5)
C(20)-C(21)-C(22)	120.2(5)
C(21)-C(22)-C(23)	119.9(5)
C(22)-C(23)-C(18)	121.0(5)
C(24)-O(25)-C(26)	118.0(3)
C(10)-C(26)-O(25)	110.3(3)
C(10)-C(26)-O(30)	126.3(3)
O(25)-C(26)-O(30)	123.3(3)
C(13)-C(27)-C(28)	104.2(3)
C(27)-C(28)-C(29)	103.6(3)
C(28)-C(29)-N(14)	103.2(3)

Table 4. Torsion angles (degrees) for bonds in X involving non-hydrogen atoms. Estimated standard deviations are one degree.

Atoms	Torsion angle (degrees)
C(1)-C(6)-C(7)-S(8)	65
C(5)-C(6)-C(7)-S(8)	-116
C(6)-C(7)-S(8)-C(9)	64
C(7)-S(8)-C(9)-C(10)	112
S(8)-C(9)-C(10)-N(11)	-81
S(8)-C(9)-C(10)-C(26)	43
C(9)-C(10)-N(11)-C(12)	-89
C(26)-C(10)-N(11)-C(12)	146
C(9)-C(10)-C(26)-O(25)	48
C(9)-C(10)-C(26)-O(30)	-130
N(11)-C(10)-C(26)-O(25)	173
N(11)-C(10)-C(26)-O(30)	-5
C(10)-N(11)-C(12)-C(13)	177
C(10)-N(11)-C(12)-O(31)	-5
N(11)-C(12)-C(13)-N(14)	-152
N(11)-C(12)-C(13)-C(27)	95

O(31)-C(12)-C(13)-C(27)	-82
C(12)-C(13)-N(14)-C(15)	55
C(12)-C(13)-N(14)-C(29)	-108
C(27)-C(13)-N(14)-C(15)	174
C(27)-C(13)-N(14)-C(29)	10
C(12)-C(13)-C(27)-C(28)	89
N(14)-C(13)-C(27)-C(28)	-29
C(13)-N(14)-C(15)-O(16)	-169
C(13)-N(14)-C(15)-O(16)	-8
C(13)-N(14)-C(15)-O(32)	10
C(29)-N(14)-C(15)-O(32)	172
C(13)-N(14)-C(29)-C(28)	12
C(15)-N(14)-C(29)-C(28)	-149
N(14)-C(15)-O(16)-C(17)	172
O(32)-C(15)-O(16)-C(17)	-7
C(15)-O(16)-C(17)-C(18)	95
O(16)-C(17)-C(18)-C(19)	-96
O(16)-C(17)-C(18)-C(23)	86
C(24)-O(25)-C(26)-C(10)	-172
C(24)-O(25)-C(26)-O(30)	6
C(13)-C(27)-C(28)-C(29)	37
C(27)-C(28)-C(29)-N(14)	-30

#### SUPPLEMENTARY PUBLICATION

##### Atomic Thermal Parameters

Anisotropic values are in the form

$$\exp\{-2 \pi^2 (U_{11} a^2 h^2 + \dots + 2U_{23} b^* c^* k l + \dots)\}$$

Atom	U11	U22	U33	U23	U13	U12
C(1)	54	63	68	9	1	-5
C(2)	61	59	75	19	2	6
C(3)	56	81	56	5	226	5

C(5)	67	57	55	4	13	4
C(6)	42	68	45	3	10	3
C(7)	50	96	63	17	3	23
S(8)	37	75	55	11	-3	-14
C(9)	24	47	57	-2	-1	0
C(10)	27	36	56	-8	7	-2
N(11)	28	40	65	-7	4	-1
C(12)	32	38	52	-2	6	-4
C(13)	32	39	68	-7	7	-3
N(14)	46	36	53	-2	7	-2
C(15)	59	45	54	-3	17	0
O(16)	70	71	56	-2	2	-1
C(17)	70	104	64	3	1	32
C(18)	48	70	54	7	-3	17
C(19)	65	67	75	6	-8	0
C(20)	61	108	75	15	6	3
C(21)	70	121	56	10	9	16
C(22)	81	86	68	-6	-12	4
C(23)	62	76	66	8	-5	-1
C(24)	104	69	71	-22	-11	-24
O(25)	66	49	63	-17	-4	-1
C(26)	42	41	48	1	4	-4
C(27)	39	49	79	-3	-4	7
C(28)	53	49	67	3	7	-8
C(29)	54	38	61	-3	11	-7
O(30)	90	64	74	-1	-35	13
O(31)	24	50	84	-18	4	-2
O(32)	94	52	84	14	19	-7