

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)

THE DESIGN OF NOVEL ANTIFOLATES
AGAINST *PNEUMOCYSTIS CARINII*

KEITH SIMPSON PHILIP
Doctor of Philosophy

THE UNIVERSITY OF ASTON IN BIRMINGHAM
January 1994

This copy of the thesis has been supplied on the 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 proper acknowledgement

The University of Aston in Birmingham
The Design of Novel Antifolates Against *Pneumocystis carinii*

by

Keith Simpson Philip

submitted for the degree of Doctor of Philosophy, January 1994

The pneumonia caused by *Pneumocystis carinii* is ultimately responsible for the death of many acquired immunodeficiency syndrome (AIDS) patients. Large doses of trimethoprim and pyrimethamine in combination with a sulphonamide and/or pentamidine suppress the infection but produce serious side-effects and seldom prevent recurrence after treatment withdrawal. However, the partial success of the aforementioned antifolates, and also trimetrexate used alone, does suggest dihydrofolate reductase (DHFR) as a target for the development of antipneumocystis agents.

From the DHFR inhibitory activities of 3'-substituted pyrimethamine analogues it was suggested that the 3'-(3'',3''-dimethyltriazene-1''-yl) substituent may be responsible for the greater activity for the *P. carinii* over the mammalian enzyme. Crystallographic and molecular modeling studies revealed considerable geometrical and electronic differences between the triazene and the chemically related formamidine functions that may account for the differences in DHFR inhibitory profiles.

Structural and electronic parameters calculated for a series of 3'-(3'',3''-disubstitutedtriazene-1''-yl) pyrimethamine analogues did not correlate with the DHFR inhibitory activities. However, the *in vitro* screening against *P. carinii* DHFR revealed that the 3''-hydroxyethyl-3''-benzyl analogue was the most active and selective.

Models of the active sites of human and *P. carinii* DHFRs were constructed using DHFR sequence and structural homology data which had identified key residues involved in substrate and cofactor binding. Low energy conformations of the 3'',3''-dimethyl and 3''-hydroxyethyl-3''-benzyl analogues, determined from nuclear magnetic resonance studies and theoretical calculations, were docked by superimposing the diaminopyrimidine fragment onto a previously docked pyrimethamine analogue. Enzyme kinetic data supported the 3''-hydroxyethyl-3''-benzyl moiety being located in the NADPH binding groove. The 3''-benzyl substituent was able to locate to within 3 Å of a valine residue in the active site of *P. carinii* DHFR thereby producing a hydrophobic contact. The equivalent residue in human DHFR is threonine, more hydrophilic and less likely to be involved in such a contact. This difference may account for the greater inhibitory activity this analogue has for *P. carinii* DHFR and provide a basis for future drug design.

From an *in vivo* model of PCP in immunosuppressed rats it was established that the 3''-hydroxyethyl-3''-benzyl analogue was able to reduce the *P. carinii* burden more effectively with increasing doses, without causing any visible signs of toxicity. However, equivalent doses were not as effective as pentamidine, a current treatment of choice for *Pneumocystis carinii* pneumonia.

Keywords: *Pneumocystis carinii*, Dihydrofolate Reductase, Molecular Modeling, Computer-Aided Drug Design, Pyrimethamine Analogue.

Acknowledgements

I would like to express my gratitude to my supervisors Dr. Carl Schwalbe and Dr. Peter Lambert for their continued enthusiasm, guidance and friendship they have displayed throughout my doctoral research. Sincerest thanks are also due to Professor Malcolm Stevens and Dr. John Slack of Aston Molecules Ltd. for their interest and positive contributions into this project.

I am especially grateful to Dr. Brian Denny, Dr. Sandy Geddes (University of Leeds), Dr. Clare Sansom (University of Leeds), Dr. Richard Rohwer and Daniel Morris for their computational expertise and invaluable assistance in the molecular modeling studies; to Dr. Sally Freeman, Dr. Bill Fraser and Ken Ross for their assistance and expertise in the nuclear magnetic resonance spectra; and to Mel Gamble for his technical assistance and patience in establishing the immunocompromised rat lung *Pneumocystis carinii* Pneumonia model. Special thanks go to Roy Tilling and Dorothy Townley for their generous assistance and friendship they have shown to me.

I would like to express my thanks to my family, especially my sister Katrina, friends and colleagues for their encouragement and support and who have made my time at Aston University so enjoyable.

Finally, I would like to acknowledge the Science and Engineering Research Council and Aston Molecules Ltd. for the financial support of this research project

Contents

Summary	2
Acknowledgements	3
Contents	4
List of Tables and Figures	10
Abbreviations	16
Chapter 1 Introduction	
1.1 A Brief Historical Account of <i>Pneumocystis carinii</i>	18
1.2 Taxonomy of <i>Pneumocystis carinii</i>	19
1.2.1 Original Taxonomy of <i>P. carinii</i>	19
1.2.2 Genetic Evidence for the Fungal Classification of <i>P. carinii</i>	19
1.2.3 Biochemical Evidence for the Fungal Classification of <i>P. carinii</i>	20
1.2.3.1 Dihydrofolate Reductase and Thymidine Synthase	20
1.2.3.2 Elongation Factor 3	20
1.2.3.3 Antigenicity of <i>P. carinii</i>	20
1.2.4 Conclusions to the Taxonomic Classification of <i>P. carinii</i>	21
1.3 The Life-Cycle of <i>Pneumocystis carinii</i>	21
1.4 Epidemiology and Pathogenicity of <i>Pneumocystis carinii</i> Pneumonia	22
1.4.1 Epidemiology	22
1.4.2 Pathogenicity	23
1.5 Antifolate Treatment and Prophylaxis of <i>Pneumocystis carinii</i> Pneumonia	24
1.5.1 Introduction	24
1.5.2 Folate Biochemistry	26
1.5.3 Antifolates used in the Treatment of <i>P. carinii</i> Pneumonia	29
1.5.3.1 Trimethoprim	29
1.5.3.2 Pyrimethamine	30
1.5.3.3 Trimetrexate	31
1.5.3.4 Piritrexim	33
1.5.3.5 Proguanil and a novel Bi-guanide	33
1.5.3.6 Dapsone	34
1.5.3.7 Conclusions for the Use of Antifolate Chemotherapy in the Treatment and Prophylaxis of <i>P. carinii</i> Pneumonia	35

1.6	Computer-Aided Drug Design	36
1.6.1	Introduction to Computer-Aided Drug Design	36
1.6.2	CADD Using a Known Receptor Structure	37
1.6.3	CADD Using a Predicted Receptor Structure	38
1.6.4	CADD Using an Unknown Receptor Structure	39
1.6.5	Dihydrofolate Reductase as a Target for CADD	39
1.6.6	Selection of a Suitable Lead Compound for the Design of Inhibitors Against <i>P. carinii</i> DHFR	40
1.7	Rationale and Objectives	44
Chapter 2 Crystallographic Studies of 2,4-Diamino-6-ethyl-5-(4'-chloro-3'-(3'',3''-dimethyltriazene-1-yl)phenyl)pyrimidine Ethanesulphonate		
2.1	Introduction	46
2.2	Materials and Methods	47
2.2.1	Sources of Chemicals	47
2.2.2	Preparation of AM90021 Crystals from the Free Base AM90020	48
2.2.3	Experimental Determination of AM90021 Crystal Density	48
2.2.4	Crystallographic Data	48
2.2.5	Data Collection and Structural Analysis of AM90021	48
2.2.6	Determination of the Planarity Between the Pyrimidine Aromatic Rings in the Unit Cell of AM90021	49
2.3	Results and Discussion	53
Chapter 3 Computational Chemical Calculations on 3,3-Dimethyl-1-(2'-chloro)phenyltriazene		
3.1	Introduction	65
3.1.1	<i>Ab Initio</i> Quantum Mechanics Programs	65
3.1.2	Semi-empirical Quantum Mechanics Programs	66
3.1.3	Molecular Mechanics Programs	66
3.2	Materials and Methods	67
3.2.1	Molecular Modeling Packages	67
3.2.2	Computational Chemistry Calculations on 3,3-Dimethyl- 1-(2'-chloro)phenyltriazene and 3,3-Dimethyl-1- (2'-chloro)phenylformamidine	68

3.2.3	Superimposing the 3,3-Dimethyl-1-(2'-chloro)phenyltriazene Structures	69
3.2.4	Determination of Molecular Similarity	69
3.3	Results and Discussion	70
Chapter 4 Biochemical Studies on Trimethoprim, Pyrimethamine and 3'-(3'',3''-disubstitutedtriazene-1''-yl) Pyrimethamine Analogues		
4.1	Introduction	77
4.2	Materials and Methods	77
4.2.1	Sources of Reagents and Enzymes Used in the Assays	77
4.2.2	Preparation of Partially Purified DHFR from <i>E. coli</i>	78
4.2.3	Preparation of Partially Purified DHFR from Rat Liver	78
4.2.4	Preparation of Reagents	78
4.2.4.1	Dihydrofolate Solution	78
4.2.4.2	Inhibitor Solution	78
4.2.5	Assay for IC ₅₀ Inhibitory Activity Against <i>Escherichia coli</i> and Rat Liver DHFR	79
4.2.6	Enzyme Kinetic Studies of AM90067 Against DHF in Rat Liver DHFR	79
4.2.7	Enzyme Kinetic Studies of AM90067 Against NADPH in Rat Liver DHFR	80
4.3	Results and Discussion	80
Chapter 5 Determination of Quantitative Structure Activity Relationships Between the 3'-(3'',3''-disubstitutedtriazene-1''-yl) Pyrimethamine Analogues		
5.1	Introduction	86
5.1.1	Molecular Similarity in Quantitative Structure Activity Relationships	86
5.1.2	Neural Networks in Quantitative Structure Activity Relationships	87
5.2	Materials and Methods	89

5.2.1	Molecular Modeling Packages	89
5.2.2	Model Compound Construction	89
5.2.3	Determination of Molecular Similarity Indices	89
5.2.4	Determination of the Molecular Volume of the Terminal Triazene Substituents	89
5.2.5	Neural Networks Procedures	90
5.3	Results and Discussion	91
Chapter 6 Conformational Analysis of 2,4-Diamino-6-ethyl-5-(4'-chloro-3'-(3'',3''-dimethyl)triazene-1''-yl)phenylpyrimidine (AM90020) and 2,4-Diamino-6-ethyl-5-(4'-chloro-3'-(3''-hydroxyethyl-3''-benzyl)triazene-1''-yl)phenylpyrimidine (AM90067)		
6.1	Introduction	96
6.2	Materials and Methods	98
6.2.1	Sources of Chemicals	99
6.2.2	Computational Software	99
6.2.3	MOPAC Calculations	99
6.2.4	MOPAC Optimisations	99
6.2.5	Construction of Protonated Model Compounds	100
6.2.6	Conformation Generation and Energy Calculation	100
6.2.6.1	Rotation of the Torsions C4P-C5P-C1P-C2P and C5D-C6D-C61D-C62D	100
6.2.6.2	Rotation of the Torsion C2P-C3P-N1-N2	100
6.2.6.3	Rotation of the Torsions N2-N3-C7-C1H and N2-N3-C8-C1B	101
6.2.6.4	Rotation of the Torsion N3-C7-C1H-O1H	101
6.2.6.5	Rotation of the Torsion C7-C1H-O1H-H10H	101
6.2.6.6	Rotation of the Torsion N3-C8-C1B-C2B	101
6.2.7	Low Energy Optimisations of Low Energy Conformers of AM90020 and AM90067	101
6.2.8	Superimposing the Conformers of AM90020 and AM90067	101
6.2.9	NMR Conformational Studies on AM90067	101
6.3	Results and Discussion	103

Chapter 7 Computer-Aided Active-Site Modeling of AM90067 in Human Liver and *Pneumocystis carinii* Dihydrofolate Reductase

7.1	Introduction	127
7.1.1	Sequence Homology Alignment	128
7.1.2	Modeling Regions with Low Structural Homology	129
7.1.3	Prediction of Side-Chain Conformation	130
7.1.4	Docking and Calculation of Interaction	130
7.2	Materials and Methods	132
7.2.1	Computational Software	132
7.2.2	DHFR Sequence Homology Studies	132
7.2.3	Construction of a Human Liver DHFR Active-Site Model	132
7.2.3.1	Secondary Structure Assignment	132
7.2.3.2	Alignment and Superimposition of Protein Structures	133
7.2.3.3	Preliminary Model Creation	133
7.2.3.4	<i>meta</i> -Azidopyrimethamine and Water Inclusion	133
7.2.4	Construction of a <i>P. carinii</i> DHFR Active-Site Model	134
7.2.4.1	Creation of a <i>P. carinii</i> DHFR Amino Acid Sequence File	134
7.2.4.2	Creation of a Molecular Structure File of <i>P. carinii</i> DHFR from a Sequence File	134
7.2.4.3	Assignment and Alignment of the Secondary Structures of the Crystalline <i>Escherichia coli</i> , <i>Lactobacillus casei</i> , Mouse Liver and Chicken Liver DHFR Structures with the Model Human DHFR	134
7.2.4.4	Alignment of <i>P. carinii</i> DHFR Amino Acid Sequence to the Five Homologous Sequences	134
7.2.4.5	Construction of the Peptide Backbone of the Active Site Model of <i>P. carinii</i> DHFR	135
7.2.4.6	Side Chain Conformation for <i>P. carinii</i> DHFR Active Site Residues With One or More Identical Matches With Known DHFR Structures	135
7.2.4.7	Side Chain Conformation of Residues Unique to <i>P. carinii</i> Dihydrofolate Reductase	135

7.2.4.8	Exceptions to 7.2.4.6 and 7.2.4.7	136
7.2.4.8.1	Serine 24	136
7.2.4.8.2	Arginine 56 and Lysine 57	136
7.2.4.8.3	Valine 154	136
7.2.4.9	<i>meta</i> -Azidopyrimethamine and Water Inclusion	136
7.2.5	Docking of AM90067 into the <i>P. carinii</i> and Human DHFR Active Site Models	136
7.2.6	Removal of Distance Residues	137
7.2.7	Calculation of the Energy of Interaction	137
7.3	Results and Discussion	138
Chapter 8 Treatment of Experimental <i>Pneumocystis carinii</i> Pneumonia using AM90067		
8.1	Introduction	156
8.2	Materials and Methods	157
8.2.1	Sources of Chemicals	157
8.2.2	Animal Protocol	157
8.2.3	Quantitative Determination of <i>P. carinii</i> Cysts in Rat Lung	158
8.2.4	Direct Immunofluorescence for <i>P. carinii</i>	158
8.3	Results and Discussion	159
Chapter 9 Concluding Remarks		165
References		169
Appendix Structure Factor Data for AM90021		189

List of Figures

1.1	The life-cycle of <i>Pneumocystis carinii</i>	22
1.2	Folate biosynthesis and the folate cycle	27
1.3a	The structure of folic acid	28
1.3b	The structure of 7,8-dihydrofolic acid	28
1.3c	The structure of 5,6,7,8-tetrahydrofolic acid	28
1.3d	The structure of N^5,N^{10} -methylene-5,6,7,8-tetrahydrofolic acid	28
1.4	The structures of trimethoprim (a) and sulphamethoxazole (b)	29
1.5	The structures of pyrimethamine (a) and sulphadoxine (b)	31
1.6a	The structure of trimetrexate	32
1.6b	The structure of leucovorin	32
1.7	The structure of piritrexim	33
1.8	The <i>in vivo</i> conversions of bi-guanides to their corresponding triazines	34
1.9	The structure of dapsone	35
1.10	Free energies (ΔG) involved in computer-aided drug design	37
1.11	Representation of the potential for the design of inhibitors against <i>P. carinii</i> DHFR based on pyrimethamine	41
1.12	Structures of pyrimethamine analogues with 3'-nitrogen substituents`	42
2.1	The structures of DAMP and DNMP	47
2.2	The structure of methylbenzoprim	47
2.3	The resolved crystal structure and atomic numbering of AM90021 with selected symmetry transformed atoms	51
2.4	PLUTO drawing of two molecules of AM90021 in the unit cell	52
2.5	ORTEP drawing of a molecule of AM90021	52
2.6	Resonance stabilisation of the triazene group	53
3.1	Phenyltriazene and phenylformamidine atomic numbering	69
4.1	Eadie-Hofstee plot for AM90067 at various concentrations of DHF	83
4.2	Eadie-Hofstee plot for AM90067 at various concentrations of NADPH	83

4.3a	A schematic representation of the hydrogen bonding between DHFR from <i>L. casei</i> and methotrexate	85
4.3b	A schematic representation of the hydrogen bonding between DHFR from <i>L. casei</i> and pyrimethamine	85
5.1	A schematic representation of a simple neural network consisting of three layers of preceptrons	88
5.2	A schematic representation of a perceptron	88
5.3	Plot of Hodgkin electrostatic similarity index against the RL/PC DHFR IC ₅₀ selectivity ratios	92
6.1	Structures of 3'-chloropyrimethamine and its 6,2'- dimethylene bridged tricyclic analogue	97
6.2	Atomic numbering for the protonated AM90067	102
6.3	Change in the molecule energy of AM90067 as the torsion angle C7-C1H-O1H-H10H rotates through 360°	104
6.4a	Change in the molecule energy of AM90067 as the torsion angle N3-C7-C1H-O1H rotates through 360°	105
6.4b	Change in the molecule energy, between -20 and 5 kcal mol ⁻¹ , of AM90067 as the torsion angle N3-C7-C1H-O1H rotates through 360°	105
6.5	Newman projections of the low energy conformations as the torsion angle N3-C7-C1H-O1H rotates through 360°	106
6.6	Change in the molecule energy of AM90067 as the torsion angle N3-C8-C1B-C2B rotates through 360°	107
6.7	Newman projections along the bond C1B-C8 of the two low energy conformations	107
6.8	Energy contour plot showing the calculated Molecule Energy when the torsion angles N2-N3-C7-C1H and N2-N3-C8-C1B are rotated 360° with respect to each other	109
6.9	Changes in the calculated heat of formation (ΔH_f) in AM90020 as the torsion angle N1-N2-N3-C7 rotates through 360°	110
6.10	Change in the calculated heat of formation (ΔH_f) in AM90067 as the torsion angle N1-N2-N3-C8 rotates through 360°	111

6.11	¹ H NMR spectra of AM90067 in DMSO at (a) 298 K and at (b) 298 K and (c) 320 K after D ₂ O exchange	113
6.12	¹³ C NMR spectra of AM90067 in DMSO at 298 K, 313 K and 353 K. At 298 K the DEPT. spectrum is also displayed	118
6.13	Changes in the calculated heat of formation (ΔH_f) in AM90020 and the N2 to C11 separation distance as the torsion angle C2P-C3P-N1-N2 rotates 360°	121
6.14	Changes in the calculated heat of formation (ΔH_f) in AM90067 and the N2 to C11 separation distance as the torsion angle C2P-C3P-N1-N2 rotates 360°	122
6.15	Energy contour plot showing the calculated Molecule Energy for AM90020 when the torsion angles C6D-C5D-C1P-C2P and C5D-C6D-C61D-C62D are rotated through 360° with respect to each other	123
6.16	Energy contour plot showing the calculated Molecule Energy for AM90067 when the torsion angles C6D-C5D-C1P-C2P and C5D-C6D-C61D-C62D are rotated through 360° with respect to each other	124
7.1	Scheme for the knowledge-based modeling of homologous proteins.	128
7.2	Four DHFR structures superimposed onto the model human DHFR	145
7.3	Superimposed structurally conserved regions for the five DHFR structures	145
7.4	AM90020 in the active site of <i>P. carinii</i> DHFR with the torsion C4D-C5D-C1P-C2P at 270° and the 3'-(3'',3''-dimethyltriazene-1''-yl) substituent in its crystallographic conformation	150
7.5	The 3'-(3'',3''-dimethyltriazene-1''-yl) substituent of AM90020 occupying the nicotinamide region of the NADPH binding site	150
7.6	AM90020 docked in the active site model of <i>P. carinii</i> DHFR with the C4D-C5D-C1P-C2P torsion set to 90°	151
7.7	AM90067 docked in the active site of <i>P. carinii</i> DHFR. Contact between the 3''-benzyl moiety and residue 154 is highlighted	153

7.8	AM90067 docked in the active site model of <i>P. carinii</i> DHFR.	153
8.1	Polynuclear <i>P. carinii</i> cyst stained with Leishmann's stain	162
8.2	Direct immunofluorescent-antibody stained <i>P. carinii</i> cyst at 520 nm	162
8.3	Scatter diagram of data points for the control and pentamidine isethionate dosing groups against the number of cysts per rat lung	164
8.4	Scatter diagram for the control and AM90067 dosing groups against the number of cysts per lung	164
9.1	The structure of AM92008	167

List of Tables

1.1	IC ₅₀ s against rat liver and <i>P. carinii</i> DHFR of clinically used DHFR inhibitors and pyrimethamine analogues	43
2.1	Intermolecular contacts in the crystal structure of AM90021	54
2.2	Experimentally determined bond torsions for AM90021 and RMS bond torsions calculated from other pyrimethamine analogues	55
2.3a	Triazene bond lengths and RMS bond lengths from phenyltriazene and phenylformamidine crystal structures	56
2.3b	Triazene bond angles for AM90021 and RMS bond angles from phenyltriazene and phenylformamidine crystal structures	56
2.3c	Bond torsions for AM90021 and RMS bond torsions from phenyltriazene and phenylformamidine crystal structures	57
2.4	Positional parameters (fractional coordinates) of non-hydrogen atoms	60
2.5	Table 2.5 Anisotropic temperature factors, and isotropic temperature factors where shown, for non-hydrogen atoms	61
2.6	Positional parameters of hydrogen atoms and isotropic temperature factors	62
2.7	Bond lengths	63
2.8	Bond angles	64
3.1a	RMS bond lengths from phenyltriazene and phenylformamidine crystal structures and calculated triazene bond lengths	72
3.1b	RMS bond angles from phenyltriazene and phenylformamidine crystal structures and calculated triazene bond angles	72
3.1c	RMS bond torsions from phenyltriazene and phenylformamidine crystal structures calculated triazene bond torsions	72
3.2a	Calculated partial charges of the phenyltriazene atoms	75
3.2b	Calculated partial charges for the phenylformamidine	75
3.3	ASP molecular similarity indices of optimised and fixed point crystal structure geometries compared with the electrostatic potential of the original 3,3-dimethyl-1-(2'-chloro)phenyltriazene calculated using STO-3G basis set	76
4.1	Biological activities for pyrimethamine, trimethoprim and non-heterocyclic triazenyl pyrimethamine analogues	81

4.2	Biological activities for pyrimethamine, trimethoprim and heterocyclic triazenyl pyrimethamine analogues	82
5.1	The triazenyl substituent volumes and the partial charges used as training parameters for neural network	94
5.2	Experimental and calculated RL/PC DHFR IC ₅₀ activity ratios of the 3'-(3'',3''-disubstitutedtriazene-1''-yl) analogues of pyrimethamine	95
6.1	Selected optimised and single point geometries of N1-N2-N3-C8 rotamers of AM90067	112
6.2	¹ H NMR data for AM90067 at 298 K in DMSO before D ₂ O exchange	114
6.3	¹³ C NMR data for AM90067 at 298 K	115
6.4	¹³ C NMR chemical shift data for AM90067 at 313 K and 353 K	115
6.5	The calculated heat of formation and the original and optimised torsion angles for the terminal triazene hydroxethyl and benzyl substituents of AM90067 when the torsion angle N1-N2-N3-C7 was 0° (1 - 12) and 180° (13 - 24)	120
6.6	The calculated heat of formation and selected original and optimised torsion angle values for AM90020 after optimisation	126
6.7	The calculated heat of formation and selected initial and optimised torsion angle values for AM90067 after optimisation	126
7.1	Sequence homology data from aligning known DHFR sequences with the <i>P. carinii</i> DHFR sequence	139
7.2	CLUSTAL multiple sequence alignment of twenty DHFR sequences	142
7.3	<i>P. carinii</i> DHFR sequence aligned with determined DHFR structures	144
7.4	Energy of interaction (ΔH_i) for the low energy conformations of AM90067 docked in the <i>P. carinii</i> DHFR active site model	154
7.5	Energy of interaction (ΔH_i) for the low energy conformations of AM90067 bound in the human DHFR active site model	155
8.1	Dosing groups and cyst counts of the animals prematurely withdrawn from treatment	161

Abbreviations

Å	-	Angstrom(s)
AIDS	-	Acquired Immunodeficiency Syndrome
AM1	-	Austin Method 1
ASP	-	automated similarity package
CADD	-	computer-aided drug design
CH ₂ -THF	-	<i>N</i> ⁵ , <i>N</i> ¹⁰ -methylenetetrahydrofolic acid
.cssr	-	crystal structure search and retrieval
°	-	degrees
δ	-	chemical shift
°C	-	degrees centigrade
DEPT	-	Distortionless Enhancement by Polarisation Transfer
ΔH _f	-	heat of formation
ΔH _i	-	energy of interaction
DHF	-	7,8-dihydrofolic acid
DHFR	-	dihydrofolate reductase
DHPS	-	7,8-dihydropteroate synthase
D _m	-	measured density
DMSO	-	dimethylsulphoxide
DNA	-	deoxyribonucleic acid
dTMP	-	deoxythymidine monophosphate
dUMP	-	deoxyuridine monophosphate
D _x	-	calculated density
e	-	electronic charge
ELISA	-	Enzyme Ligand Immunosorbent Assay
e.s.d.	-	estimated standard deviation
FA	-	folic acid
FTIC	-	fluorescein isothiocyanate
g	-	gramme(s)
g	-	gravity
HIV	-	human immunodeficient virus
hr	-	hour
Hz	-	hertz
<i>J</i>	-	coupling constant
K	-	Kelvin
kcal	-	kilocalories
litre	-	litre
M	-	Molar
mg	-	milligramme(s)
MHz	-	megahertz
min	-	minute(s)
ml	-	millilitre(s)
μl	-	microlitre(s)
mM	-	millimolar
μM	-	micromolar
MNDO	-	Modified Neglect of Diatomic Overlap
mol	-	mole
.msf	-	molecular structure file
MW	-	molecular weight
MZP	-	<i>meta</i> -azidopyrimethamine
NADPH	-	nicotinamide adenosine dinucleotide phosphate (reduced form)
NADP ⁺	-	nicotinamide adenosine dinucleotide phosphate

nm	-	nanometres
NMR	-	Nuclear Magnetic Resonance
P	-	probability
p-ABA	-	<i>para</i> -aminobenzoic acid
PCP	-	<i>Pneumocystis carinii</i> Pneumonia
.pdb	-	protein databank
PM3	-	Parameter Method 3
pO ₂	-	partial pressure of oxygen
ppm	-	parts per million
QSAR	-	quantitative structure-activity relationship
RL/EC	-	rat liver to <i>Escherichia coli</i>
RL/PC	-	rat liver to <i>Pneumocystis carinii</i>
RMS	-	root mean square
RNA	-	ribonucleic acid
[S]	-	substrate concentration
s.d.	-	standard deviation
SHMT	-	serine hydroxymethyl transferase
THF	-	5,6,7,8-tetrahydrofolic acid
TMP	-	trimethoprim
TS	-	thymidylate synthase
UV	-	ultraviolet
V	-	rate of reaction
vol	-	volume
Z	-	molecules in unit cell

Chapter 1 Introduction

1.1 A Brief Historical Account of *Pneumocystis carinii*

The earliest descriptions of *Pneumocystis carinii* (Chagas 1909) mistook it for part of the trypanosome life-cycle, though subsequent investigations (Delanoe & Delanoe 1912) confirmed *P. carinii* as a distinct genus and species. It was not until 1952 that any pathogenic association was made (Vanek & Jirovec 1952) when it became realised that the outbreaks of interstitial plasma cell pneumonia, a disorder common in premature and malnourished children in crowded institutions during and after World War II, were caused by *P. carinii*. The increased medical interest led to the discovery that the antitrypanosomal agent pentamidine was active against *P. carinii*, significantly improving the survival of infants afflicted with *Pneumocystis carinii* pneumonia (PCP) (Ivaldy & Paldy 1958). By the 1970's *P. carinii* was recognised as a major opportunistic pathogen amongst children with primary immunodeficiency disorders and patients of all ages receiving immunosuppressive therapy for the treatment of cancer, organ transplantation and other diseases. Development of PCP in rats by the administration of corticosteroids (Frenkel *et al.* 1966) has allowed for a greater understanding of the disease pathogenesis and transmission, as well as providing a model to test novel agents. This led to the introduction of trimethoprim-sulphamethoxazole (co-trimoxazole) for the treatment and prophylaxis of PCP (Hughes 1976). Co-trimoxazole was so successful that temporarily the urgency to understand more about this organism was reduced. However in the early 1980's, amongst otherwise healthy male homosexuals there was a significantly higher incidence of PCP, a major factor that led to the recognition of Acquired Immunodeficiency Syndrome (AIDS). Unlike non-AIDS immunosuppressed patients, many patients with AIDS could not tolerate the large doses of co-trimoxazole necessary to treat PCP and frequently suffered relapses within six to nine months of treatment withdrawal (Gazzard 1989). Hence, considerable research is being carried out to improve our understanding of this organism, thereby possibly discovering a target from which safe and effective rationally designed antipneumocystis drugs can be developed.

1.2 Taxonomic Classification of *Pneumocystis carinii*

1.2.1 Original Taxonomy

Since its discovery the taxonomic position of *P. carinii* has always been debated. Based on inconclusive characterisation criteria, morphological similarities between life cycle stages, susceptibility and resistance to antiprotozoal and antifungal agents respectively, and the inability to cultivate the organism in either cell free or typical fungal media, *P. carinii* was generally accepted as a protozoan. Doubt over this classification existed. This was due to the staining of *P. carinii* with toluidine blue and methenamine silver, both generally used for staining the cell walls of fungi, and examination of the *P. carinii* cell (Vavra & Kureca 1970) which revealed an ultrastructure with a greater fungal than protozoan resemblance. Because of the difficulties in obtaining this organism in sufficiently large quantities more conclusive genetic and biochemical data were, until recently, unavailable.

1.2.2 Genetic Evidence for the Fungal Classification of *P. carinii*

Ribosomal ribonucleic acid (rRNA) sequences have been used widely and considered the method of choice to assess phylogenetic relationships. Analysis of sequences of 16S small rRNA (srRNA), (Edman *et al.* 1988), a 5S srRNA (Watanabe *et al.* 1989) and 18S srRNA (Wakefield *et al.* 1992, Wakefield *et al.* 1993) gave strong evidence supporting classification of *P. carinii* as a fungus. However the researchers were inconsistent in determining which fungal class most closely resembled *P. carinii*. Edman *et al.* (1988) concluded that *P. carinii* was closely related to the Ascomycota *Saccharomyces cerevisiae*, whereas Watanabe *et al.* (1989) suggested there was a closer relationship, not with the Ascomycota, but with the Protista, a smaller group of fungi containing the Rhizopoda, Myxomycota and the Zygomycota. Wakefield *et al.* (1993), despite criticisms from Taylor and Bowman (1993) agreeing with Edman's classification, showed *P. carinii* to have a greater affinity to another common fungal class, the Basidiomycota, and in particular the Ustomycetes red yeast fungi.

1.2.3 Biochemical Evidence for the Taxonomic Classification of *P. carinii*

1.2.3.1 Dihydrofolate Reductase and Thymidylate Synthase Characterisation

Two enzymes, dihydrofolate reductase (DHFR) and thymidylate synthase (TS) are implicated directly in deoxyribonucleic acid (DNA) biosynthesis (figure 1.2). Data derived from the molecular cloning of the DHFR (Edman *et al.* 1989a) and the TS (Edman *et al.* 1989b) showed that two separate genes, on different chromosomes, produced two distinct enzymes. This contrasted with the situation in the parasitic protozoa *Leishmania* (Grumant *et al.* 1986), *Plasmodia* (Zolg *et al.* 1989) and *Toxoplasma* (Kovacs *et al.* 1989) where a single gene expresses a bifunctional protein. Although this does not confirm *P. carinii*'s fungal nature it does provide considerable weight against its protozoal classification. In addition, analysis of the DHFR amino acid sequence showed greatest homology with the *S. cerevisiae* enzyme (Edman *et al.* 1989a).

1.2.3.2 Elongation Factor 3

Protein synthesis in eukaryotic cells requires two soluble protein elongation factors (EF-1 and EF-2) for the peptide elongation cycle. In several genera of fungi a third factor (EF-3) is required. Using immunological methods (Jackson *et al.* 1991) EF-3 could not be detected in *P. carinii*, but more recently Yomawong *et al.* (1992) discovered an EF-3 gene. EF-3 is not common to all fungi and the taxonomic value of this discovery remains uncertain.

1.2.3.3 Antigenicity

The cross-reactivity of antibodies raised against *P. carinii* does not confirm its phylogenetic relationship in its own right, but can provide strong evidence to support a particular classification. A monoclonal antibody derived from rat *P. carinii* was able to cross-react with fifteen antigens from fifty two fungi (Lundgren *et al.* 1992). The strongest cross-reactivity was with *S. cerevisiae*. Furthermore the antigens from six protozoa tested failed to exhibit any cross-reactivity with the *P. carinii* antibody.

1.2.4 Conclusions of the Taxonomic Classification of *P. carinii*

Recent molecular and biochemical data strongly suggest that *P. carinii* is a fungus, though as yet the data are inconclusive as to which, if any, fungal group it most resembles. As more is understood about *P. carinii* and standard approaches for phylogeneticity are devised, then a more accurate representation of the relationships of this organism with other eukaryotes can be drawn.

1.3 Life Cycle of *Pneumocystis carinii*

A generally accepted life cycle has yet to be agreed upon. However on the basis of electron microscopic studies it has been hypothesised to include at least three basic stages (figure 1.1) (Dei-Cas *et al.* 1992): a mononucleated trophozoite, 1 - 4 μm in length; a precyst approximately 5 μm in length; and a mature cyst, 5 - 8 μm long, containing up to eight intracystic bodies which will evolve into trophozoites. Apart from physical size there are believed to be antigenic and metabolic differences between the stages, as well as the progressive thickening of the cell wall from two to three layers as the trophozoite evolves to a mature cyst. A more complex life cycle (Yoshida 1989) attempts to make logical explanations for observations noted through histologic studies, including the possibility of sexual reproduction. Neither of these models accounts for the plausible existence of an environmental phase implicated in the transmission of the organism between hosts, though complete elucidation of the *P. carinii* life cycle could be achieved with the development of an *in vitro* cultivation, thereby removing interference from host cells.



Aston University

Content has been removed for copyright reasons

Figure 1.1 The life cycle of *Pneumocystis carinii* (Dei-Cas *et al* 1992)
The different stages are represented as seen on transmission electron microscopy.
AL: alveolar lumen; CL: capillary lumen; RBC: red blood cell within a pulmonary
capillary blood-vessel; T: trophozoite attached to the alveolar epithelium.

1.4 Epidemiology and Pathogenesis of *Pneumocystis carinii* Pneumonia

1.4.1 Epidemiology

A series of experiments using the immunosuppressed rat model (Hendley and Weller 1971) and the Severely Compromised Immunodeficient (SCID) mouse (Walzer *et al* 1977) demonstrated that in germ-free animals PCP would develop if in distant contact and would develop more rapidly when housed with infected animals. This suggested an airborne transmission of the organism. Supportive research (Hughes 1982, Hughes *et al* 1983) found that food or water was not a crucial factor for acquiring the infection in steroid treated germ-free rats. Although airborne transmission has yet to be confirmed in humans, outbreaks or clusters of infections in crowded institutions, hospitals (Fenlon *et al* 1985) and organ transplant units (Santiago-Delpin *et al* 1988) does suggest a person to person spread. The fact that over 75 % of children by four years are seropositive to *P. carinii* antigens (Cushion *et al* 1991) implies the organism is

commonly present in the environment, supporting the airborne hypothesis. This also adds weight to the previously suggested possibility of an undiscovered, dormant stage in the life cycle that can persist without any host interaction. Until recently, it was believed that in the immunocompromised PCP developed, after the first exposure, from the reactivation of a latent organism. Wakefield *et al.* (1990), using the very sensitive polymerase chain reaction technique were unable to show the presence of *P. carinii* in immunocompetent patients. However, samples from patients with histologically documented cases of infection gave positive signals. These findings suggested that the reactivation theory is not the mechanism leading to fulminant infection and that the infection originates from recent acquisition of airborne organisms.

A definitive study of the global distribution of *P. carinii* has still to be completed. However, a review (Hughes 1987) reported incidences of infection in temperate, tropical and polar countries alike, without any geoclimatic preferences. Throughout the world prevalence does vary, but because of problems with disease recognition and surveillance in developing countries, it remains to be seen whether there are actual differences in the virulence of *Pneumocystis* strains or the host.

1.4.2 Pathogenesis of *P. carinii*

From a very early age immunity to *P. carinii* can be seen in most people. It is therefore logical that the one condition vital for sufficient proliferation of the organism to cause the symptoms of pneumonia is suppression of the immune system. Understanding of the invasive process has increased dramatically due to the technological advancements in tissue evaluation techniques, as well as the greater number of cases reported and studied. In its trophozoite form *P. carinii* associates in the alveoli with type I pneumocytes (Yoneda & Walzer 1983). Electron microscopy studies have indicated that although both cells remain intact there is a “damaging or aggressive” association (Ham *et al.* 1971) due to changes in the host cell surface causing a diminished functionality (Yoneda & Walzer 1984). Proliferation of *P. carinii* trophozoites and the formation of cysts, coupled with the host’s desquamation of the cell lining, eventually lead to the alveoli filling with a foamy exudate (Kim *et al.* 1987). Using methenamine silver stain the presence of *P. carinii* can be readily detected. The classic symptoms of patients with

Pneumocystis carinii Pneumonia are a dry, nonproductive cough, fever, weight loss and a shortness of breath (Gazzard 1989). Disease progression, characterised by ventilatory impairment, is caused by the filling of the interstitial spaces and alveoli by the organism, serous fluids, and lymphocytes. Subsequent development of adult respiratory distress syndrome may ensue over a period of weeks and the patient may die of respiratory failure (Berkowitz 1985).

With a greater number of patients undergoing chemically induced immunosuppression as part of the necessary post-operative organ transplant chemotherapy, cancer chemotherapy or as a direct consequence of developing AIDS from contraction of the human immunodeficiency virus (HIV) virus, so *P. carinii* has become a major pathogen to these patients and a common cause of mortality. Although molecular biological knowledge has deduced that *P. carinii* is most likely fungal in nature, the organism is resistant to classical antifungal chemotherapy. However the demonstration of *de novo* folate biosynthesis (Kovacs *et al.* 1989) and the sensitivity to co-trimoxazole suggests that this is a potential target for directing antipneumocystis chemotherapy.

1.5 Antifolate Treatment and Prophylaxis of *Pneumocystis carinii* Pneumonia

1.5.1 Introduction

Very few advances have been made in finding a specific cure for HIV infections (McLeod & Hammer 1992). Therefore treatment and prevention of the occurrence and reoccurrence of PCP remains the principal objective in improving the life-expectancy and quality of life for patients with full-blown AIDS (Martin *et al.* 1992). Although antimicrobial prophylaxis and chemotherapy are readily available (Bernard *et al.* 1992) there is still a need to develop more effective, better tolerated antimicrobial agents (Medina *et al.* 1990, Sattler & Feinberg 1992). This is highlighted by the fact that nearly 80% of patients with HIV infections have suffered at least one episode of PCP (Murray *et al.* 1984), of which there was a 20% mortality rate during the first exposure (Zackrisson & Tsou 1991). Also there are limitations with the current conventional antipneumocystis chemotherapeutic agents. In AIDS patients, co-trimoxazole and pentamidine cause severe toxicity reactions, require large doses and long dosing

periods to have any therapeutic efficacy and have a high incidence of treatment failure (Gazzard 1989). Finally, with the anticipated life-expectancy of AIDS patients increasing through antiretroviral therapy, it is correct to expect an increased occurrence of PCP (Volberding 1991).

Remarkable advances have been made to combat this otherwise fatal infection (Smith & Gazzard 1991). These have included new regimens and reformulations of the conventional therapies aimed at retaining efficacy, but reducing the severity and incidence of toxic effects. The development of novel agents has progressed considerably, but is hindered by the difficulties in cultivating the organism. This has meant the evaluation of new entities has become a complex and time consuming process. Evaluation and re-evaluation of novel and current antimicrobials has consistently shown *P. carinii* to exhibit a greater susceptibility to antiprotozoal and antiparasitic agents (Cushion *et al.* 1991) than to antifungals or antibacterials. Hence one rational approach to develop rapidly novel antipneumocystis agents was to by-pass the early developmental processes and test current and novel antiprotozoal agents (Jones *et al.* 1990). Certain compounds have shown very good clinical potential as antipneumocystis agents (Dohn *et al.* 1992), whereas others have demonstrated a synergistic activity when combined with another broad-spectrum antimicrobial (Toma 1991) or provided a novel lead compound for further development (Schmatz *et al.* 1992). Other rational methods for discovering novel antipneumocystis compounds have utilised the increased molecular and biological data available for *P. carinii*, identified enzymes that could be exploited as chemotherapeutic targets and tested known inhibitors against *P. carinii in vitro* and *in vivo*. Randomly screening classes of broad spectrum antimicrobials against experimental PCP has had some success in discovering new lead compounds, but so far none with any clinical potential.

Thus it remains to be seen whether any of these novel antipneumocystis agents will have a therapeutic impact, especially in AIDS patients. However, with more compounds now being evaluated and recognised as having a clinical potential, then their different toxicity profiles will at least increase the choice of alternative agents when poor therapeutic responses are observed.

1.5.2 Folate Biochemistry

Because folate synthesis and cycle are indirectly involved in DNA biosynthesis (figure 1.2), it is a vital biochemical process for proliferating organisms and an ideal target for chemotherapy. In mammalian tissues dihydrofolate reductase (DHFR) catalyses the reduced nicotinamide adenosine dinucleotide phosphate dependent reduction of folic acid (FA) (figure 1.3a) to 7,8-dihydrofolic acid (DHF) (figure 1.3b). Dietary or biochemically synthesised DHF is then reduced by DHFR to 5,6,7,8-tetrahydrofolate (THF) (figure 1.3c) before the enzyme serine hydroxymethyl transferase (SHMT) introduces a N^5,N^{10} -methylene bridge to the THF. 5,10-MethyleneTHF (CH_2 -THF) (figure 1.3d) acts as a source of carbon for biosynthetic pathways involved in the addition of a single carbon atom to a biochemical substrate. In most of these processes removal of the methylene unit produces THF which, via SHMT, can readily be regenerated back to CH_2 -THF. In these circumstances CH_2 -THF is only required at co-factor levels. However in one biological pathway, the thymidylate synthase (TS) assisted conversion of deoxyuridine monophosphate (dUMP) to deoxythymidine monophosphate (dTMP), a base precursor in DNA biosynthesis, CH_2 -THF also acts as a hydrogen donor and is therefore converted to DHF. Before levels of CH_2 -THF can be restored, DHFR must reconvert the DHF to THF, hence for this process higher substrate levels of CH_2 -THF are required. Therefore inhibition of DHFR will eventually cause the depletion of CH_2 -THF reserves necessary for dTMP biosynthesis, thereby preventing DNA replication and eventually resulting in a “thymine-less death.” Although DHFR is common to all species, its potential as a target for antipneumocystis chemotherapy was realised when kinetic differences between the *P. carinii* and mammalian DHFRs suggested the active sites of the two enzymes were structurally different (Kovacs *et al.* 1990).

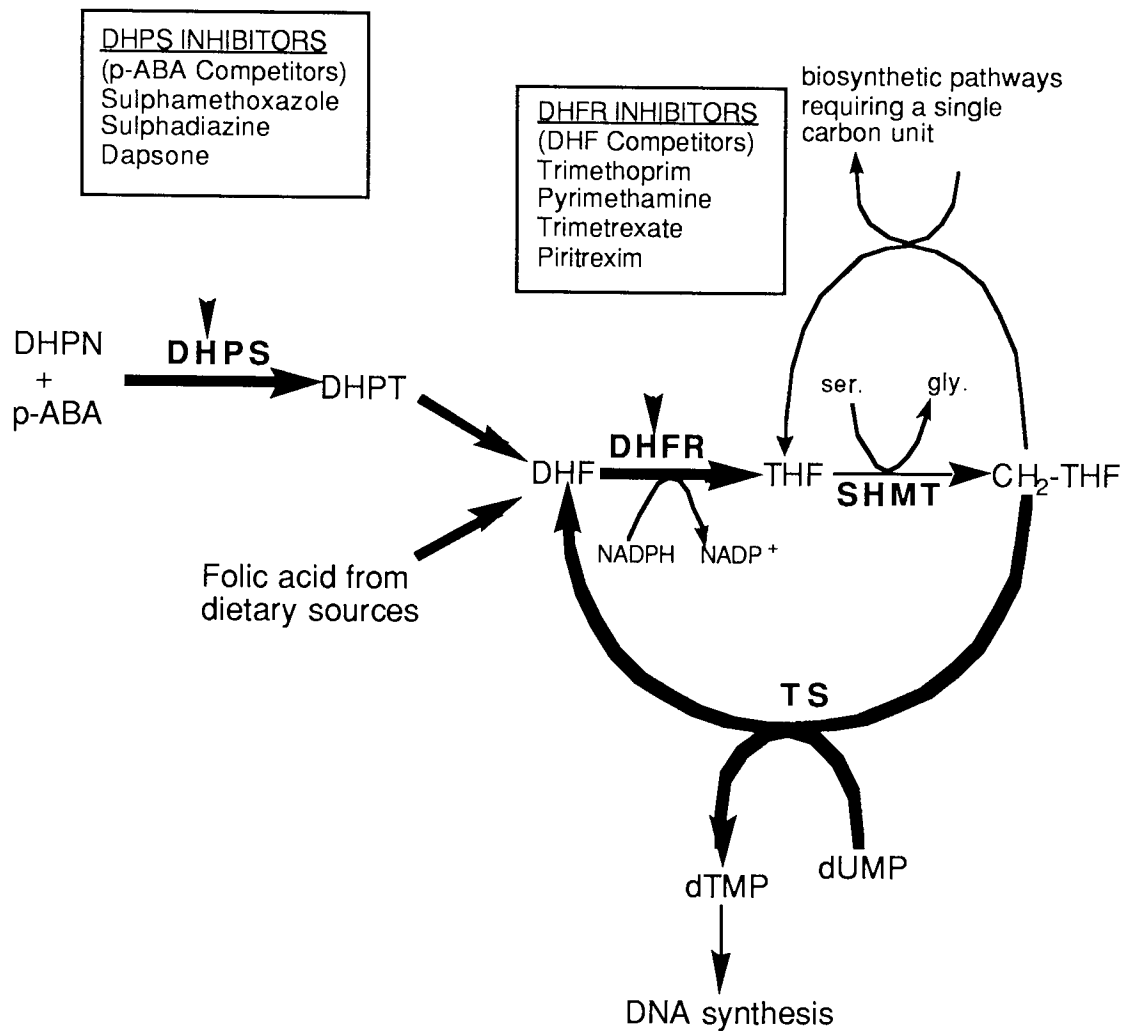


Figure 1.2 Folate biosynthesis and the folate cycle

Abbreviations : Enzymes :- **DHPS**, dihydropteroate synthase; **DHFR**, dihydrofolate reductase; **SHMT**, serine hydroxymethyl transferase, **TS**, thymidylate synthase.

Substrates :- DHPN, dihydropteridine; p-ABA, *para*-aminobenzoic acid; DHPT, dihydropteroate; DHF dihydrofolate; NADPH, nicotinamide adenosine dinucleotide phosphate (reduced form); NADP⁺, nicotinamide adenosine dinucleotide phosphate; THF, tetrahydrofolate; CH₂-THF, 5,10-methylenetetrahydrofolate; ser., serine; gly., glycine; dUMP, deoxyuridine monophosphate; dTMP, deoxythymidine monophosphate.

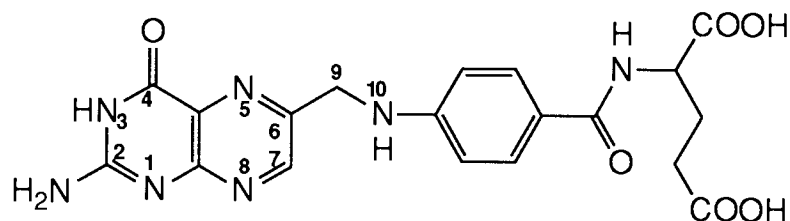


Figure 1.3a The structure of folic acid

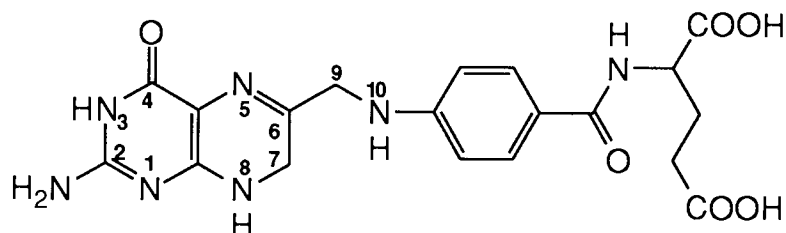


Figure 1.3b The structure of 7,8-dihydrofolic acid

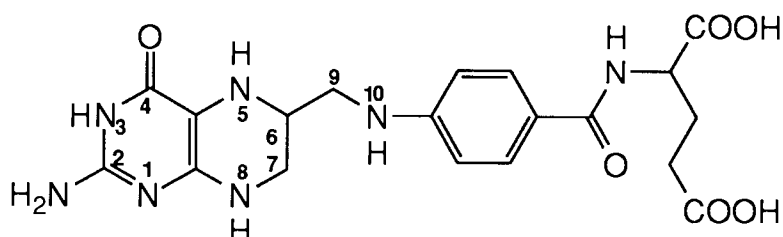


Figure 1.3c The structure of 5,6,7,8-tetrahydrofolic acid

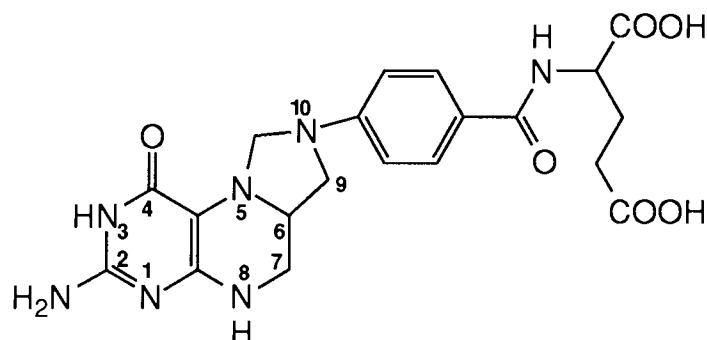


Figure 1.3d The structure of N^5,N^{10} -methylene-5,6,7,8-tetrahydrofolic acid

Mammalian cells do not have the ability to synthesise FA, but rely on the active transport into the cells of folate structures from dietary sources. Many microorganisms, including *P. carinii*, do not have this transport facility (Allegra *et al.* 1987b), but are

capable of *de novo* folate biosynthesis (Kovacs *et al* 1989). FA is produced by the enzyme dihydropteroate synthase (DHPS) using the two principal components *para*-aminobenzoic acid (p-ABA) and dihydropteridine (DHP). Sulphonamides, competitive p-ABA antagonists, inhibit FA biosynthesis. Their selective antimicrobial action is a result of the absence of this biochemical process in mammalian tissue. It is believed that there is a therapeutic advantage to be gained by combining a p-ABA antagonist with an DHFR inhibitor, thereby inhibiting two separate stages in the folate metabolism.

1.5.3 Antifolates used in the Treatment of *P. carinii* Pneumonia

1.5.3.1 Trimethoprim

In vitro trimethoprim (figure 1.4a), despite having a poor inhibitory activity profile against non-bacterial DHFRs, demonstrated a greater inhibition of *P. carinii* as opposed to mammalian DHFR (Broughton & Queener 1991). *In vivo* when trimethoprim was used alone it was ineffective against steroid-induced experimental PCP in rats (Hughes *et al.* 1977). However, when trimethoprim was co-administered with the p-ABA antagonists sulphamethoxazole (Hughes *et al.* 1974) (figure 1.4b) or dapsone (Lee *et al.* 1989) (figure 1.9) a pronounced antipneumocystis effect was observed.

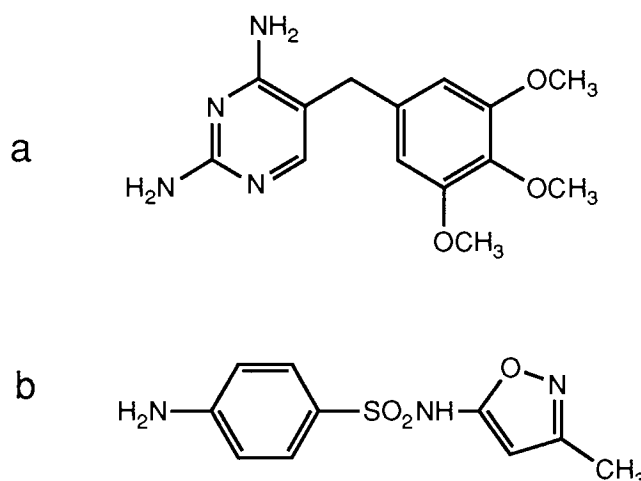


Figure 1.4 The structures of trimethoprim (a) and sulphamethoxazole (b)

Co-trimoxazole is a combination of trimethoprim and sulphamethoxazole in a ratio of one part to five parts. It has remained the treatment of choice since Hughes (1976) originally demonstrated its efficacy against childhood *Pneumocystis* infections. Shortly

afterwards it began to be used prophylactically (Hughes *et al.* 1977) and today it is regarded as a benchmark against which other treatments of PCP are compared.

The optimal regimen of co-trimoxazole for a maximum therapeutic effect and minimal toxicity in *Pneumocystis* infections remains undecided. However, the twice daily 640:3200 mg dose of co-trimoxazole that has been used effectively to treat PCP in patients with AIDS (Eeftinck Schaffenkerk *et al.* 1990) is significantly larger than that required to treat bacterial infections in the immunocompetent (80:400 - 160:800 mg twice daily). Furthermore, there are limitations with co-trimoxazole for the treatment of PCP in patients with AIDS. Relapses within nine months of treatment withdrawal are common (Gazzard 1989). Also, in patients with AIDS, more so than other immunocompromised patients with PCP, dose-related adverse and toxic reactions to co-trimoxazole are observed. These include skin rashes (Coker *et al.* 1992), liver function changes (Johnson *et al.* 1990), blood dyscrasias agranulocytosis and leukopenia (Gordin *et al.* 1984), fever (Ruskin & La Riviere 1991) and hypotension (Coker *et al.* 1992). It has been suggested that the sulphamethoxazole is responsible for the serious side-effects (Alberti-Flori *et al.* 1989). However, because trimethoprim alone is ineffective, co-trimoxazole remains an integral component of antipneumocystis chemotherapy.

In summary, the large doses of co-trimoxazole can be regarded as an effective treatment for PCP. However, in AIDS patients, a high incidence of adverse reactions and frequent relapses suggest that more effective antifolates against *P. carinii* are necessary.

1.5.3.2 Pyrimethamine

When used alone, pyrimethamine (figure 1.5a) exhibits a modest antipneumocystis activity (Walzer *et al.* 1992), though not of any clinical significance. Like trimethoprim, it is far more effective when co-administered with a p-ABA antagonist. It was Frenkel *et al.* (1966) who originally demonstrated that the combination of pyrimethamine and sulphadoxine (figure 1.5b) is an effective treatment against experimental PCP. Subsequent investigations and clinical trials showed this combination to be the first suitable alternative to pentamidine for PCP (Hughes *et al.* 1974), though it was soon disregarded when co-trimoxazole became the treatment of choice.

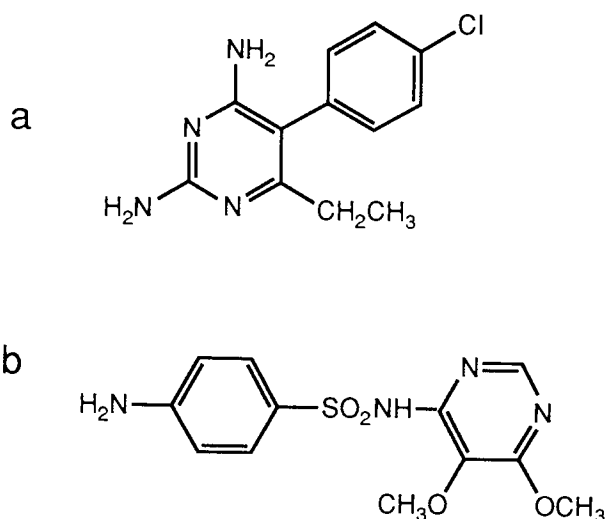


Figure 1.5 The structures of pyrimethamine (a) and sulphadoxine (b)

Pyrimethamine-sulphadoxine was reconsidered for PCP treatment in patients with AIDS after clinical trial follow-up studies revealed that very few patients receiving this combination to treat or prevent central nervous system *Toxoplasma gondii* infections contracted PCP (Heald *et al.* 1991). Although it has not yet become a front line treatment for PCP, pyrimethamine-sulphadoxine has been used successfully in *P. carinii* chemoprophylaxis (Danneman *et al.* 1992). Further, reassessment has involved trials to determine if there would be any therapeutic advantage in combining pyrimethamine with other antimicrobial agents, including clindamycin (Ruf *et al.* 1993) and dapsone (Clotet *et al.* 1991).

1.5.3.3 Trimetrexate

Trimetrexate (figure 1.6a) is a very potent lipid-soluble DHFR inhibitor that was used originally for its antineoplastic properties in methotrexate-resistant tumours (Lin & Bertino 1991). Subsequently, *in vitro*, it was shown to be a very potent inhibitor of *P. carinii* DHFR, approximately 1500 times more potent than trimethoprim (Allegra *et al.* 1987a). Hence, it was recognised as having a great potential as an antipneumocystis agent. Furthermore, due to its powerful DHFR inhibitory activity, co-administration of a p-ABA antagonist was not necessary. Because it was believed that p-ABA antagonists were responsible for the many adverse reactions observed in patients with AIDS, (van der Ven 1991) using only one antifolate component, a DHFR inhibitor, may provide an improvement in the treatment of PCP.

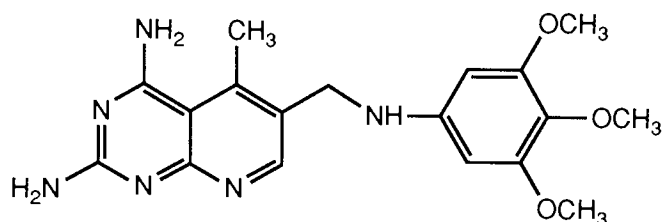


Figure 1.6a The structure of trimetrexate

Phase I-II escalation trials have determined the optimal doses of trimetrexate needed to produce a maximum efficacy to toxicity ratio, hence the most clinically acceptable regimen for the treatment PCP in AIDS patients (Sattler *et al.* 1990). A high incidence of haematologic toxicities was observed, in particular neutropenia and thrombocytopenia, but, by dosage modification, the treatment could continue. Other adverse effects that were noted included hypotension and a rash formation (Sattler *et al.* 1990). Thus any improved efficacy of treating PCP with a DHFR inhibitor alone was marred by the emergence of other adverse effects.

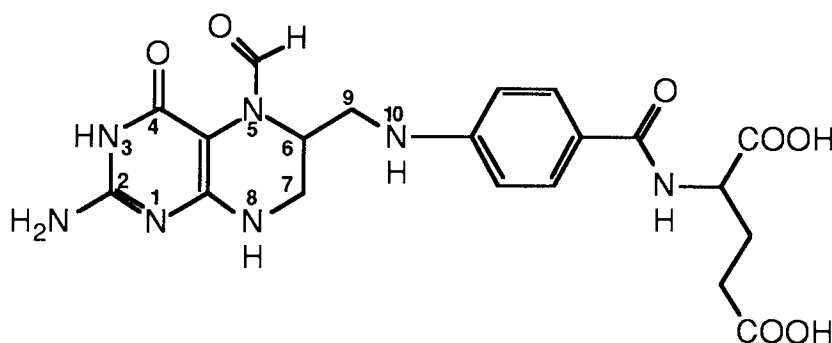


Figure 1.6b The structure of leucovorin

Reduction of trimetrexate toxicity was partially achieved by the co-administration of the reduced folate leucovorin (figure 1.6b). Cell membrane transport studies with radiolabelled folates and antifolates have shown that, unlike mammalian cells, *P. carinii* cannot actively transport classical folate structures across its cell membrane (Allegra *et al.* 1987a). Instead it relies on *de novo* synthesis as a source of folic acid (Kovacs *et al.* 1989) (figure 1.2). Trimetrexate, being lipid soluble and therefore able to diffuse across the membrane, does not require this active transport process to penetrate into the cells. Thus, despite it having a similar inhibitory activity toward mammalian

DHFR, a major therapeutic advantage was achieved by combining trimetrexate with leucovorin (Allegra *et al.* 1987b). Leucovorin can be actively transported across the host cell membrane, thereby allowing the host selective regeneration of THF levels, and a rescue therapy for the trimetrexate toxicity. *P. carinii*, without this transport system, is left exposed to the full effects of trimetrexate.

Trimetrexate in combination with leucovorin has been shown to be very useful in antipneumocystis therapy. However, due to its potency it requires careful monitoring, and therefore its use has been restricted to cases that have not responded to the conventional treatments, co-trimoxazole and pentamidine.

1.5.3.4 Piritrexim

Piritrexim (figure 1.7) is another potent lipid soluble inhibitor of DHFR that has been combined with leucovorin and tested *in vivo* against experimental PCP. Structurally it is very similar to trimetrexate. It has a shorter half life than trimetrexate (Kovacs *et al.* 1987), which was considered a therapeutic advantage because its more rapid clearance would result in a reduction in the frequency and severity of adverse reactions due to prolonged exposure. However, in phase I clinical trials in patients with AIDS, with leucovorin co-administration, piritrexim produced an unacceptable haematological toxicity. This has precluded it from other trials for the treatment of PCP in patients with AIDS (Falloon *et al.* 1990).

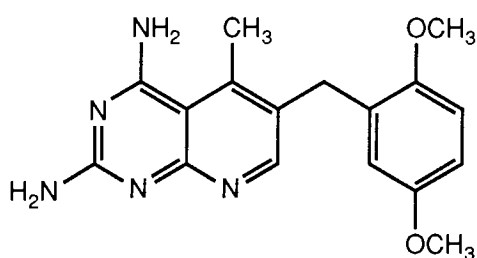


Figure 1.7 The structure of piritrexim

1.5.3.5 Proguanil and a novel Bi-guanide

Another class of DHFR inhibitors that has been screened for antipneumocystis activity is the triazines. Because of poor systemic absorption *in vivo*, triazines are administered as bi-guanide pro-drugs which are subsequently converted to their active form once

absorbed (figure 1.8). Proguanil, converted *in vivo* to cycloguanil, is used commonly in malaria chemotherapy (Greenwood 1989). Other classes of antimalarials have exhibited potent antipneumocystis activity. Although proguanil failed to produce any activity in rats, there is still sufficient justification to screen other bi-guanides.

In vitro the triazine 4,6-diamino-1,2-dihydro-2,2-dimethyl-1-(2,4,5-trichlorophenoxypropyloxy)-1,3,5-triazine hydrochloride (PS-15) is an active inhibitor of DHFR. *In vivo*, due to poor systemic absorption, PS-15 failed to elicit any discernible antipneumocystis activity. However, its corresponding biguanide N-3-(2,4,5-trichlorophenoxypropyloxy)-N'-(1-methylethyl)imidocarboimidic diamide hydrochloride (WR99210), metabolised *in vivo* to PS-15 was as efficacious as co-trimoxazole at treating experimental PCP prophylactically and curatively, at considerably lower doses, without any noticeable toxicity (Hughes *et al.* 1993). WR99210 / PS-15 chemotherapy of PCP is still at the developmental stages, and therefore it remains to be seen whether the triazine PS-15, unlike trimetrexate and piritrexim, can exhibit all the features of an ideal *P. carinii* DHFR inhibitor: the selectivity of trimethoprim, the potency of trimetrexate and produce no serious toxicities.

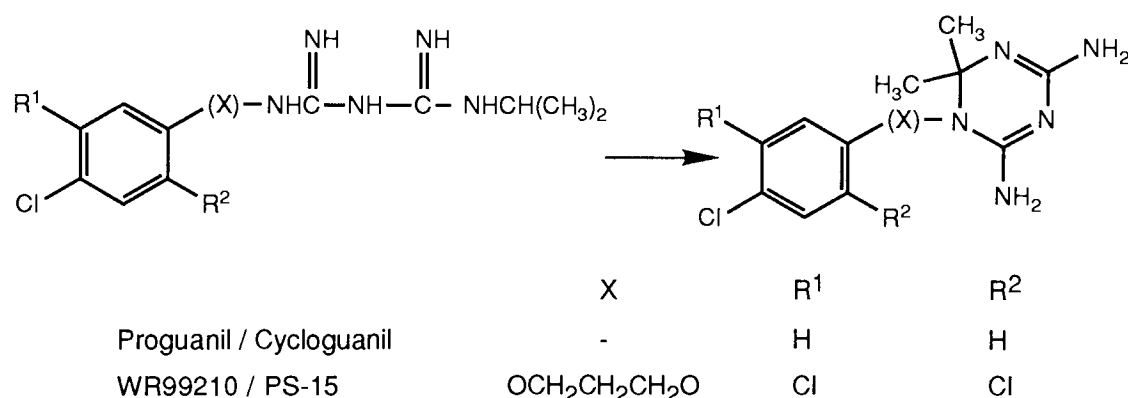


Figure 1.8 The *in vivo* conversion of bi-guanides to their corresponding triazines

1.5.3.6 Dapsone

The success of the DHPS inhibitor dapsone (figure 1.9) in malaria and leprosy chemotherapy warranted its investigation for the treatment of PCP. However, it has been shown that, when used alone, a large daily dose of dapsone provides suboptimal

therapy (Safrin *et al.* 1991). Furthermore, due to ventilatory insufficiency, there is a substantial reduction in the partial pressure of oxygen (PO₂) in arterial blood in PCP. Dose-related methaemoglobinaemia, a very common side-effect of dapsone, may further reduce the PO₂ sufficiently to precipitate prematurely the life-threatening respiratory failure that is often the cause of death from PCP (Gallant *et al.* 1991). When smaller doses were administered in combination with either trimethoprim (Medina *et al.* 1990) or pyrimethamine (Clotet *et al.* 1991) a profound antipneumocystis effect was observed. Further studies have shown that it is as well tolerated and as efficacious as co-trimoxazole in the treatment of mild to moderately severe PCP (Bernard *et al.* 1992). Therefore, the use of dapsone in the management of PCP in patients with AIDS has been restricted to smaller, prophylactic doses, in combination with one of the aforementioned DHFR inhibitors.

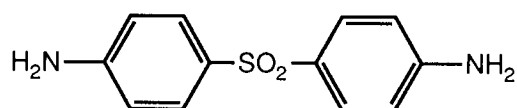


Figure 1.9 The structure of dapsone

1.5.3.7 Conclusions for the use of Antifolate Chemotherapy in the Treatment and Prophylaxis of *P. carinii* Pneumonia

The successful clinical use of DHFR inhibitors and p-ABA antagonists, alone or in combination, clearly demonstrates antifolate chemotherapy as an ideal approach for directing antipneumocystis treatment. This is due to the fact that *de novo* folic acid biosynthesis has been demonstrated in *P. carinii*. Therefore inhibition can ultimately lead to the “thymine-less” death of the target organism. Although co-trimoxazole remains one of the first line treatments of PCP, its suitability remains in doubt because of high incidence of toxic reactions due to the p-ABA antagonists, causing patient intolerability. The more potent lipophilic DHFR inhibitors that have successfully treated PCP alone have also exhibited unwanted side-effects. However this may be due to inappropriate use, as these compounds do not demonstrate any greater activity for *P. carinii* DHFR. The potent inhibitory, species specific differential activities that various chemical classes of DHFR inhibitors exhibit can be exploited for the design of more effective antipneumocystis agents. Therefore, there is enormous clinical potential for the design of an active and therapeutically acceptable inhibitor of *P. carinii* DHFR.

1.6 Computer-Aided Drug Design

1.6.1 Introduction to Computer-Aided Drug Design

Many drugs have been discovered by developing a biological test and screening large numbers of compounds. A degree of rationality was introduced to this process when more sophisticated analytical equipment allowed rapid structural elucidation. This enabled medicinal chemists to ascertain whether particular molecular characteristics were common and responsible for compounds exhibiting particular biological activity. The characteristics often investigated were basic chemical properties such as electron distribution, interatomic distances, molecular volume and hydrogen bonding capacity. Furthermore, by determining which, if any, of these parameters were implicated in biological activity it was possible to target synthesis towards particular compounds (Hansch *et al.* 1962). These were designed specifically to incorporate and enhance any molecular feature believed to be responsible for the desired biological activity. Thus, compounds would be synthesised with the aim of maximising any activity.

Occasionally alterations to the chemical structure did not produce the predicted increase in biological activity. Often, this was a result of the chemical modifications influencing the interaction of the compound with receptor ligands (Chan & Roth 1991). Thus, methods that approached the rational drug design procedure from a biological, not chemical, perspective were required to increase the predictability of the structural characteristics with the biological activity profile.

With improving computer technology, high resolution graphics have enabled the three dimensional visualisation of atomic coordinates as molecular models. More sophisticated computer graphics have enabled the molecular modeller to replace the mechanical wire macromolecular models with detailed on-screen representations. Improvements in the processing speed of computers have allowed complex chemical calculations, associated with the molecular energy, the electronic distribution and the molecular geometry, to be routinely completed. Increased memory capacity has meant that considerable macromolecular structural data can be stored and readily retrieved. Thus, a molecular modeller with access to these facilities can investigate a compound's steric and electronic properties and deduce any favourable interactions with the target macromolecule. The ease with which a compound interacts with a target macromolecule

relates directly to its biological activity (Richards 1990). The knowledge of how a compound exerts its biological activity can be extended, within reason, to structurally similar homologues. Hence, with the visualisation of the molecular interaction, it is possible to establish which of the novel entities can fit within the size constraints of the target receptor. Furthermore, molecular modelling packages can evaluate the extent to which a compound possesses a complementary electronic surface to interact favourably with the target receptor surface (Brooks *et al.* 1983). Computer-aided drug design (CADD), therefore, has developed into a powerful investigative tool in drug discovery. Research of this nature, however, is dependent on the availability of receptor structure information. Thus, CADD can be sub-divided into three areas, where :-

- (i) the receptor structure is known
- (ii) the receptor structure can be predicted
- (iii) the receptor structure remains unknown

Despite recent advances in CADD technology, prior structural knowledge and biological testing to discover a lead active compound from which the rational design of inhibitors with improved biological activity can originate is also necessary (Moon & Howe 1991).

1.6.2 CADD Using a Known Receptor Structure

Macromolecular receptors with experimentally determined high resolution atomic structural data include DNA fragments and enzymes. Using these structures and docking novel ligands, it is possible to calculate the change in energy when a novel entity (drug B) interacts and complexes with a macromolecular structure as outlined in figure 1.10.

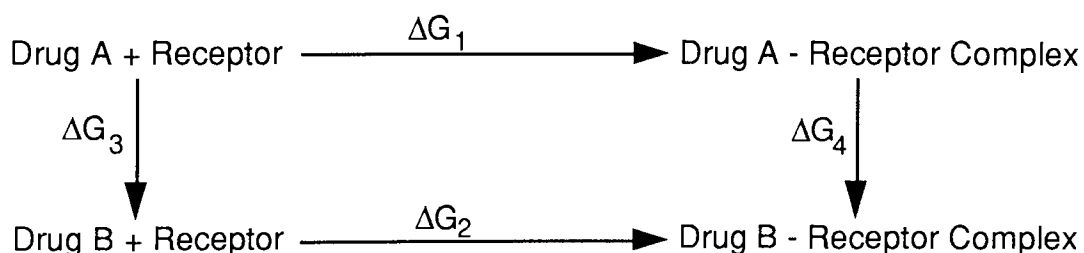


Figure 1.10 Free energies (ΔG) involved in computer aided drug design (From Richards 1992)

ΔG_1 , the change in the binding free energy when drug A is brought to the receptor from

infinity, can be determined experimentally. ΔG_3 and ΔG_4 are the changes in free energy involved in changing A to B in isolation and when complexed with the receptor respectively. In reality, these changes cannot occur but, using CADD techniques it is possible to simulate them. From the alterations to any ionic, van der Waals or hydrogen bond contacts, it is possible to calculate the changes in free energy (ΔG_3 and ΔG_4) (Andrews *et al.* 1984, Andrews 1986), claimed by Richards (1992) to be within ± 1 kcal mol⁻¹ accuracy. Thus, without any experimental analysis, it is possible to calculate accurately the change in the binding free energy for drug B (ΔG_2) that would occur as it forms a complex with the target receptor. A value of how favourably the potential ligand, drug B, binds to the target receptor site can be equated directly to its biological activity (Andrews 1986) and thereby assess, prior to synthesis, its potential usefulness.

1.6.3 CADD Using a Predicted Receptor Structure

The advent of DNA base sequencing techniques and the subsequent translation has caused an explosion in the number of determined primary polypeptide structures. It is this data that has extended the possibilities of CADD (Sali & Blundell 1990). Despite sensationalist claims, it is still beyond our capability to predict a tertiary folded protein structure directly from the primary amino acid sequence (Kolaskar & Kulkarni-Kale 1992). Lesk & Chothia (1980) showed that active site geometries of homologous proteins have remained very similar through molecular evolution. This is because of the considerable implications any alterations would have on substrate binding. Thus rational methods have been used to devise a knowledge-based approach to model protein structures (Chothia & Lesk 1986, Sali & Blundell 1990, Havel & Snow 1991). These techniques will be developed further in chapter 5.

This type of modelling has elucidated macromolecular model structures for important enzymes targets in CADD, including an HIV protease model (Weber 1990). Furthermore, model enzyme structures have given a considerable insight at a molecular level of substrate (Petukhov & Mazur 1992) and inhibitor (Yeh *et al.* 1991) binding conformation, the selective activity of a ligand between homologous proteins (Folkers *et al.* 1991) as well as ligand design (Ferenczy & Garrett 1989) that, at the time were only possible by the prediction of the active site geometries.

1.6.4 CADD Using an Unknown Receptor Structure

On the basis that the structure of a drug is essentially complementary to its binding site and that a few important interactions between the receptor and the drug are responsible for the conformation adopted, methods have been devised to project a basic image of the receptor site (Gilbert & Hyde 1992). These involve considering the relative activities of a number of ligands known to act at the same site and superimposing conserved structural features. This may require lengthy conformational analysis of flexible compounds to ensure that the correct three dimensional binding conformations are superimposed. A picture of the common or basic structural unit necessary for binding and relevant biological activity is envisaged. This entity is often referred to as the pharmacophore. From this, deductions can be made about the location of binding site groups, such as hydrogen bond donors. Introducing larger groups to the basic inhibitory skeleton can provide some information concerning the spatial constraints, such as hydrophobic pockets, within the receptor site. At best, a fuzzy and patchwork view of the receptor site can be obtained, from which the biological activity of novel agents can be predicted by evaluating how well they can fit into the pharmacophore model and the spatial constraints. Because of the assumption in the methodology, it is difficult to assess the accuracy of a pharmacophore model. However Lloyd and Andrews (1986) have been able to define a common structural basis for central nervous system acting drugs picked at random. This can be a very time-consuming process and therefore should be used only when there are no structural data available.

1.6.5 Dihydrofolate Reductase as a Target Receptor for CADD

DHFR is an ideal target receptor for antimicrobial chemotherapy because of the direct consequences inhibition has in preventing DNA biosynthesis, ultimately leading to cell death, and the remarkable differential activity the chemical classes of DHFR inhibitors exhibit against various species' DHFR (Blaney *et al.* 1984). As molecular biological knowledge increases, other potential chemotherapeutic strategies for targeting against *P. carinii* have become identified, including the use of pneumocandins to inhibit the biosynthesis of the important cell wall constituent β -D-1,3-glucan (Bartizal *et al.* 1992) and thiobendazole to inhibit microtubule synthesis (Bartlett *et al.* 1992). However, for both of these targets, as yet there have been no published data concerning receptor

structure. Furthermore, determination of a pneumocandin pharmacophore would be impractical because of its complex chemical structure.

The structures of the binary complexes of human DHFR with DHF and human DHFR with 5-deazafolate are available (Davies *et al.* 1990). More recently, the crystal structure of the human DHFR, DHF and NADPH ternary complex has been solved (Cody *et al.* 1992). Also the DHFR structures in binary and ternary complexes with DHF, NADPH and various inhibitors have been determined for mouse liver (Stammers *et al.* 1987), chicken liver (Matthews *et al.* 1985), *Escherichia coli* (Bolin *et al.* 1982) and *Lactobacillus casei* (Bolin *et al.* 1982) DHFRs. Although the crystal structure for *P. carinii* DHFR is being solved currently (Stammers *et al.* 1993) there are no three dimensional data available. However, the amino acid sequence has been determined from the DNA sequence (Edman *et al.* 1989a). Hence, with the multitude of structural data available for DHFR structures homologous to *P. carinii* DHFR, CADD techniques can be applied to create a reasonable representation of the active site of this target enzyme.

1.6.6 Selection of a Suitable Lead Compound for the Design of a Selective Inhibitor of *Pneumocystis carinii* DHFR

For the design of an inhibitor against *P. carinii* DHFR with clinical potential two criteria must be satisfied. The potent inhibitory activity against *P. carinii* DHFR must be maximised, and simultaneously, any activity against mammalian DHFR must be kept to a minimum. Examination of the current antifolate and antimicrobial chemotherapy against PCP shows that *P. carinii* is consistently more susceptible to antimalarial (Falloon *et al.* 1991) and antiparasitic (Montgomery *et al.* 1987) agents than antifungal agents. Thus, even though pyrimethamine is less potent and has a lower differential activity towards *P. carinii* DHFR than other DHFR inhibitors (table 1.1), it was selected as an original lead compound because of its excellent antimalarial activity together with some demonstrable *in vivo* antipneumocystis activity (Walzer *et al.* 1992).

Many pyrimethamine analogues have been synthesised and evaluated as potential antineoplastic agents. Unlike classical folate and antifolate structures that are actively transported across the cell membrane, their lipophilic nature allows them to diffuse

across cell membranes. Hence, pyrimethamine analogues have had some clinical significance in the treatment of methotrexate-resistant tumours that have a modified folate transport system (Baker *et al.* 1991). Structure-activity studies on the numerous analogues of pyrimethamine that have been synthesised have shown that the 2,4-diaminopyrimidine pharmacophore is essential for its antifolate activity (Cody & Zakrzewski 1982). Also, the greatest antifolate activity was obtained when the substituent in position 6 of the pyrimidine heterocycle was a small hydrophobic group (Roth *et al.* 1983). Modifying the hydrophobic group to become incorporated in a 6,2'-dimethylene bridge in tricyclic antifolate structures caused a greater increase in inhibitory activity towards mammalian than *Pneumocystis* DHFR (Rosowsky *et al.* 1989). With the pyrimidine heterocycle and substituents so important for DHFR inhibitory activity, it was decided to investigate the improvement of antipneumocystis activity of pyrimethamine by the addition of substituents to the phenyl ring (figure 1.11).

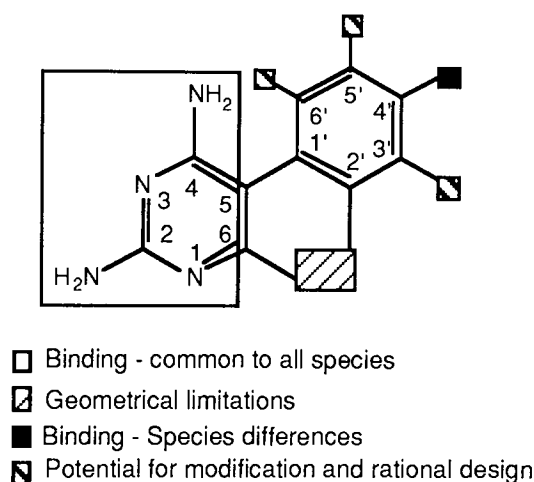
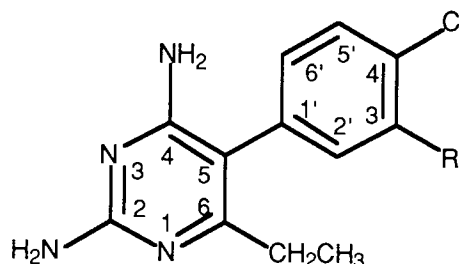


Figure 1.11 Representation of the potential for the design of inhibitors against *P. carinii* DHFR based on pyrimethamine

The presence of the chlorine at position 4' of the 5-phenyl ring has a marked antimalarial effect and is believed to be implicated in binding to the active site (Burchall & Hitchings 1965). Hence, for the design of novel antipneumocystis agents based on pyrimethamine, this substituent was retained. Bliss *et al.* (1987) synthesised pyrimethamine analogues embracing a hydrophobic nitrogen-containing functional group incorporated at position 3' of the 5-phenyl ring. Crucial to their design was the

incorporation of a lipophilic moiety that could be transformed, either metabolically or chemically into a polar diazonium group. These compounds were expected to have a relatively short biological half-life. Also, further metabolism would yield the arylamine, with markedly reduced biological activity (Sansom *et al.* 1989) and rapidly cleared from the body. Thus, it was anticipated that these compounds would not exhibit the chronic toxicity associated with other antifolates with excessively long half-lives (Graziano *et al.* 1992).



R	=	-H, pyrimethamine
R	=	-N ₃ , 3'-azidopyrimethamine, NCI319947
R	=	-N ⁺ =N, 3'-diazoniumpyrimethamine, NCI330463
R	=	-N=N-N(CH ₃) ₂ , NCI330465 (AM90020)
R	=	-N=CH-N(CH ₃) ₂ , NCI372950

Figure 1.12 Structures of pyrimethamine analogues with 3'-nitrogen substituents

The partial successes of clinically significant folate antagonists in the treatment of PCP led Broughton and Queener (1991) to test many analogues from the different classes of DHFR inhibitors against the isolated mammalian and *P. carinii* enzymes. Included in this study were pyrimethamine and the analogues shown in figure 1.12. Pyrimethamine produced IC₅₀ values of 3.65 μ M and 2.3 μ M against *P. carinii* and rat liver DHFR respectively, giving a rat liver/*P. carinii* IC₅₀ activity ratio of 0.63 (table 1.1). Activity ratios of a similar order of magnitude were obtained for the azido and diazonium analogues. Another pyrimethamine analogue with a substituted nitrogen function in the 3' position, 2,4-diamino-6-ethyl-5-(4'-chloro-3'-(3'',3''-dimethyltriazene-1''-yl))phenylpyrimidine (AM90020) produced IC₅₀ values of 2.8 μ M and 18.9 μ M and a selectivity ratio of 6.75. Another pyrimethamine analogue, related to AM90020, 2,4-diamino-6-ethyl-5-(4'-chloro-3'-(3'',3''-dimethylformamidine-1''-yl))phenylpyrimidine gave values of 10.6 μ M, 3.5 μ M and 0.33 respectively. Therefore, it appeared that the

incorporation of the 3'',3''-dimethyltriazene-1''-yl moiety into the pyrimethamine structure adjacent to the chlorine was responsible for a 20 fold increase in the selectivity toward *P. carinii* DHFR. This increase is primarily a result of a reduction in activity toward the rat liver DHFR. Although there appears to be only minor structural differences between the triazenyl and the formamidinyl moieties they were sufficient to cause a considerable change in the biological activities toward the two enzymes. It was hypothesised that unique steric or electronic characteristics of the triazene group were responsible for the significant differences in selectivity of AM90020.

Inhibitor	IC ₅₀ (μM) for DHFR from :-		IC ₅₀ ratio (RL/PC)
	Rat Liver (RL)	<i>P. carinii</i> (PC)	
Trimethoprim	390	12	32.5
Trimetrexate	0.0030	0.042	0.07
Piritrexim	0.0036	0.019	0.19
Pyrimethamine	2.3	3.65	0.63
NCI319947	0.33	1.33	0.25
NCI330463	0.032	0.19	0.17
NCI3304635 (AM90020)	18.9	2.8	6.75
NCI372950	3.5	10.6	0.33

Table 1.1 IC₅₀s against rat liver and *P. carinii* DHFR of clinically used DHFR inhibitors and pyrimethamine analogues (data from Bartlett & Queener 1991)

In conclusion, AM90020 appears to be a very promising lead compound from which improvements in the selectivity and activity against *P. carinii* DHFR, can be achieved by making only minor alterations to the inherent chemical structure. In terms of molecular design, two options for the medicinal chemist are available :-

- (i) adding substituents to other positions in the phenyl ring
- (ii) changing the identity of the terminal triazene groups.

Broughton and Queener (1991) showed that introducing other substituents at the 4' position of the phenyl ring had an adverse effect on the inhibitory activity towards *P. carinii*. There are no data reporting the effect that adding substituents at position 2' has on the biological activity. Therefore, because it appears that the 3'-triazenyl substituent

is responsible for improved selectivity towards *P. carinii* DHFR, it would be valuable to investigate the effect that different terminal 3''-substituents have on the biological activity of AM90020. Furthermore, inherent in the pharmacophore is the phenyltriazene moiety. Studies have shown that the 3'',3''-dimethyltriazene-1''-yl moiety can be metabolised to the corresponding arylamine (Rouzer *et al.* 1993, Spassova & Golovinsky 1985). Hence, a similar metabolism for AM90020 analogues would produce 3'-aminopyrimethamine, which has a lower inhibitory activity and is rapidly excreted (Sansom *et al.* 1989).

In conclusion, AM90020 possesses the desirable biological activity, the potential for a low toxicity profile and a suitable structure from which other analogues can originate for the rational design of a selective inhibitor of *P. carinii* DHFR.

1.7 Rationale and Objectives

The previous sections have highlighted that *P. carinii* is a serious pathogen amongst immunocompromised patients. Due to the high incidence of adverse reactions and relapse within nine months of treatment the antimicrobial compounds that are used currently to treat *P. carinii* infections are inadequate, and there is a need to continue to develop novel antipneumocystis agents.

As discussed in section 1.5 the use of antifolate chemotherapy in the treatment and prophylaxis of PCP would suggest that folate synthesis and metabolism is an ideal target against which more suitable antimicrobial agents could be developed. This is supported by the characterisation of *de novo* folate biosynthesis in *P. carinii* (Kovacs *et al.* 1989) and enzyme inhibitory kinetic studies have shown that there are structural differences in the active sites of mammalian and *P. carinii* DHFRs (Allegra *et al.* 1987a). Structure-activity studies have suggested that the 3'-(3'',3''-dimethyltriazene-1''-yl) substituent of pyrimethamine analogue AM90020 is responsible for the greater selective activity towards *P. carinii* over mammalian DHFR. The objective of this research was to ascertain reasons for this observation, with the aim of developing analogues of AM90020 with a greater selective activity for *P. carinii* DHFR that could be evaluated for the treatment of PCP.

Owing to the DHFR inhibitory profile of AM90020, study of this compound in greater detail is warranted with regard to its steric and electronic characteristics. This requires

elucidation of the x-ray crystallographic structure of AM90020. Accurately determined atomic coordinates could then be used in quantum mechanical calculations to determine the energetically favourable conformations and the electronic distribution. Similar studies on other dihydrofolate reductase inhibitors have suggested that any selective inhibitory activity is a result of conformational changes when bound in the active site (Andrews *et al.* 1986, Welsh 1990). Therefore with structural data for the target enzyme it may be possible to examine the potential interactions of AM90020 and other 3'-(3'',3''-disubstitutedtriazene-1''-yl) analogues of AM90020 can have when bound in the active sites. The human DHFR crystallographic coordinates are available (Davies *et al.* 1990). Unfortunately, as yet, there are no three dimensional data for the *P. carinii* DHFR structure. However, there are enough DHFR structural data available for other species such that homology modeling techniques can be applied to construct a representation of the active site of *P. carinii* DHFR. The structural and electronic data obtained from the quantum mechanical calculations on AM90020 and its analogues can be applied to determine if any quantitative structure-activity relationships (QSAR) exist. The use of CADD and QSAR methods should determine at a molecular level why AM90020 has a greater differential activity towards *P. carinii* over mammalian DHFR. This in turn may lead to rational design and synthesis of other compounds with the desired biological activity.

Chapter 2 Crystallographic Studies on 2,4-Diamino-6-ethyl-5-(4'-chloro-3'-(3'',3''-dimethyltriazen-1-yl))phenylpyrimidine Ethanesulphonate

2.1 Introduction

The first steps in CADD involve obtaining crystallographic atomic coordinates either by experimentation or from searching literature or chemical structure databases such as the Cambridge Crystallographic Database (Allen *et al.* 1991). The crystal structure, as a low energy conformation, also provides an ideal initial geometry for conformational analysis (Denny *et al.* 1992, Sansom *et al.* 1989). Additional modeling techniques, as discussed in chapter 3, have used crystallographic geometries as a template on which to construct analogues and determine parameters for QSAR studies (So & Richards 1992). As crystallographic structural data have become available for target enzymes (Lewis & Moereels 1992, Domenighini *et al.* 1991), it has been possible to dock into the active site of DHFR from solved crystal structures. For example the crystal structures of the antifolate 2,4-diamino-5-(1'-adamantyl)-6-methylpyrimidine (DAMP) (figure 2.1) and its less active analogue where 5-(1'-naphthyl) replaced the adamantyl group (DNMP) (figure 2.1) were docked in the active site of *L. casei* DHFR. Molecular graphic studies were able to reveal that steric clashes between the naphthalene ring in its crystallographic conformation and active site amino acid side chains were responsible for the DNMP inactivity (Cody & Zakrzewski 1982). Similarly, the experimentally determined atomic coordinates of the potent DHFR inhibitor methylbenzoprim (MBP) (figure 2.2) (Meek *et al.* 1987) were used as a starting point to examine its conformational flexibility and then study its interaction in the active site of mammalian DHFR (Denny *et al.* 1992).

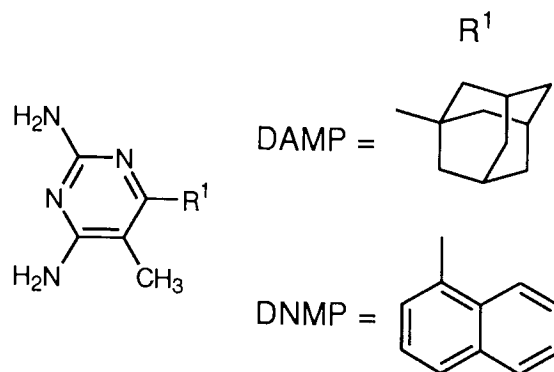


Figure 2.1 Structures of DAMP and DNMP

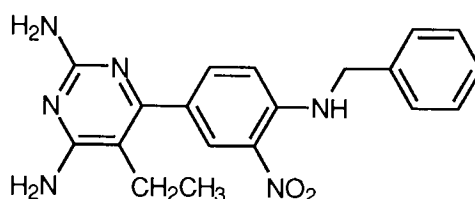


Figure 2.2 The structure of methylbenzoprim

For the development of analogues of AM90020 as antipneumocystis agents, it was considered necessary to obtain accurate structural data for this lead compound. Because antifolates bind to DHFR as the protonated cation, a salt of AM90020 was the appropriate form to study. Considerations to crystal quality lead to the selection of a specimen of the ethanesulphonate salt (AM90021). The atomic coordinates could then be used to construct analogues, obtain structure-activity and conformational data. This, with docking studies, could assist in proposing potential interactions.

2.2 Materials and Methods

2.2.1 Sources of Chemicals

The following reagents were purchased from or supplied by the sources indicated
Crystals of 2,4-diamino-6-ethyl-5-(4'-chloro-3'-(3'',3''-dimethyltriazene-1''-yl))phenyl-pyrimidine ethanesulphonate (AM90021)

Aston Molecules Ltd., Gosta Green, Birmingham, West Midlands

Carbon tetrachloride

Chlorobenzene

Ethanesulphonic acid

Isopropyl alcohol

BDH Chemicals Ltd., Atherstone, Warwickshire

2.2.2 Preparation of AM90021 Crystals from the Free Base AM90020

AM90020 (0.3g) was dissolved in isopropyl alcohol (10ml) at room temperature, to which one equivalent of ethanesulphonic acid was added. AM90021 crystals grew as rosette structures at 5°C. A specimen crystal; measuring 0.7 x 0.5 x 0.3 mm, possessing shiny surfaces, but no defined shape, was mounted onto a glass filament with epoxy adhesive. The filament was secured to the X-ray diffractometer goniometer head by the hot wax method.

2.2.3 Experimental Determination of AM90021 Crystal Density

This was measured by flotation in carbon tetrachloride and chlorobenzene.

2.2.4 Crystallographic Data

AM90021, $C_{14}H_{19}N_7Cl^+.C_2H_5O_3S^-.C_3H_8O$, MW = 490.255, triclinic, $a = 8.438(1)$ Å, $b = 9.933(1)$ Å, $c = 16.969(2)$ Å, $\alpha = 99.86(1)^\circ$, $\beta = 92.92(1)^\circ$, $\gamma = 113.92(2)^\circ$, vol = 1267.41 \AA^3 , $Z = 2$, $D_x = 1.276$, $D_m = 1.27$, $F(000) = 452$, $\mu = 2.20 \text{ cm}^{-1}$, temperature $21(2)^\circ\text{C}$, space group = $P\bar{1}$.

2.2.5 Data Collection and Structural Analysis of AM90021

Data were collected on an Enraf-Nonius CAD4 diffractometer with monochromated Mo-K α radiation, $\lambda = 0.71069$ Å. Unit cell dimensions were obtained from least squares analysis of setting angles of 17 unique reflections with $1 \leq \theta \leq 25$. Intensity data were collected by the ω -2 θ scan technique with ω scan range of $1.35 + 0.35 \tan \theta$ and a ω scan speed of 0.83 to 3.33 deg min⁻¹. Each hour three intensity monitor reflections were measured and in every 100 experimental reflections three orientation monitor reflections were located. Significant alterations to the orientation did not occur, but by the end of the data collection the intensity standards had reduced by

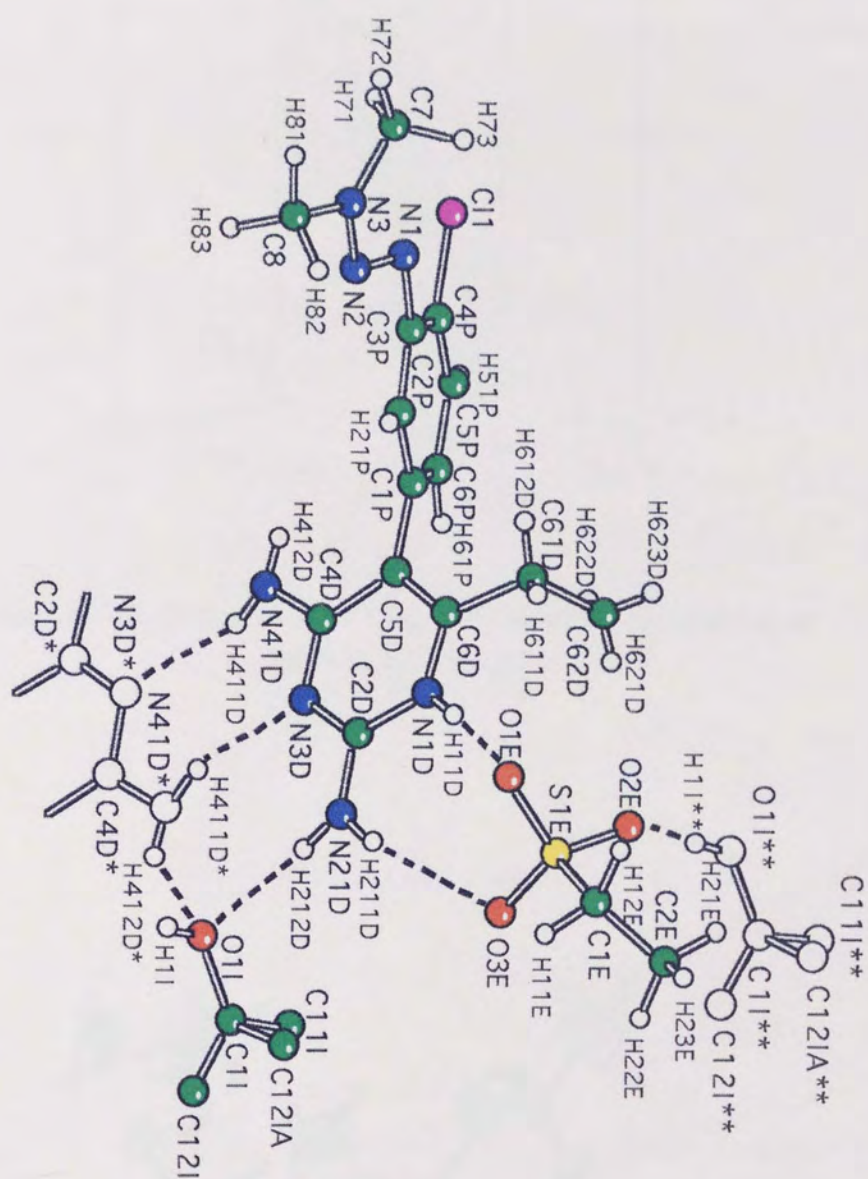
approximately 30%, indicating some decomposition had occurred. This was partially compensated for by using a linear and isotropic decay estimate. A total of 4752 reflections were collected for $0 \leq h \leq 10$, $-11 \leq k \leq 11$ and $-20 \leq l \leq 20$ in the range of $2 < \theta < 25^\circ$ and were merged to give 4420 unique reflections ($R_{\text{int}} = 0.0188$) of which 2991 were deemed observed with $F_o > 3\sigma(F_o)$. Estimated standard deviations (σ) were based on counting statistics and an allowance of $0.02 F_o$ for the minimum expected instability. All non-hydrogen atoms were located by direct methods using the SHELX program (Sheldrick 1976). Disorder in the isopropanol molecule of crystallisation was observed and it was assumed that an estimated 20% adopted an alternative, less favourable, methyl group position. With the exception of the hydroxyl hydrogen, this prevented the location of the isopropanol hydrogens. The two terminal triazene methyl groups were treated as rotatable functions with ideal geometry and their hydrogens calculated in fixed positions using the AFIX command, which was also used to calculate the positions of the two methylene and two of the three methyl hydrogens of the Ethanesulphonate anion. The remaining hydrogen atoms were located from difference electron density maps after least squares refinement of non-hydrogen atom positions. The S-O bond lengths did not refine to acceptable geometry, so they were restrained to $1.50 \pm 0.02 \text{ \AA}$ using the DFIX command. Further least squares refinement of all the determined atomic co-ordinates, the anisotropic thermal parameters for non-hydrogen atoms and the isotropic temperature factors for the hydrogen atoms was carried out.

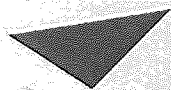
In the weighting scheme $\omega = k / [\sigma(F_o) + g(F_o)^2]$, for the function $\sum \omega (|F_o| - |F_c|)^2$, the parameters converged to give $k = 2.4861$ and g refined to 0.000535 at the discrepancy indices of $R = 0.0631$ and $R\omega = 0.0698$. When least squares refinement terminated, the shifts in positional parameters of individual atoms were < 0.01 of the estimated standard deviation and no parameter shift exceeded 0.25 e.s.d.

2.2.6 Determination of the Planarity Between the Pyrimidine Aromatic Rings in the Unit Cell of AM90021

The final positional parameters of AM90021 and its inverted transformation were written to a .cssr formatted file and visualised using the molecular modeling package NEMESIS (Oxford Molecular Ltd. 1991). The package then used atoms N1D, C2D,

N3D, C4D, C5D and C6D to calculate centroids and the best fit plane through the original and the inverted diaminopyrimidine rings. The minimum distance from the centroids to the best fit plane and the centroid separation distance were calculated. From this, using simple fundamental trigonometry the angle, defined as the 3 - x, 1 - y, 1 - z transformed pyrimidine centroid, the original pyrimidine centroid and the nearest point on the best fit plane through the pyrimidine plane to the transformed centroid, was determined.

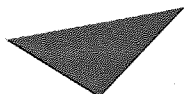




Aston University

Content has been removed for copyright reasons

Figure 2.4 PLUTO (Motherwell & Clegg 1978) drawing of two molecules of AM90021 in the unit cell



Aston University

Content has been removed for copyright reasons

Figure 2.5 ORTEP (Johnson 1976) drawing of a molecule of AM90021

2.3 Results and Discussion

The final positional parameters for non-hydrogen atoms are given in table 2.4 with their respective anisotropic and isotropic thermal parameters in table 2.5. For hydrogen atoms the final positional and isotropic thermal parameters are given in table 2.6. The experimentally determined bond lengths and bond angles are presented in tables 2.7. and 2.8 respectively, and important intermolecular contacts in table 2.1. Observed and calculated structure factors are given in the Appendix. The disorder in the isopropyl alcohol molecule of crystallisation refined in the two positions as described.

The narrow range of C-N bond lengths (1.324(5) - 1.358(5) Å) indicates that the p-electrons from every nitrogen atom in the 2,4-diaminopyrimidine heterocycle participate in the delocalised π system. The N1-N2 and N2-N3 triazene bond lengths, 1.269(5) Å and 1.324(5) Å respectively, differ significantly from representative 1.24 Å N-N and 1.47 Å N=N bond lengths (Sandorfy 1970) and suggest these bonds possessed both single and double bond characteristics as a result of resonance as proposed by Akhtar *et al.* (1968) (figure 2.6).



Content has been removed for copyright reasons

Figure 2.6 Resonance stabilisation of the triazene group as proposed by Akhtar *et al.* (1968)

Studies on other molecules containing the diaminopyrimidine nuclei have shown association in the crystalline state via a dimeric N-H ---- N interaction (Schwalbe & Williams 1982). Protonated 2,4-diaminopyrimidine antifolates have formed these base pair centres or pseudocentres of inversions with the 4-amino groups linking N-3 of an inverted pyrimidine nucleus and vice versa (Schwalbe & Williams 1987). This was also observed in AM90021. Consistent with other diaminopyrimidine crystal structures, the original and transformed aromatic nuclei are not exactly coplanar (Schwalbe & Cody 1983). This is emphasised by the angle between the original and 3 - x, 1 - y, 1 - z transformed centroids and the point closest to the centroid on the best-fit plane, (see section 2.2.6) calculated to be 5.3°. For the two pyrimidine nuclei to be a co-planar system, this angle would be 0°. This transformation also orientates the 4-amino group

sufficiently close to the isopropyl alcohol molecule of crystallisation to allow the possibility of an N41D*-H412D* ---- O1I as well as the N21D*-H212D* ---- O1I interaction.

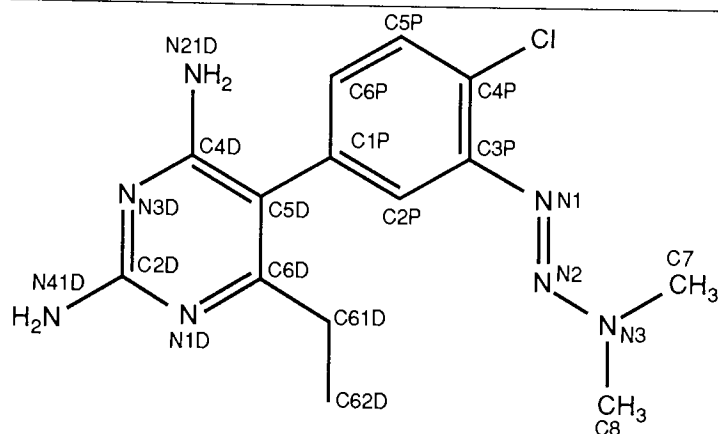
Contact	Contact Distance (Å)	Contact Angle (°)
N1D-H11D ---- O1E	2.741	172.8
N21D-H211D ---- O3E	3.069	155.1
N21D-H212D ---- O1I	2.965	174.5
N41D-H411D ---- N3D*	3.039	161.6
N41D*-H412D* ---- O1I	2.922	138.9
O1I**-H1I** ---- O2E	2.693	175.5

Table 2.1 Intermolecular contacts in the crystal structure of AM90021.

Symmetry transformations :- * 3 - x, 1 - y, 1 - z; and ** 1 + x, y, z.

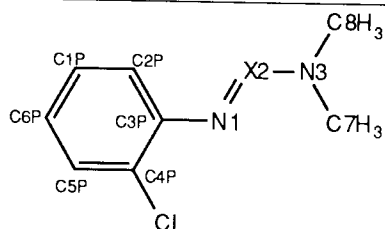
A similar dimeric interaction has been observed involving the 2-amino group and N1D in crystallographic structures of non-protonated diaminopyrimidines (Griffin *et al.* 1989). However protonation of the title compound prevents this interaction, but allows N1D of the pyrimidine ring and the 2-amino group to form contacts with two of the three ethanesulphonyl oxygens. The remaining ethanesulphonyl oxygen forms a very close O1I**-H1I** ---- O2E contact. The contact lengths of approximately 2.7 Å and 3.0 Å agree very well with mean values determined by Taylor & Kennard (1984), who went on to suggest that the non-linearity of many contacts was a consequence of spatial factors. They concluded that hydrogen bond contact angles as low as 90° could exist. Therefore, it is possible that the contact N41D*-H412D* ---- O1I is a true hydrogen bond.

The torsions listed in table 2.2 show that the 6-ethyl and the 5-substituted phenyl substituents, consistent with the other protonated crystallographic structures of pyrimethamine and its analogues, are not co-planar with the 2,4-diaminopyrimidine system.



Torsion Angle	Torsion Angle in AM90021 (°)	Pyrimethamine (RMS)*
C5D-C6D-C61D-C62D	-94.4	±82.6
C4D-C5D-C1P-C2P	107.1	±103.5

Table 2.2 Experimentally determined bond torsions for AM90021 and RMS bond torsions calculated from other pyrimethamine analogues. *Structures used:- **1** pyrimethamine acetate (Hill 1982), **2** pyrimethamine hydrochloride (Griffin *et al.* 1989), **3** pyrimethamine salicylate (two independent molecules) (Hill & Schwalbe, unpublished data), **4** 2,4-diamino-5-(3'-amino-4'-chloro)phenyl-6-ethylpyrimidine (Bryant 1984), **5** 2,4-diamino-5-(3'-nitro-4'-N-methyl-benzylamino)phenyl-6-ethylpyrimidine ethanesulphonate (Griffin *et al.* 1989), **6** 2,4-diamino-5-(3'-azido-4'-chloro)phenyl-6-ethylpyrimidine ethanesulphonate (Bryant *et al.* 1986).



X2 = N, phenyltriazenes

X2 = CH, phenylformamidine

Bond	Bond Distance for AM90021(Å)	Triazene (RMS) (Å)	Formamidine (RMS) (Å)
C1P-N1	1.412(5)	1.421(7)	1.396(17)
N1-X2	1.269(5)	1.271(6)	1.290(11)
X2-N3	1.324(5)	1.321(9)	1.337(11)
N3-C7	1.425(8)	1.447(12)	1.453(17)
N3-C8	1.429(7)	1.458(11)	1.466(13)

Table 2.3a Triazene bond lengths for AM90021 and RMS bond lengths from phenyltriazenes and phenylformamidine crystal structures with estimated standard deviations (AM90021) and standard deviations in parentheses.

Bond Angle	Bond Angle for AM90021 (°)	Triazene RMS (°)	Formamidine RMS (°)
C2P-C3P-N1	125.0(4)	124.7(5)	124.0(7)
C4P-C3P-N1	118.2(4)	116.6(13)	117.2(8)
C3P-N1-X2	112.0(4)	112.0(4)	117.4(17)
N1-X2-N3	114.8(4)	114.4(7)	122.9(7)
X2-N3-C7	120.8(5)	122.9(11)	120.9(17)
X2-N3-C8	117.4(5)	115.5(14)	121.4(14)
C7-N3-C8	121.8(5)	121.0(13)	117.6(23)

Table 2.3b Triazene bond angles for AM90021 and RMS bond angles from phenyltriazenes and phenylformamidine crystal structures with e.s.d. (AM90021) and s.d. in parentheses

Bond Torsion	Bond Torsion for AM90021 (°)	Triazene RMS (°)	Formamidine RMS (°)
C2P-C3P-N1-X2	-22.4	±15.0	±32.5
C4P-C3P-N1-X2	159.4	±167.6	±154.0
C3P-N1-X2-N3	176.6	±178.7	±176.6
N1-X2-N3-C7	-0.8	±2.8	±5.6
N1-X2-N3-C8	177.8	±176.7	±174.5

Table 2.3c Bond torsions for AM90021 and RMS bond torsions from phenyltriazenes and phenylformamidine crystal structures

Triazene RMS values were calculated from: **1** *para*-nitrodiazoaminobenzene (Kondrashev 1974); **2** 2-(3,3-dimethyltriazeno)phenyl-1-carboxamide (Edwards *et al.* 1977); **3** (2,6-*cis*-dimethylpiperidyl)diazobenzene (Lunazzi *et al.* 1978); **4** 1,6-bis(*para*-chlorophenyl)-3,3-diacetylhexaaza-1,5-diene (Mackay *et al.* 1982); **5** 3-methyl-1-*para*-tolyltriazenes (Randall & Schwalbe 1984); **6** *para*-(3,3-dimethyltriazeno)benzonitrile (Fronczek *et al.* 1988); **7** methyl-4-(3-ethyl-3-hydroxymethyltriazen-1-yl)benzoate (Chopra & Schwalbe 1991); **8** ethyl-4-(3-methyl-3-hydroxymethyltriazen-1-yl)benzoate (Simmonds *et al.* 1988).

Formamidine RMS values were calculated from: **1** *N,N*-dimethyl-*N'*-(6-chloropyridazinyl)-3-formamidine (Leban, 1975); **2** *N'*-(4-chloro-*ortho*-tolyl)-*N,N*-dimethylformamidine (Gifkins & Jacobson, 1980); **3** N1-(hexamethylene-C1-C6)-N2-*para*-nitrophenylformamidine (Krajewski *et al.* 1981); **4** *N,N*-dimethyl-*N'*-*para*-nitrophenylformamidine (Ciszak *et al.* 1989); **5** trans-1,2-bis(2-*N,N*-dimethylaminemethylenamino-4-trifluoromethylphenylthio)ethane (Chen, 1990)

As previously mentioned, steric interactions between the triazene and formamidine groups and DHFR may be responsible for the notable differences in enzyme inhibitory activities. Tables 2.3a, 2.3b and 2.3c list the AM90021 bond distances, bond angles and torsion angles and the equivalent root mean square (RMS) geometric parameters for a range of phenyltriazenes and phenylformamidines taken from the Cambridge Structure Database (Allen *et al.* 1991) (see legend to table). For AM90021 SHELX did not observe any contacts involving the triazene atoms. Thus it appears the triazene is free to adopt the determined conformation and has not been forced to by crystal packing arrangements. Mean values have not been quoted because it is more appropriate to use RMS values with negative torsion angles. These tables show that the experimentally determined triazene structural features of AM90021 are consistent with those of other solved phenyltriazene crystal structures. Large thermal motion parameters for the two terminal triazene carbons, as listed in table 2.5, account for the N3-C7 and N3-C8 bond distances being considerably shorter. This motion adds weight to the suggestion that the triazene conformation is not a result of space constrictions. Generally, the RMS bond distances are very similar, with a 0.025 Å longer C1P-N1 phenyltriazene bond distance being the largest discrepancy.

Interference by the triazene and formamidine moieties with the aromatic ring is responsible for the large C2P-C3P-N1 bond angles; it is often reduced by the rotation of the C2P-C3P-N1-X2 torsion angle away from planarity. In the AM90021 crystal structure the magnitude of the aforementioned bond angle is unexpectedly unaffected by the presence of the “bulky” *ortho*-chlorine substituent on the other adjacent ring position. Examination of recommended van der Waals radii (Bondi 1964) for aryl chlorine (1.77 Å) and doubly bonded nitrogen (1.55 Å) and an interatomic Cl --- N1 separation distance of 2.933 Å clearly illustrates the necessity of some structural compromise to allow this overlap. The C3P-N1-X2 and N1-X2-N3 RMS bond angles in the phenyltriazene, 112.0(4)° and 114.4(7)° respectively, are considerably more compressed than the equivalent RMS phenylformamidine angles (117.4(17)° and 122.9(7)°). It is believed that this phenomenon is caused by the presence in the triazene of additional lone pair electrons on the central nitrogen which, coupled with the electron rich environment of the triazene moiety, decreases these bond angles from normal trigonal geometry to minimise the repulsive intramolecular electrostatic forces. These

lone pair electrons are not present in the formamidine, hence constricted bond angles are not necessary.

Summation of the X2-N3-C7, X2-N3-C8 and the C7-N3-C8 bond angle to approximately 360° and the scalar values for the N1-X2-N3-C7 and the N1-X2-N3-C8 bond torsions to approximately 180° strongly suggests that atoms X2, N3, C7 and C8 form a planar system. However, because the torsion angles are not 0° and 180° , N1 is not co-planar with the terminal triazene methyl carbons. It is also observed that there are significant differences between the N2-N3-C8 RMS phenyltriazene and phenylformamidine bond angles, 115.5° and 121.4° respectively. This may be a result of a knock-on effect from the strained triazene bond angle causing C7 to orientate closer to N1 in the triazene than formamidine; RMS N1 ---- C7 separation distances were 2.594(70) Å and 2.798(45) Å. A larger triazene X2-N3-C7 RMS bond angle strongly suggests interatomic N1 ---- C7 repulsion in the phenyltriazene, which, in order to maintain a suitable C7 ---- C8 separation, causes greater straining of the X2-N3-C8 angle.

The crystallographic determination of the structure of AM90021 has provided a suitable origin for the molecular modelling of analogues of AM90020 (free base of AM90021) because Hitchings & Burchall (1965) have shown that DHFR inhibitors that bind in the active site are also protonated at N1D. The geometric parameters determined were very similar to other solved pyrimethamine and phenyltriazene structures. Comparison of the triazene parameters to equivalent phenylformamidine parameters shows that there are major structural differences that can account for the considerable variation in DHFR inhibitory activity in AM90020 and NCI372950.

Atom	X / A	Y / B	Z / C	Site Occupancy Factor *
N1D	9968(4)	1834(4)	5219(2)	
C2D	11693(5)	2561(4)	5517(2)	
N21D	12196(5)	2325(4)	6213(2)	
N3D	12838(5)	3492(4)	5135(2)	
C4D	12256(5)	3658(4)	4426(2)	
N41D	13450(5)	4540(4)	4042(2)	
C5D	10451(5)	2928(4)	4089(2)	
C6D	9329(5)	2018(4)	4520(2)	
C61D	7382(5)	1247(5)	4314(3)	
C62D	6527(6)	2172(6)	4754(3)	
C1P	9875(5)	3139(4)	3293(2)	
C2P	9379(5)	2009(4)	2619(2)	
C3P	8855(5)	2151(4)	1866(2)	
C4P	8854(6)	3528(5)	1808(3)	
C5P	9339(7)	4686(5)	2464(3)	
C6P	9862(6)	4515(5)	3211(3)	
N1	8282(6)	997(4)	1163(2)	
N2	8835(6)	-1(4)	1217(2)	
N3	8263(8)	-1156(5)	593(2)	
C7	7097(12)	-1235(7)	-69(4)	
C8	8890(11)	-2292(7)	605(4)	
CL1	8251(2)	3802(2)	870(1)	
S1E	7806(2)	533(1)	6875(1)	
O1E	8075(5)	-187(3)	6104(2)	
O2E	6622(5)	1211(4)	6746(2)	
O3E	9407(5)	1508(4)	7356(2)	
C1E	6702(8)	-1352(6)	7579(4)	
C2E	5862(12)	597(10)	8016(5)	
O1I	15734(4)	3533(3)	7139(2)	
C1I	15672(12)	3932(9)	7961(4)	
C11I	15094(16)	5146(10)	8116(5)	
C12I	16592(16)	3687(13)	8553(5)	0.7940(44)
C12AI	14364(25)	2658(20)	8236(13)	0.2060(44)

Table 2.4 Positional parameters (fractional coordinates) ($\times 10^4$) of non-hydrogen atoms with standard deviations in parentheses.* Taken as unity unless otherwise specified

ATOM	U11	U22	U33	U23	U13	U12	Uiso
N1D	361(17)	443(16)	520(18)	47(14)	56(15)	71(14)	481.3
C2D	408(21)	431(18)	518(20)	-9(16)	17(17)	178(17)	467.8
N21D	481(20)	724(21)	614(21)	253(17)	34(17)	164(17)	623.1
N3D	338(16)	461(16)	516(17)	47(14)	48(14)	125(14)	460.4
C4D	363(19)	391(17)	460(19)	-6(15)	18(16)	138(16)	425.5
N41D	415(18)	574(19)	511(2)	113(15)	-13(15)	121(16)	532.5
C5D	386(20)	406(17)	511(19)	-37(15)	-10(17)	160(16)	456.7
C6D	394(20)	372(17)	496(20)	-43(15)	11(17)	136(16)	448.5
C61D	367(21)	497(20)	655(23)	8(19)	-14(19)	80(18)	558.3
C62D	461(24)	876(28)	837(28)	-37(25)	76(23)	264(22)	759.3
C1P	323(19)	454(18)	591(20)	42(16)	-7(16)	153(16)	471.0
C2P	477(21)	388(18)	584(22)	50(17)	15(18)	179(19)	493.0
C3P	476(21)	494(19)	519(20)	60(16)	-34(17)	168(17)	518.3
C4P	619(24)	631(21)	621(23)	126(19)	-35(20)	291(20)	619.6
C5P	697(26)	538(21)	856(27)	119(21)	-35(23)	309(20)	689.5
C6P	546(23)	530(21)	682(24)	20(19)	-73(20)	269(19)	591.8
CL1	1313(13)	956(9)	767(8)	254(7)	-108(8)	634(9)	959.7
N1	868(25)	623(20)	597(20)	63(17)	-56(19)	303(19)	714.1
N2	986(27)	668(21)	552(20)	-7(17)	-1(20)	335(21)	761.4
N3	1511(34)	755(23)	596(22)	-89(24)	-86(24)	452(25)	1001
C7	1924(41)	1012(32)	766(29)	-170(27)	-496(32)	504(33)	1345
C8	1859(40)	795(28)	999(32)	-78(26)	-13(33)	737(29)	1199
S1E	619(7)	516(6)	771(7)	260(5)	184(6)	256(5)	609.8
O1E	1015(23)	599(17)	1009(22)	154(16)	470(19)	312(17)	872.7
O2E	1184(25)	940(21)	1109(22)	415(19)	305(20)	776(20)	935.9
O3E	885(24)	897(22)	971(23)	311(19)	-133(20)	-4(20)	1039
C1E	1050(32)	950(28)	1430(34)	729(26)	676(28)	569(25)	1012
C2E	1741(40)	1864(39)	1450(36)	1072(32)	731(34)	892(35)	1526
O1I	680(19)	689(17)	684(17)	166(14)	75(15)	393(15)	647.1
C1I	1960(40)	1695(37)	655(28)	92(29)	120(31)	1245(32)	1300
C11I	3054(45)	1511(37)	1362(39)	-239(34)	301(40)	1459(36)	1880
C12I	2165(44)	2052(42)	697(34)	185(35)	15(37)	1339(39)	1508
C12AI							74(4)

Table 2.5 Anisotropic temperature factors, and isotropic temperature factors where shown, with standard deviations in parentheses for non-hydrogen atoms ($\times 10^4$).

Atom	X / A	Y / B	Z / C	Uiso
H11D	9372(34)	1284(29)	5451(18)	328(44)
H211D	11610(36)	1930(32)	6528(21)	511(46)
H212D	13190(41)	2713(37)	6499(27)	996(48)
H411D	13131(35)	4788(30)	3636(21)	451(45)
H412D	14344(36)	5015(32)	4358(21)	512(46)
H611D	7202(36)	1109(32)	3793(21)	538(45)
H612D	7002(37)	362(34)	4466(21)	628(46)
H621D	6612(40)	2241(37)	5231(25)	837(47)
H622D	7143(39)	3134(36)	4613(24)	835(47)
H623D	5405(40)	1691(36)	4640(25)	846(47)
H21P	9340(34)	1294(30)	2678(19)	335(44)
H51P	9437(39)	5531(36)	2397(24)	811(47)
H61P	10278(36)	5291(32)	3732(22)	654(45)
H71	7695	-261	-333	
H72	6786	-2246	-521	
H73	5913	-1252	152	
H81	8322	-3166	66	
H82	8507	-2767	1128	
H83	10298	-1808	640	
H11E	5718	-1352	7579	
H12E	7643	-1890	6910	
H21E	4962	-152	7831	
H22E	6808	208	8511	
H23E	5140	-1741	8184	1836(49)
H1I	16001	2622	7006	1217(48)

Table 2.6 Positional parameters of hydrogen atoms and isotropic temperature factors
($\times 10^4$) with standard deviations in parentheses

Bond	Bond Length (Å)	Bond	Bond Length (Å)
N1D-C2D	1.358(5)	N1D-H11D	0.760(27)
C2D-N21D	1.324(5)	N21D-H211D	0.791(32)
N21D-H212D	0.848(34)	C2D-N3D	1.324(5)
N3D-C4D	1.338(5)	C4D-N41D	1.324(5)
N41D-H411D	0.838(38)	N41D-N412D	0.814(28)
C4D-C5D	1.430(5)	C5D-C6D	1.360(5)
C5D-C1P	1.490(5)	C6D-C61D	1.496(5)
C6D-N1D	1.349(5)	C61D-H611D	0.867(34)
C61D-H612D	0.894(31)	C61D-C62D	1.509(7)
C62D-H621D	0.799(40)	C62D-H622D	0.963(33)
C62D-H623D	0.863(29)	C1P-C2P	1.368(5)
C2P-H21P	0.721(28)	C2P-C3P	1.380(5)
C3P-N1	1.412(5)	C3P-C4P	1.389(6)
C4P-C1I	1.745(4)	C4P-C5P	1.365(6)
C5P-H51P	0.837(35)	C5P-C6P	1.381(6)
C6P-H61P	1.008(32)	C6P-C1P	1.402(5)
N1-N2	1.269(5)	N2-N3	1.324(5)
N3-C7	1.425(8)	N3-C8	1.429(7)
S1E-O1E	1.456(3)	S1E-O2E	1.440(4)
S1E-O3E	1.409(4)	S1E-C1E	1.774(5)
C1E-C2E	1.437(9)	C2E-H23E	1.145(8)
O1I-H1I	0.973(3)	O1I-C1I	1.394(6)
C1I-C11I	1.464(9)	C1I-C12I	1.470(15)
C1I-C12AI	1.352(9)		

Table 2.7 Bond lengths (Å) with standard deviations in parentheses

Bond Angle	Angle (°)	Bond Angle	Angle (°)
C2D-N1D-C6D	122.0(3)	C2D-N1-H11D	117.5(22)
C6D-N1D-H11D	120.5(22)	N1D-C2D-N3D	121.5(4)
N1D-C2D-N21D	117.7(4)	N3D-C2D-N21D	121.5(4)
C2D-N21D-H211D	128.6(23)	C2D-N21D-H212D	130.7(29)
H211D-N21D-H212D	99.4(35)	C2D-N3D-C4D	117.9(3)
N3D-C4D-C5D	122.7(3)	N3D-C4D-N41D	116.3(3)
C5D-C4D-N41D	121.0(4)	C4D-N41D-H411D	119.5(19)
C4D-N41D-H412D	108.1(23)	H411D-N41D-H412D	127.3(30)
C4D-C5D-C6D	116.7(3)	C4D-C5D-C1P	120.5(3)
C6D-C5D-C1P	122.8(3)	N1D-C6D-C5D	119.1(3)
N1D-C6D-C61D	115.2(3)	C5D-C6D-C61D	125.6(4)
C6D-C61D-C62D	111.1(3)	C6D-C61D-H611D	104.4(19)
C6D-C61D-H612D	108.4(20)	H611D-C61D-H612D	110.3(30)
C62D-C61D-H611D	113.9(23)	C62D-C61D-H612D	108.5(23)
C61D-C62D-H621D	114.4(29)	C61D-C62D-H622D	101.8(23)
C61D-C62D-H623D	109.4(25)	H621D-C62D-H622D	111.6(34)
H621D-C62D-H623D	99.0(36)	H622D-C62D-H623D	121.1(35)
C2P-C1P-C5D	121.0(3)	C6P-C1P-C5D	120.7(3)
C2P-C1P-C6P	118.2(4)	C1P-C2P-C3P	123.4(4)
C1P-C2P-H21P	116.6(25)	C3P-C2P-H21P	119.8(25)
C2P-C3P-C4P	116.8(4)	C2P-C3P-N1	125.0(4)
C4P-C3P-N1	118.2(4)	C3P-C4P-C5P	121.7(4)
C3P-C4P-C11	119.3(3)	C5P-C4P-C11	119.0(3)
C4P-C5P-C6P	120.3(4)	C4P-C5P-C51P	119.0(28)
C6P-C5P-C51P	120.3(27)	C5P-C6P-C1P	119.6(4)
C5P-C6P-H61P	127.5(20)	C1P-C6P-H61P	112.9(20)
C3P-N1-N2	112.0(4)	N1-N2-N3	114.4(4)
N2-N3-C7	120.8(5)	N2-N3-C8	117.4(5)
C7-N3-C8	121.8(5)	O1E-S1E-O2E	109.9(2)
O1E-S1E-O3E	111.6(3)	O2E-S1E-O3E	114.9(3)
C1E-S1E-O1E	114.3(2)	C1E-S1E-O2E	106.7(3)
C1E-S1E-O3E	108.8(3)	S1E-C1E-C2E	113.8(5)
C1E-C2E-H23E	101.6(7)	H11-O11-C11	112.4(4)
O11-C11-C111	110.1(6)	O11-C11-C121	110.2(11)
O11-C11-C12AI	124.7(7)	C111-C11-C121	105.9(11)
C111-C11-C12AI	121.0(7)	C121-C11-C12AI	75.9(10)

Table 2.8 Bond angles with standard deviations in parentheses

Chapter 3 Computational Chemical Calculations on 3,3-Dimethyl-1-(2'-chloro)phenyltriazene

3.1 Introduction

CADD is based upon the visualisation or determination of how a potential inhibitor is capable of interacting with specific residues located in the receptor site. Calculating the energy of interaction is based on the electronic properties and the three dimensional configuration of the ligand with respect to receptor site residues. As a consequence there is considerable importance associated with the determination, either experimentally or theoretically, of the electronic distribution. Experimental determination involves crystallographic procedures. This is often time consuming and the results may have limited significance when applied to biological systems at 37°C. Therefore, theoretical methods (empirical, semi-empirical and *ab initio* programs) are used to calculate the electronic and steric properties on isolated molecules through to complex molecular systems.

3.1.1 *Ab Initio* Quantum Mechanics Programs

The starting point of quantum mechanical chemistry is the Schroedinger equation (3.1)

$$H \psi = E \psi \quad \text{Equation 3.1}$$

where E is the energy of the molecule, ψ is the wave function and H the Hamiltonian. The Hartree-Fock equation (3.2)

$$H \phi = \epsilon \phi \quad \text{Equation 3.2}$$

provides solutions for each individual electron in a molecule (ϕ), where ϵ is the orbital energy. This equation may be solved for atoms or very small molecules. Very good approximations for E and ψ can be obtained from the summations of the individual values of ϵ and ϕ (ϵ_i and ϕ_i), from which steric or electronic functions can be calculated respectively (Richards 1990). These non-parameterised solutions are termed *ab initio*. There are simplifying assumptions in *ab initio* theory including the definition of the atomic orbitals (basis set) in the usual Gaussian methods as a number of probability functions. The coefficients of these functions are selected to give as good a fit as possible to the corresponding orbital. For each additional orbital there is a fourth power

increase in the number of evaluable terms in the calculation. Therefore, although *ab initio* programs are the most sophisticated and accurate, they are computationally demanding. Thus, before they can be used routinely for geometry optimisations and electronic charge distribution calculations further advances in computational technology are required.

3.1.2 Semi-empirical Quantum Mechanics Programs

Semi-empirical calculations were designed originally as an intermediate to molecular mechanics and *ab initio* techniques. Like molecular mechanics they use experimentally determined parameters, but because they consider the only valence electrons in the calculation, they are also quantum mechanical in nature (Stewart 1992). However, with the extensive use of approximations, the need to evaluate the large number of terms in the *ab initio* calculations is avoided, thereby reducing considerably the computational time. Although early methods were parametrised to consider aromatic delocalisation and conjugation, an inability to reproduce hydrogen bonding was fatal for studying biological systems (Thiel 1988). Dewar *et al.* (1985) developed additional parameters for the Austin Model 1 (AM1) that corrected for the excessive repulsion at van der Waals' distances in the modified neglect of diatomic overlap (MNDO) basis set (Dewar & Thiel 1977); hydrogen bonding could be calculated more accurately. The next package to be developed, PM3 (Stewart 1989), introduced more reference data for the optimisation routines. However, because of inconsistent results this parameter set has not been generally accepted (Dewar *et al.* 1990).

Semi-empirical methods lack the extreme accuracy of *ab initio* calculations and the speed of molecular mechanics. However, the excessive computational demands of *ab initio* calculations often limit them to simplified molecular systems. The ability to handle more realistic representations of complex situations makes the semi-empirical approximations and parameters useful sources of quantum mechanical data with chemically useful accuracy (Stewart 1992).

3.1.3 Molecular Mechanics Programs

This approach is based on the concept that the energy of the molecule can be categorised into the summation of discrete individual components, collectively known

as the forcefield, as shown in equation 3.3.

$$E_{\text{molecule}} = E_{\text{steric}} + E_{\text{bond length}} + E_{\text{bond angle}} + E_{\text{bond torsion}} + E_{\text{polar}} \quad \text{Equation 3.3}$$

Quantum mechanical aspects are not considered. Therefore the simplicity of this concept makes it ideal for situations where computational efforts can be expended on other procedures, such as macromolecular study or conformational analysis (Brooks *et al.* 1983). However, this simplified view of a molecule results in a very poor capability to handle non-ground state geometries (Stewart 1992), polar interactions, and conjugation (Loftus 1990). Thus, molecular mechanics and semi-empirical molecular orbital programs will remain complementary techniques.

Several studies on various classes of compounds have tried to evaluate and compare these various packages. Studies in substituted aromatic amines (Smith *et al.* 1992, Bonatti *et al.* 1992) have shown that the semiempirical methods, especially using the AM1 parameters, were superior to molecular mechanics when compared to *ab initio* data. In theoretical acetyltriene decomposition studies, AM1 results provided a very good estimate for 3-21G basis set optimised bond lengths and proton affinity calculations (Ozment *et al.* 1991). However the authors acknowledged the limitations of the AM1 parameters for predicting triene bond angles.

Ultimately the choice of method will depend on the size of the molecular system and the computer memory capacity and speed available. As the system becomes more complex, because of the vast amount of time required to complete the calculation, it is often impractical to use more rigorous programs. Therefore rational selection of the most appropriate method and programs based on time and accuracy is essential for any further studies concerning the similarities between molecules exhibiting potent or weak biological activities or the ligand-receptor electrostatic interactions in the active site.

3.2 Materials and Methods

3.2.1 Molecular Modeling Packages

ChemX (Chemical Design Ltd., Chipping Norton 1992) was mounted on a VAX/VMS 8650 mainframe and was used as an interface to the semi-empirical quantum mechanical

program MOPAC (Stewart 1990a), from which the MNDO, AM1 and PM3 parameter sets were used.

Access to the COSMIC molecular mechanics forcefield was via the package NEMESIS (Oxford Molecular Ltd., Oxford 1991), mounted on an Apple Macintosh IIfx.

GAMESS and ASP (Automated Similarity Package, Oxford Molecular Ltd., Oxford 1991) were mounted on a Sun SPARCstation and used for the *ab initio* and molecular similarity calculations respectively.

3.2.2 Computational Chemistry Calculations on 3,3-Dimethyl-1-(2'-chloro)phenyltriazene and 3,3-Dimethyl-1-(2'-chloro)phenylformamidine

Starting coordinates for the model compound 3,3-dimethyl-1-(2'-chloro)phenyltriazene were obtained from the crystal structure coordinates of AM90021 by removing the ethanesulphonate anion and the isopropyl alcohol molecule of crystallisation and replacing the 2,4-diamino-6-ethylpyrimidine fragment with a hydrogen atom, bond length 1.08 Å. A model structure of 3,3-dimethyl-1-(2'-chloro)phenylformamidine was constructed from the crystallographic coordinates for *N'*-(4-chloro-*ortho*-tolyl)-*N,N*-dimethylformamidine (Gifkins & Jacobson 1980). The 4-chloro and *ortho* methyl substituents were removed and replaced with a hydrogen atom, 1.08 Å from the crystal structure coordinate of C6P and a chlorine atom 1.74 Å from C4P.

The single point, self-consistent field (SCF), *ab initio* calculations, using the STO-3G basis set in the GAMESS program, and the MNDO, AM1 and PM3 semi-empirical calculations in MOPAC V6.0, were performed. The semi-empirical calculations continued until Herbert's test was satisfied. These calculations were then repeated with complete geometry optimisations. It was necessary to specify the keywords "GEO-OK" and "PRECISE" for the semi-empirical calculations.

The molecular mechanics routines, using the COSMIC parameters, were carried out using the modeling package NEMESIS. For the geometry optimisation, iterations continued until the root mean square (RMS) difference in geometry of the new, compared to the previous set of atomic coordinates, was consistently less than 0.01.

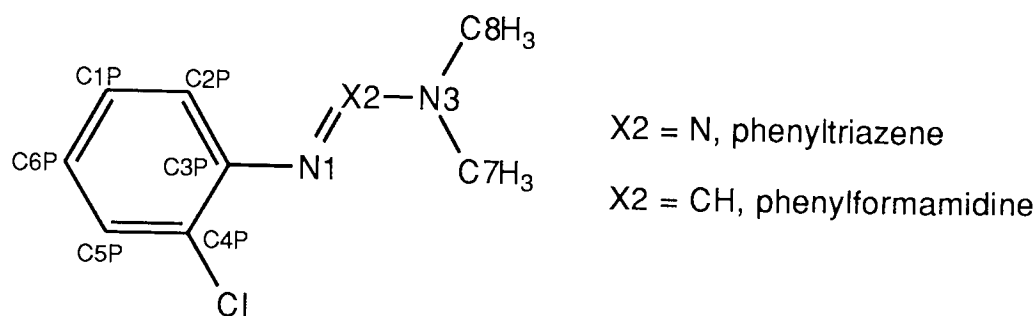


Figure 3.1 Phenyltriazene and Phenylformamidine atomic numbering

3.2.3 Superimposing the 3,3-Dimethyl-1-(2'-chloro)phenyltriazene Structures

Output from the calculations was edited into a .cssr (crystal Structure Search and Retrieval) formatted file. In the optimised structures the bond torsion C2P-C3P-N1-N2 was adjusted to -22.4° (figure 3.1). The structures were superimposed by using the "RIGID FIT" facility in the "FITTING" option in ChemX. Atoms C1P, C3P, C4P and C6P from each optimised structure were matched to the equivalent atoms in the original 3,3-dimethyl-1-(2'-chloro)phenyltriazene structure using a restraining constant of 10. The new coordinates were then saved in an updated .cssr file.

3.2.4 Determination of Molecular Similarity

The structure and charges from single point 3,3-dimethyl-1-(2'chloro)phenyltriazene *ab initio* calculation was used as the "FIXED" molecule. The superimposed single point and optimised structures were then entered. The Hodgkin electrostatic potential molecular similarity indices in ASP were determined using a "GRID EXTENT" of 4 Å and a "GRID INCREMENT" of 0.25 Å. Geometry optimisations and charge calculations were not specified.

3.3 Results and Discussion

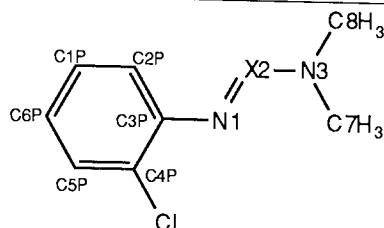
The aim of this comparison study was to develop an optimal method, in terms of accuracy and computational limitations, that could calculate the charge distribution and optimise the molecular geometry of higher members of the analogous series. Hence, with accurately determined steric and electronic parameters for these analogous compounds, more meaningful interpretations could be made concerning QSAR correlations and the interaction in the DHFR active sites.

It must be remembered that the optimisations have been calculated for an isolated molecule in the gaseous state, whereas the crystal structures that are being used as a comparison are highly ordered, compact solid structures. Some of the geometrical differences may be a result of these factors, though SHELX did not reveal any close contacts amongst the phenyltriazene atoms due to unit cell packing in AM90021.

As discussed previously there are unusual, but consistent characteristics in the phenyltriazene bond distances, angles and torsions (listed in table 3.1). It would be expected that these features would be recognised and predicted by a suitable geometry optimisation package. All the semi-empirical quantum mechanical calculations produce bond length data consistent with each other and very similar to the RMS values from a series of experimentally determined triazene structures (table 3.1a). The only anomaly is the 1.388 Å N2-N3 bond length predicted by PM3, which is considerably longer than the 1.321 Å experimental value. The STO-3G quantum mechanics and molecular mechanics calculations are not so consistent with the experimentally determined geometry. However, the *ab initio* calculated geometry agrees with 3-21G calculations on alkyl triazenes (Schmiedekamp *et al.* 1991). These methods gave N2-N3 bond distances, 1.387 Å and 1.393 Å respectively, that are considerably larger than the aforementioned RMS value. Whereas the *ab initio* method overpredicts the C3P-N1 bond length (1.421 Å) by approximately 0.05 Å, COSMIC predicts a bond length shorter than the experimentally determined distance by a similar amount.

STO-3G, PM3, AM1 and, to a lesser extent, COSMIC are able to predict the interference between the aromatic ring and the triazene group. This is shown by a C2P-C3P-N1 bond angle that is significantly greater than 120° (table 3.1b) and, with the exception of COSMIC, a magnitude for the bond torsion C2P-C3P-N1-N2 (table 3.1c) that is greater than 0°. COSMIC suggests that the triazene function and the aromatic ring

are co-planar. The bond angles calculated using the MNDO parameters suggest there is no interference. This is a consequence of the optimised MNDO geometry orientating the triazene function perpendicular to the aromatic ring. In this situation interaction between the two moieties is minimal, therefore interference is unlikely. However, this also demonstrates the limitations that the earlier semi-empirical quantum mechanics packages have in handling conjugated and delocalised molecular systems. Because of the necessity to employ default parameters, COSMIC suggests that the triazene and aromatic moieties are co-planar, thus demonstrating that this molecular mechanics force field is too limited to deal adequately with the complexity of the phenyltriazene. Another characteristic feature of the experimental phenyltriazene geometry is the tightness of the two bond angles C3P-N1-N2 and N1-N2-N3, 112.0° and 114.4° respectively. The *ab initio* package produced bond angles, 109.7° and 111.4° , that are closest to the RMS values. Of the other packages, only MNDO produces bond angles similar to those determined experimentally. The crystallographic studies in the previous chapter show that the triazene geometry is unaffected by the chlorine substituent bonded to the adjacent carbon. MNDO is the only package to produce a C4P-C3P-N1 bond angle greater than 120° , suggesting that there is interference and further illustrating the limitations of this parameter set.



X2 = N, phenyltriazene

X2 = CH, phenylformamidine

Bond	triazene (RMS)	formamidine (RMS)	STO-3G	PM3	AM1	MNDO	COSMIC
C3P-N1	1.421(7)	1.396(21)	1.468	1.443	1.427	1.434	1.363
N1-X2	1.271(6)	1.290(11)	1.286	1.236	1.246	1.240	1.293
X2-N3	1.321(9)	1.337(11)	1.387	1.388	1.339	1.333	1.393
N3-C7	1.447(12)	1.453(17)	1.471	1.488	1.458	1.480	1.454
N3-C8	1.466(13)	1.458(11)	1.471	1.489	1.460	1.470	1.452

Table 3.1a RMS bond lengths from phenyltriazene and phenylformamidine crystal structures and calculated triazene bond lengths (Å) and s.d. in parentheses.

Angle	triazene (RMS)	formamidine (RMS)	STO-3G	PM3	AM1	MNDO	COSMIC
C2P-C3P-N1	124.7(5)	124.0(7)	123.4	124.0	125.1	119.3	122.3
C4P-C3P-N1	116.6(13)	117.2(8)	118.0	116.1	116.9	121.4	119.5
C3P-N1-X2	112.0(4)	117.4(17)	109.7	119.7	119.5	115.6	114.1
N1-X2-N3	114.4(7)	122.9(7)	111.4	117.1	119.4	116.7	120.8
X2-N3-C7	122.9(11)	120.9(17)	115.0	120.8	124.0	124.7	119.5
X2-N3-C8	115.5(14)	121.4(14)	120.3	112.1	115.2	112.8	117.8
C7-N3-C8	121.0(13)	117.6(23)	124.4	113.0	113.3	118.2	122.1

Table 3.1b RMS bond angles from phenyltriazene and phenylformamidine crystal structures and calculated bond angles (°) and s.d. in parentheses

Torsion	triazene (RMS)	formamidine (RMS)	STO-3G	PM3	AM1	MNDO	COSMIC
C2P-C3P-N1-X2	±15.0	±32.5	24.5	-22.0	-46.1	89.7	0.0
C4P-C3P-N1-X2	±167.6	±154.0	-158.0	160.3	139.4	-97.0	-179.9
C3P-N1-X2-N3	±178.7	±176.6	-178.9	174.7	173.4	-176.0	179.9
N1-X2-N3-C7	±5.6	±2.8	2.5	23.8	9.6	-11.2	4.4
N1-X2-N3-C8	±174.5	±176.7	-179.5	161.0	157.1	-167.4	175.4

Table 3.1c RMS bond torsions from phenyltriazene and phenylformamidine crystal structures calculated triazene bond torsions (°)

Crystallographic data show that the terminal phenyltriazene methyl carbons are coplanar with N2 and N3. This could be confirmed by the addition of the final three bond angles (table 3.1b) producing approximately 360°. The planarity could not be extended further to include N1, shown by a N1-N2-N3-C7 RMS torsion angle of $\pm 5.6^\circ$. STO-3G predicts a terminal planar geometry very similar to the RMS geometry. However, the semi-empirical packages could not demonstrate this planarity. The planarity is a consequence of resonance allowing the N2-N3 bond to adopt double bond characteristics (Akhtar *et al.* 1968). The failure of the semi-empirical methods to demonstrate this planarity further illustrates their inability to predict electron delocalisation effects. COSMIC, by virtue of using default parameters on the triazene moiety, is able to produce bond angles for these four atoms that resemble those determined experimentally.

Disregarding the physical differences between environments in which the experimental and theoretical geometries are determined these studies indicate that even the more rigorous programs cannot predict some of the geometrical characteristics of the phenyltriazene. When these DHFR inhibitors are bound in the active site it is likely they would be subjected to a similar restrictive spatial environment as these crystal structures. Therefore, it was decided that for future study the bond angles and distances of the phenyltriazene moiety should be constrained to the experimentally determined RMS values.

Because of the quantum mechanical nature, *ab initio* programs are the most accurate for determining the molecular electronic properties from which partial charges can be calculated. The excessive computing demands limit the *ab initio* calculations to small molecules. Therefore, it is necessary to determine which of the less sophisticated packages is most accurate in calculating electronic properties of the phenyltriazene when compared to *ab initio* results, which could then be applied to the pyrimethamine analogues. The 3-21G basis set is often used as a basis for comparison purposes (Clark 1985). However, geometry optimisation failed to converge for either model structure using this basis set in GAUSSIAN 92 on a Cray supercomputer, carried out by C.H. Schwalbe at Oxford University. Thus for comparison purposes the STO-3G calculated Lowrie partial charges are the most accurate available.

With the exception of N1 the empirically calculated COSMIC charges are considerably

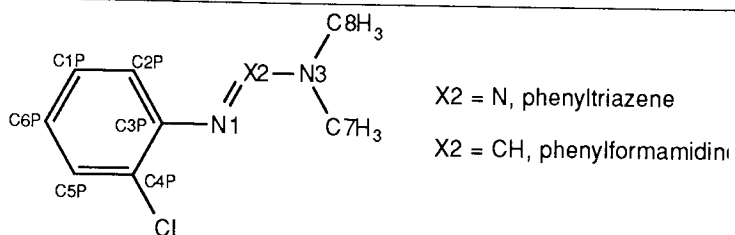
more accurate than those calculated using MNDO (table 3.2). This confirms the limitations of MNDO for the study of the delocalised triazene system. With the exception of one atom each, C3P for AM1 and N3 for PM3, the two other semi-empirical packages produced calculated partial charges very similar to the STO-3G charges. For AM1, the partial charge on C3P is 0.068 electronic charge (e) less than the *ab initio* result, whereas for PM3 the calculated partial charge on N3 is 0.309 e greater. This considerable discrepancy suggests there are limitations with the PM3 parameter set for studying phenyltriazenes. In conclusion, the AM1 calculated partial charges have greatest similarity to the *ab initio* calculation, and thus are most applicable to phenyltriazenes when it is not possible to use *ab initio* methods.

With the exception of the PM3 results, comparison of the phenyltriazene and phenylformamidine charges (tables 3.2a and 3.2b) consistently show more negative values on N1 and N3 of the phenylformamidine and a more positive C2 charge. Allowing for the greater electronegativity of N2 this result is expected. These differences in electrostatic potential could account for the species selectivity of DHFR inhibitors as suggested by Andrews *et al.* (1986).

In the ASP molecular similarity calculations the advantage of using electrostatic field is that this can be calculated from point charges, thus allowing partial charges calculated from other programs to be used. Studies have shown that a 4 Å grid size with a 1 Å increment is optimal for the comparison of indices (Burt *et al.* 1990). Although a 0.25 Å grid increment substantially increases the computational demands, it is preferred because it has been shown to consider the molecular shape as well (Meyer & Richards 1991). STO-3G charges calculated for the original 3,3-dimethyl-1-(2'-chloro)phenyltriazene structure were used as references for the optimised and other single point calculations. Re-orientating the C2P-C3P-N1-N2 torsion back into the original crystallographic conformation was used to reduce the bias of shape in the similarity calculation.

The similarity indices (table 3.3) show that single point calculations consistently produced higher values than the equivalent optimised calculation. As the Hodgkin index accounts for shape and electrostatic similarity (Good *et al.* 1993a), this is to be expected. The index from the STO-3G optimisation is closest to the reference compound. The next highest similarity index, 0.868, confirms the suitability of using

AM1, with constrained geometry, for further studies of the pyrimethamine triazene analogues.



Atom	STO-3G	PM3	AM1	MNDO	COSMIC
C3P	0.074	0.011	0.006	0.048	0.057
N1	-0.178	-0.178	-0.183	-0.203	-0.286
N2	-0.015	-0.037	0.041	0.079	0.054
N3	-0.205	0.114	-0.230	-0.345	-0.090
C7	-0.074	-0.109	-0.116	0.195	-0.045
C8	-0.068	-0.075	-0.078	0.213	-0.045

Table 3.2a Calculated partial charges of the phenyltriazene atoms

Atom	STO-3G	PM3	AM1	MNDO	COSMIC
C3P	0.085	-0.019	0.028	0.087	0.041
N1	-0.293	-0.139	-0.222	-0.325	-0.417
C2	0.180	-0.061	0.063	0.263	0.226
N3	-0.348	-0.251	-0.251	-0.415	-0.160
C7	-0.074	-0.086	-0.096	0.221	-0.049
C8	-0.077	-0.092	-0.096	0.208	-0.067

Table 3.2b Calculated partial charges for the phenylformamidine

Optimisation Basis Set	<u>Hodgkin Electrostatic Potential Similarity Indices</u>					
	Phenyltriazene			Phenylformamidine		
	Optimised	Single point	RMS	Optimised	Single Point	RMS
STO-3G	0.901	1.000	0.005	0.376	-	0.005
PM3	0.652	0.695	0.006	0.448	0.466	0.005
AM1	0.824	0.868	0.018	0.437	0.455	0.002
MNDO	0.797	0.858	0.003	0.357	0.374	0.001
COSMIC	0.793	0.860	0.004	0.329	0.349	0.002

Table 3.3 ASP molecular similarity indices of optimised and fixed point crystal structure geometries compared with the electrostatic potential of the original 3,3-dimethyl-1-(2'-chloro)phenyltriazene calculated using STO-3G the basis set

Chapter 4 Biochemical Studies on Trimethoprim, Pyrimethamine and 3'-(3'',3''-disubstitutedtriazene-1''-yl) Pyrimethamine Analogues

4.1 Introduction

As discussed in chapter 1, DHFR catalyses the NADPH-dependent reduction of DHF to THF. The most commonly used method for assaying this reaction is the UV spectrophotometric assay developed by Osborn and Huennekens (1958) in which a decrease in absorbance at 340 nm is measured over a time period as DHF and NADPH are converted to THF and NADP⁺ respectively. The change in absorbance at 340 nm records the loss of NADPH. In the presence of a DHFR inhibitor, the rate of NADPH oxidation decreases. As a means of quantifying and comparing this inhibitory activity the IC₅₀ for each compound is determined. This is defined as the concentration of inhibitor necessary to reduce the reaction rate to 50% of the uninhibited rate. It can be obtained from a plot of log₁₀[concentration of inhibitor] versus inhibitory activity. Because the IC₅₀ is determined from a dose-response correlation, it provides a value for biological activity suitable for use in QSAR studies (Dunn 1990).

Denny *et al* (1992) were able to correlate the competitive inhibition of DHFR by methyl benzoprim against NADPH with molecular modeling studies of the docked inhibitor. Thus, classical enzyme kinetic studies can also provide descriptive information concerning how the inhibitor binds to the macromolecule with respect to the substrates.

4.2 Materials and Methods

Initial screening of activity of the compounds against *P. carinii* and rat liver DHFR was carried out by Dr. S.F. Queener of the National Institute of Health, Maryland, U.S.A., using procedures described by Broughton & Queener (1991). More detailed studies were carried out as described below.

4.2.1 Source of Reagents and Enzymes Used in the Assays

The following reagents were purchased from or supplied by the sources indicated

Aston Molecules Ltd., Gosta Green, West Midlands

3'-(3'',3''-disubstitutedtriazene-1''-yl) pyrimethamine analogues (AM90020-AM90074)
(structures shown in tables 4.1 and 4.2)

Sigma Chemical Company, Poole, Dorset

Dihydrofolic acid (90%)

Reduced nicotinamide adenosine dinucleotide phosphate (97%) sodium salt

Microbiology Technology Laboratory, Centre for Applied Microbiology and Research,
Porton Down, Wiltshire

Frozen cell paste of *Escherichia coli* MRE 600 (EC/600/38, Lot 5). Cells were grown aerobically by continuous culture at 37°C under conditions of glycerol limitation and a dilution rate of 0.78 hr⁻¹.

4.2.2 Preparation of Partially Purified DHFR from *Escherichia coli*

Partially purified DHFR from *E. coli* was prepared using the procedure from Burchall and Hitchings (1965).

4.2.3 Preparation of Partially Purified DHFR from Rat Liver

This was prepared using the method of Bertino and Fischer (1964).

4.2.4 Preparation of Reagents

DHF, NADPH and inhibitor solutions were prepared freshly on the day of the assay.

4.2.4.1 Dihydrofolate Solution

DHF was suspended in 0.25 M 2-mercaptoethanol to give a final concentration of 1 mgml⁻¹. 0.1M NaOH was added dropwise to the suspension which was vigorously shaken until complete dissolution of the DHF occurred.

4.2.4.2 Inhibitor Solution

Inhibitor solutions of concentration between 6.25 mg l⁻¹ and 1000 mg l⁻¹ were prepared by dilution of a freshly prepared 4000 mg l⁻¹ stock solution. All stock solutions and dilutions were prepared in DMSO.

4.2.5 Assay for IC₅₀ Inhibitory Activity Against *Escherichia coli* and Rat Liver DHFR

The spectrophotometric assay was carried out using the following program on an Anthos 2001 Plate Reader (Anthos Labtec Instruments). Absorbance readings were taken every 3 minutes at 340 nm, over 30 minutes, with 10 seconds of shaking between each reading. Each assay was carried out in one well of a Dynatech Immulon 96 well ELISA plate. To each well 50 µl of prepared enzyme, 50 µl of 0.15 M potassium phosphate buffer and 20 µl of inhibitor solution were added and incubated at 37°C for 5 minutes. 40 µl of NADPH, 2 mg ml⁻¹ in 0.15M potassium phosphate buffer pH 7, was added, then left for 1 minute before the DHF solution was added to initiate the reaction. The ELISA plate was then placed in the chamber of the plate reader and assay left to run. The time curve for each reaction was plotted and all initial rates determined. To obtain a measurement of maximum activity of the enzyme in the absence of inhibitor, the inhibitor solution was replaced by 20 µl of DMSO.

The partially purified enzyme preparations contained some non-specific NADPH oxidase activity. Thus, controls were required to assess the level of this activity in the absence of DHF, which was replaced with potassium phosphate buffer. Controls were also carried out to determine if the inhibitors possessed any activity against the NADPH oxidase. Therefore, at each inhibitor concentration, the assay was repeated with the DHF solution replaced with the buffer.

4.2.6 Enzyme Kinetic Studies of AM90067 Against DHF in Rat Liver Dihydrofolate Reductase

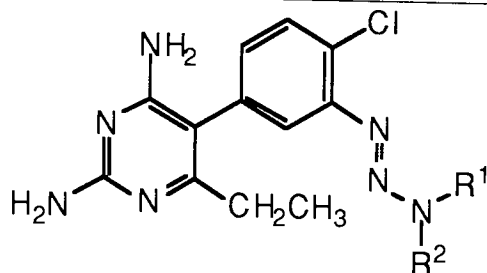
This was carried out as described in section 4.2.5, using rat liver DHFR (50 µl) with a range of concentrations of DHF and fixed concentration of NADPH (40 µl of 2 mg ml⁻¹). DHF (40 µl) solution was used to give the following final concentrations; 0.02, 0.03, 0.04, 0.08, 0.12, 0.20, 0.30, 0.40 and 0.80 mg ml⁻¹. The assays were repeated in the presence of AM90067 at final concentrations of 5 x 10⁻⁷ M, 1 x 10⁻⁶ M and 2.5 x 10⁻⁶ M.

4.2.7 Enzyme Kinetic Studies of AM90067 against NADPH in Rat Liver Dihydrofolate Reductase

This was carried out as described in 4.2.6 using a fixed concentration of DHF (40 μl of 1 mg ml^{-1}) over the following range of final NADPH concentrations: 0.02, 0.03, 0.04, 0.08, 0.12, 0.20, 0.30, 0.40 and 0.80 mg ml^{-1} . The assay was repeated in the presence of AM90067 at final concentrations of 5×10^{-7} M, 1×10^{-6} M and 2.5×10^{-6} M.

4.3 Results and Discussions

Inhibitory activity was expressed as the percentage inhibition of the initial rate of the enzyme assay. To validate the assay procedure, the IC_{50} s for trimethoprim and pyrimethamine were determined against rat liver DHFR. The result compares closely with those obtained by Broughton and Queener (1991). Also the pyrimethamine and trimethoprim IC_{50} s against *E. coli* and rat liver DHFR are very similar ($\pm 10\%$) to those reported by other workers (Broughton & Queener 1991, Griffin & Lowe 1992, Chan & Roth 1991). The IC_{50} s and the ratio of IC_{50} s of rat liver to *P. carinii* DHFR (RL/PC) and rat liver to *E. coli* DHFR (RL/EC) are given in tables 4.1 and 4.2. Ideally an RL/PC IC_{50} ratio of similar magnitude to the RL/EC ratio of trimethoprim would be necessary for any clinical potential as antipneumocystis DHFR inhibitors. Some of the AM compounds (11/20) exhibit a RL/PC selectivity ratio less than 1. Most pronounced selectivity is displayed by AM90067 with a RL/PC ratio of 21, comparable to that of trimethoprim. Since the series of compounds differs only in the triazene substituent this pattern of results suggests that the chemical nature of the terminal triazene groups has a profound effect on the inhibitory profile. Of the heterocyclic substituted compounds AM90026 has the largest RL/PC IC_{50} ratio. Thus, there appears to be a consistency between the third main chain atom from N3 being oxygen and a favourable RL/PC IC_{50} ratio. Therefore, it would be logical to employ QSAR methods to determine any correlation between the physicochemical characteristics of the terminal substituents and the biological activities.



Inhibitor			IC ₅₀ (μM) for DHFR from			IC ₅₀ Ratios	
	R ¹	R ²	Rat Liver (RL)	<i>P. carinii</i> (PC)*	<i>E. coli</i> (EC)	RL/PC	RL/EC
pyrimethamine	-	-	5.5	3.65	-	1.5	-
trimethoprim	-	-	390	12	0.022	32.5	~17000
AM90020	Me	Me	18.9*	2.8	91	6.75	0.21
AM90022	Et	Et	27	29.3	101	0.92	0.27
AM90024	Me	Pr(i)	4.8	10.2	16	0.47	0.30
AM90029	Me	Et	202*	14.1	102	14.3	2.0
AM90035	Me	Bz	2.3	2.8	-	0.82	-
AM90063	EtOH	EtOH	3.4	0.44	45	7.7	0.08
AM90066	Me	EtOH	3.5	3.3	24	1.1	0.14
AM90067	Bz	EtOH	5.3	0.26	3.5	21	1.5
AM90068	Et	EtOH	4.2	11.5	25	0.37	0.17
AM90069	Pr(n)	EtOH	6.2	2.5	17	2.48	0.36
AM90070	Me	EtN(Et) ₂	10.2	18.2	9.6	0.56	1.06
AM90071	Me		1.9	0.57	26	0.33	0.07
AM90072	Et	C ₆ H ₁₁ (c)	0.74*	4.1	41	0.18	0.02
AM90073	Bu(t)	EtOH	14.2	4.9	19	2.9	0.75
AM90074	H	Benzoyl-L-glutamate	0.67	2.1	-	0.32	-

Table 4.1 Biological activities for pyrimethamine, trimethoprim and non-heterocyclic triazenyl pyrimethamine analogues

Abbreviations: Me, methyl (-CH₃); Et, ethyl (-CH₂CH₃); Pr(i), isopropyl (-CH(CH₃)₂); Bz, benzyl (-CH₂C₆H₅); Bu(t), tertiary butyl (-C(CH₃)₃); EtOH, hydroxyethyl (-CH₂CH₂OH); Pr(n), propyl (-CH₂CH₂CH₃); C₆H₁₁(c), cyclohexyl; EtN(Et)₂, diethylethylamine (-CH₂CH₂N(CH₂CH₃)₂).

* Data supplied by Dr. S.F. Queener, National Institute of Health, Maryland, U.S.A.

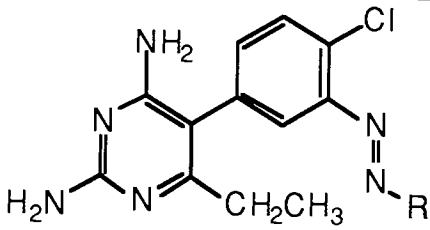
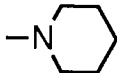
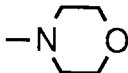
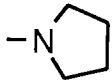
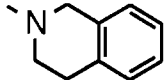
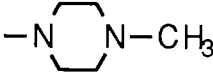
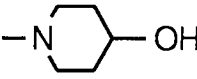
						
Inhibitor	R	IC ₅₀ (μM) for DHFR from			IC ₅₀ Ratios	
		Rat Liver (RL)	<i>P. carinii</i> (PC)*	<i>E. coli</i> (EC)	RL/PC	RL/EC
AM90023		4.0	40.4	8.5	0.10	0.47
AM90026		26.3*	1.7	50	15.3	0.53
AM90034		3.9	>11	12	-	0.33
AM90036		2.0	5.0	0.75	0.4	2.7
AM90038		3.7	8.5	11	0.44	0.34
AM90064		5.0	3.3	6.0	1.5	0.83

Table 4.2 Biological activities for heterocyclic substituted triazene pyrimethamine analogues

* Data supplied by Dr. S.F. Queener, National Institute of Health, Maryland, U.S.A.

Determination of the type of enzyme inhibition (i.e. competitive, non-competitive or uncompetitive) indicates how the inhibitor binds to the active site with respect to the substrate. From this, assumptions can be made concerning the docking of AM90067 into the active site models of mammalian and *P. carinii* DHFR. Over the concentration ranges investigated, figure 4.1 shows that AM90067 is a competitive inhibitor of DHF in rat liver DHFR. Figure 4.2 shows that at low concentrations (5×10^{-7} - 1×10^{-6} M)

AM90067 is also a competitive inhibitor of NADPH in rat liver DHFR. However, the results suggest that at the highest concentrations of AM90067 (2.5×10^{-6} M) non-competitive inhibition occurs. Thus when AM90067 is bound to DHFR it can partially occupy the binding sites of both DHF and NADPH.

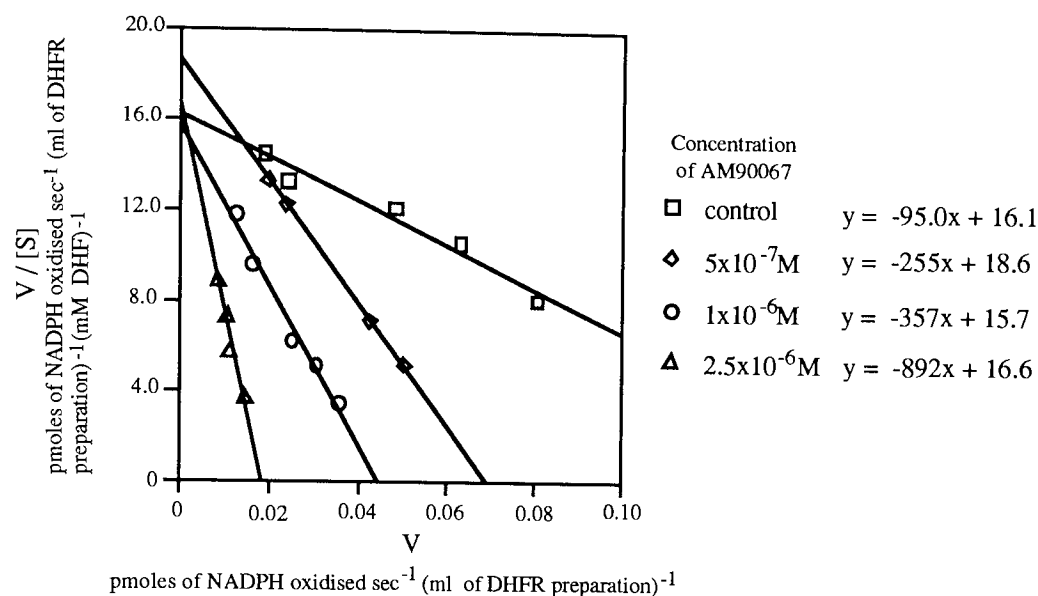


Figure 4.1 Eadie-Hofstee plot for AM90067 at various concentrations of DHF using rat liver DHFR. The Eadie-Hofstee plot ($V/[S] : V$) shows the rate of reaction (V) on the x-axis and the rate of reaction divided by the DHF concentration ($[S]$) on the y-axis. The assay was carried out at a fixed rate saturating concentration of NADPH.

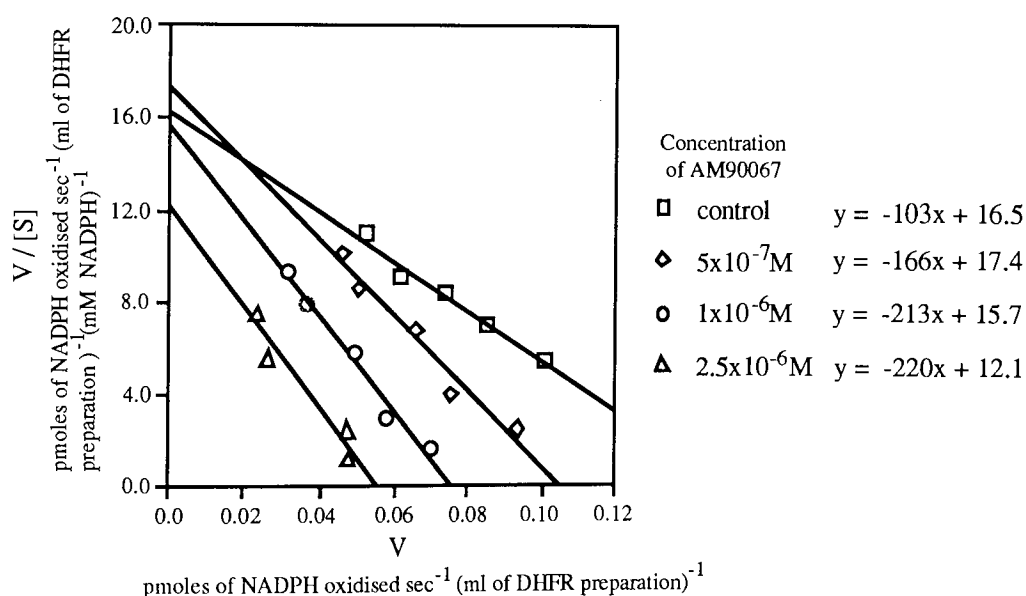


Figure 4.2 Eadie-Hofstee plot for AM90067 at various concentrations of NADPH using rat liver DHFR. The assay was carried out at a fixed rate saturating concentration of DHF.

X-ray crystallographic studies of Baker *et al.* (1983) show that the diaminopyrimidine moiety in pyrimethamine binds to key active site residues of *Lactobacillus casei* DHFR in an identical manner to the equivalent atoms of methotrexate (figures 4.3a and 4.3b), a DHF analogue and potent competitive inhibitor of DHFR. If a similar mode of binding occurs for AM90067 in rat liver DHFR, the observed competitive inhibition with respect to DHF would be explained.

In order for reduction of the pteridine moiety of DHF at positions 7 and 8 to occur the NADPH cofactor must bind in close proximity. This has been confirmed from the crystal structures of the ternary complexes of *E. coli* and *L. casei* DHFRs with methotrexate and NADPH (Filman *et al.* 1982). Binding studies on the pyrimethamine analogue 2,4-diamino-6-ethyl-5-(3'-fluoro-4'-nitro)phenylpyrimidine in *L. casei* DHFR (Birdsall *et al.* 1990) showed a change in fluorine environment in the presence of NADPH. It was suggested that steric interactions between the fluorine and the NADPH forced the C4D-C5D-C1P-C2P torsion (see figure 2.1, atomic numbering as for AM90021) around 180°, thereby orientating the fluorine to a different environment. In AM90067 the 3'-(3"-benzyl-3"-hydroxyethyltriazene-1-yl) group is considerably larger than the fluorine and therefore possibly unable to adopt a second conformation that would permit NADPH binding. Thus, the results from figure 4.2. suggest that this bulky substituent competes with NADPH for occupancy of the active site binding region, a feature which will be developed in subsequent docking studies.

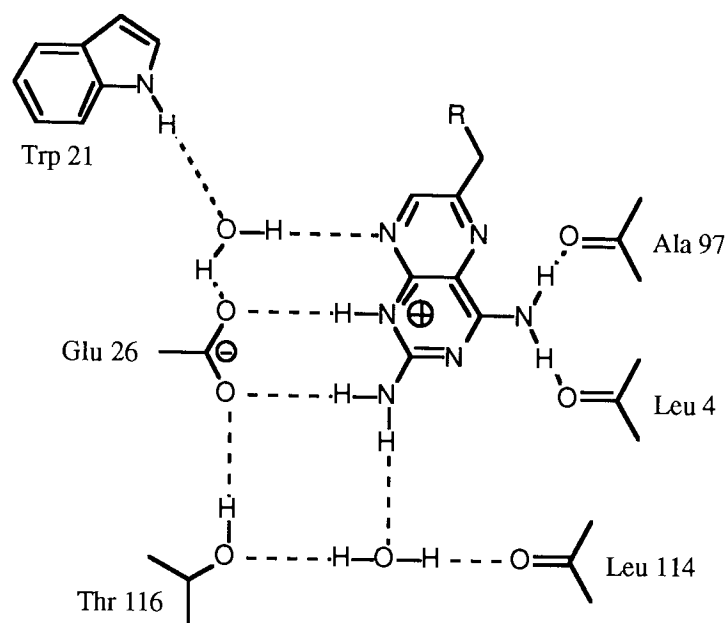


Figure 4.3a A schematic representation of the hydrogen bonding between DHFR from *L. casei* and methotrexate

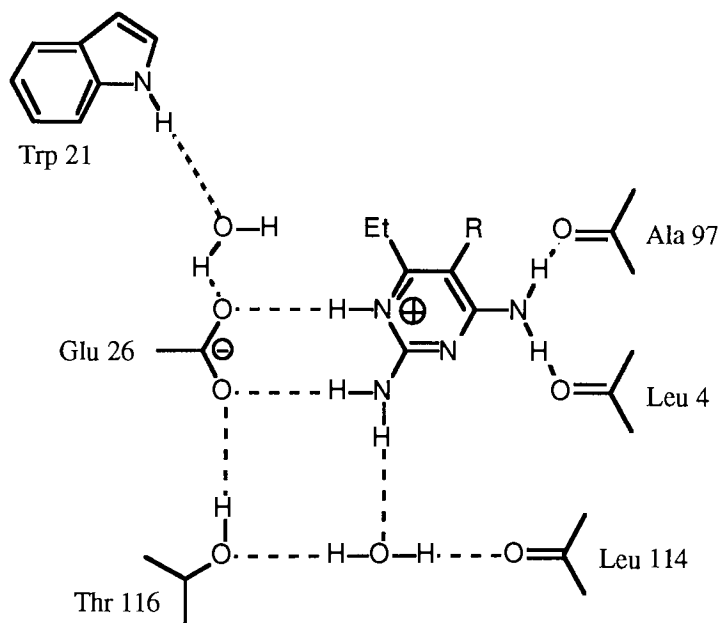


Figure 4.3b A schematic representation of the hydrogen bonding between DHFR from *L. casei* and pyrimethamine

Chapter 5 Determination of Quantitative Structure Activity Relationships Among the 3'-(3'',3''-disubstitutedtriazene-1''-yl) Pyrimethamine Analogues

5.1 Introduction

Without the aid of active site structural data, it is still possible to make quantifiable estimates of the biological activity of a compound by investigating the effect that certain physicochemical properties, such as Hammett's potential (σ), Hansch's hydrophobicity (π) and Taft's molecular volume (E) of a series of analogous compounds, have on the inherent biological activity. By correlating the substituent effect contributing to the overall physicochemical property with the biological activity it is possible to determine which parameters have the greatest influence and can then be incorporated into the next generation of analogues to be synthesised. Furthermore, since molecular modeling packages are capable of calculating structural and physicochemical parameters, with good correlations, it may be possible to predict the biological activity for an untested compound (Hansch *et al.* 1962). Thus novel compounds can be quickly assessed for the possibility of possessing the desired biological activity. The methods in determining quantitative structure activity relationships (QSAR) have progressed considerably, with two new techniques, molecular similarity (Good *et al.* 1993a & b) and neural networks (Aoyama & Ichikawa 1991a) receiving considerable attention.

5.1.1 Molecular Similarity

For comparing molecular similarity, superimposing the standard ball and stick representations visualised on a graphics monitor has limited value. The similarities in the atomic conformation revealed cannot forecast the potential a ligand has to interact at the receptor site. As mentioned previously this is also reliant on the inhibitor possessing an electronic distribution complementary to the active site. Hence, numerical indices have been developed to measure the overall steric and electrostatic similarity between pairs of molecules to be correlated with biological activity (Carbo *et al.* 1980).

The computer program ASP (Automated Similarity Package, Oxford Molecular 1990) has been designed to compare the electrostatic potential, P_A and P_B , of two superimposed molecules A and B. The Hodgkin index H_{AB} (equation 5.1) (Hodgkin &

Richards 1987) employed was created to have a greater sensitivity toward the magnitude than the shape of the electrostatic potential.

$$H_{AB} = \frac{2 \int P_A P_B dv}{\int P_A^2 dv + \int P_B^2 dv} \quad \text{equation 5.1}$$

Comparative Molecular Field Analysis methodology, as a technique applicable in the design of novel DHFR inhibitors, was established originally by Hopfinger (1980). Using ASP Burt *et al.* (1990) were able to demonstrate a correlation between molecular similarity and biological activity in series of insecticides. More recently, it has been demonstrated that the technique can be applied to several molecular series to validate structure activity relationships (Good *et al.* 1993b).

5.1.2 Neural Networks in QSAR

Traditional QSAR methods use multiple regression analysis to fit a mathematical function to calculated physicochemical properties (X, Y, ...) and experimentally determined biological data (A) (Hansch *et al.* 1962)

$$A = C_0 + C_1X + C_2X^2 + C_3Y + C_4Y^2 + \dots \quad \text{Equation 5.2}$$

where C_0, C_1, \dots , minimise the variance between the data and the model. However, multiple linear regression analysis of QSAR data is restricted to first and second order terms only and does not consider cross-product terms. The consequence of this is the production of a model with very limited flexibility, unable to cope with outlying data.

In recent years there has been growing interest in the use of neural networks (figure 5.1) in QSAR studies. A neural network consists arbitrarily of n layers of perceptrons (figure 5.2), the basic unit of a network, each connected to the previous layer. The perceptron calculates the weighted sum of inputs and uses a threshold function to convert this to output in the same way that neuron excitation is controlled from input from other axons. The main strength of the neural networks is their ability to perform complex nonlinear mapping of the physicochemical parameters onto the biological activity with no prior knowledge of the functions involved (Aoyama & Ichikawa 1991b). Recent QSAR studies on various classes of DHFR inhibitors have shown the

neural network to produce a better correlation with experimental results than multiple linear regression (Andrea & Kalayeh 1991, So & Richards 1992)

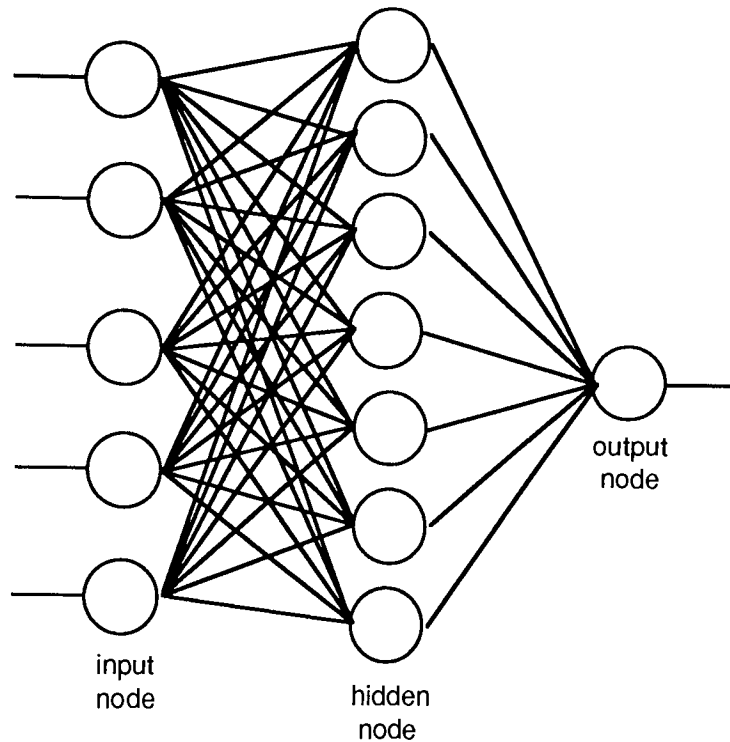


Figure 5.1 A schematic representation of a simple neural network, consisting of three layers of perceptrons (figure 5.2), as used in this study

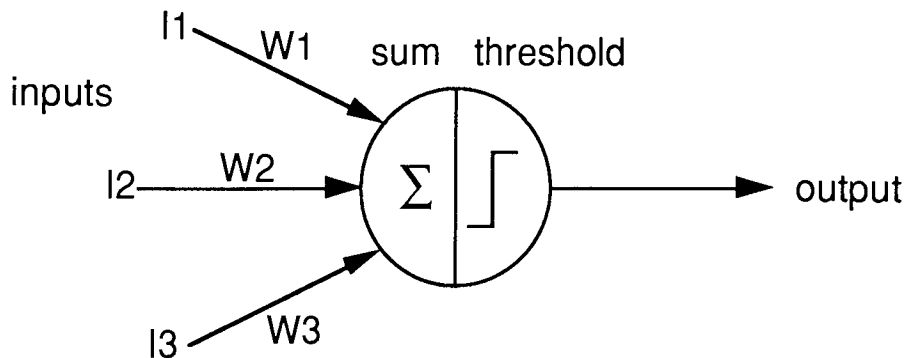


Figure 5.2 A schematic representation of a perceptron. Summation of the inputs I1, I2 and I3, weighted by W1, W2 and W3 respectively, will produce an output if it is greater than the threshold value.

5.2 Materials and Methods

5.2.1 Molecular Modeling Packages

ChemX was used as an interface to MOPAC from which the AM1 parameter set was used. ASP and the neural network and interface program, Aspirin/MIGRAINES V6.0 (Leighton, The MITRE Corporation 1992) were mounted on SUN SPARCstations.

5.2.2 Model Compound Construction

From the crystal structure coordinates of AM90021, the ethanesulphonate anion and the isopropyl alcohol molecule of crystallisation were removed. Additional functional groups were added to the remaining protonated AM90020 structure from the fragment library in ChemX using standard bond lengths and angles. The initial geometry of any given substituent was identical. AM1 MOPAC optimisation calculations used the restraints suggested in chapter 3 with the keywords "GEO-OK" and "CHARGE=1" specified.

5.2.3 Determination of Molecular Similarity Indices

Torsion angles of the optimised geometries were rotated so that the atoms occupied the same space as the equivalent atoms in AM90020. The optimised geometries were then superimposed onto the AM90020 structure using the "RIGID FIT" RMS facility in ChemX, with N21D, N41D, C61D and Cl1 (figure 2.1) nominated as the equivalent atoms. The superimposed geometries were saved to .cssr formatted files which were read into ASP. The Hodgkin electrostatic potential similarity indices were calculated using the optimised protonated AM90067 structure as the "FIXED" molecule with a "GRID EXTENT" of 4 Å and a "GRID INCREMENT" of 0.5 Å.

5.2.4 Determination of Molecular Volume of the Terminal Triazene Substituents

The two methyl groups were removed from the optimised protonated AM90020 coordinates. The resulting structure was saved to a protein databank (.pdb) formatted file. The torsion angles from the optimised pyrimethamine analogues were orientated into an identical conformation to AM90020, then modified to contain only the one

terminal substituent and saved as a .pdb file. This was repeated for the second substituent where necessary. In Quanta the corresponding molecular structure file (.msf) files were created. In the "Applications" menu "**molecular similarity**" was selected. The structure without terminal substituents was superimposed onto the other structures in turn by using N21D, N41D, C61D and C11 as equivalent atoms in the "match atoms" facility as part of a "rigid body to target" least squares fit. The "molecular volume" facility was able to calculate the difference in volumes between two superimposed structures, the substituent volume.

5.2.5 Neural Network Procedures

The Aspirin/MIGRAINES neural network and interface programs were used to map the five parameters listed in table 5.1 on to the RL/PC DHFR inhibitory ratios. All the data points were adjusted so that the new values were between 0 and 1. Prior to the mapping the following lines of text were added to the login command file (.crshc) so that the correct environment was set up for using Aspirin :-

```
setenv NNTOOLS /usr/local/aspirin
setenv MACHTYPE sun_sparc
setenv MANPATH $NNTOOLS/man:$MANPATH
set path = ($path $NNTOOLS/bin/$MACHTYPE)
```

The following control file (*.aspirin) was created to instruct Aspirin that there was one output layer, five input layers in the data file and to use a three layer neural network (figure 5.1):-

```
{
    Output Layer->      OutputLayer
    InputSize->         5
    Components->
    {
        PdpNode Hidden 3
        {
            InputsFrom-> $INPUTS
        }
        InputsFrom-> Hidden
    }
}
```

Using the command :-

bpmake

the control file is translated into a C formatted file, which was then compiled into a machine executable program named (*). The mapping of the chemical parameters was completed by typing machine executable program name (*) and the following qualifiers -l, to instruct the neural network it is a learning data set, and -d, to instruct the neural network to a data format file (*.df). This file, with the following format :-

ReadFile(*.data Ascii)

instructed the neural network that the data were stored in the file *.data and were in Ascii format.

5.3 Results and Discussion

For the molecular similarity calculations a 4.0 Å “GRID EXTENT” and a 0.5 Å “GRID INCREMENT” were used to take account of both the molecular shape and the electrostatic distribution. AM90067 was used as the “FIXED” molecule because it had the largest RL/PC IC₅₀ ratio in the series and therefore any correlation between the similarity indices and the inhibitory activity ratio would be easier to detect. Figure 5.3 shows that the values of the ASP calculated molecular similarity indices do not correlate with the RL/PC inhibitory activity ratios. Because it is only in the identity of the terminal triazene substituents that these compounds differ, it may be that chemically these compounds are too similar for ASP to detect any differences. However, when the calculations were repeated with abbreviated model structures that included the 5-(4'-chloro-3'-(3'',3''-disubstitutedtriazene-1''-yl)phenyl fragment only there was no significant improvement in the regression co-efficient obtained.

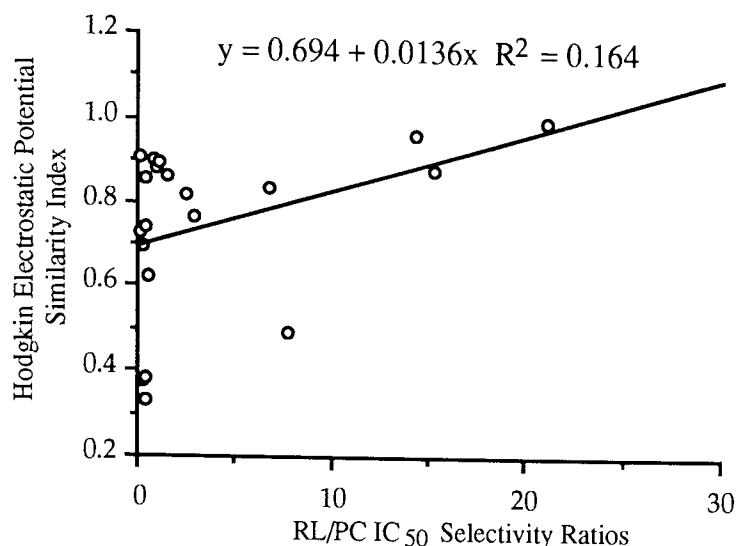


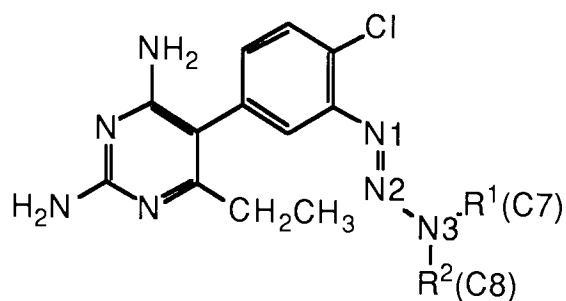
Figure 5.3 Plot of Hodgkin electrostatic similarity index, calculated as described in section 5.2.3, against the RL/PC DHFR IC₅₀ selectivity ratios, reported in tables 4.1 and 4.2

Table 5.2 lists the structural and electronic parameters that were mapped onto the RL/PC inhibitory activity ratios using the Aspirin neural network. The triazene substituent molecular volume and the partial charges on the atoms N3, C7 and C8 were selected as it had been reasoned that the triazene substituent was responsible for the greater selective inhibitory activity towards *P. carinii* DHFR and that the triazene substituents may have an influence on the DHFR inhibitory profile.

A three layer neural network was used to map the chemical parameters onto the biological data because the training set was small and using more layers would only increase the computational time, without significantly improving the precision of the output. Aspirin was able to give a very good correlation between experimental and calculated RL/PC values ($R^2 = 0.853$) (table 5.2). However the RMS of the differences between the actual and calculated values was 0.869. This suggested that in the situation where there was a low RL/PC ratio the neural network was not sufficiently accurate in its convergence of the structural data onto the biological data. The major reasons for obtaining poor convergence using a neural network have been outlined by Aoyama & Ichikawa (1991a) and better convergence may be obtained with knowledge of the degree of contribution any input parameter has on the biological activity. Although a neural network can establish a correlation using only a small training set (Andrea & Kalayeh 1991), the better structure-activity relationships obtained by So and Richards

(1991) and Andrea and Kalayeh (1991) for other classes of DHFR inhibitors used considerably larger data sets.

In summary, neither QSAR technique has been able to produce a useful correlation with biological activity. Therefore they have been unable to determine the significance of the triazene function and its terminal substituents for *P. carinii* DHFR inhibitory activity. With structural knowledge of the active site it may be possible to determine how these pyrimethamine analogues align and interact in the active site. From this, the importance of the triazene function for *P. carinii* DHFR inhibitory activity may be resolved.



Inhibitor	Volume of R ¹ (Å ³)*	Volume of R ² (Å ³)*	Charge on N3 (e)	Charge on C7 (e)	Charge on C8 (e)
AM90020	22.63	22.63	-0.238	-0.087	-0.047
AM90022	37.50	37.50	-0.241	-0.041	-0.001
AM90024	51.50	22.63	-0.226	-0.009	0.035
AM90029	37.50	22.63	-0.228	-0.110	-0.015
AM90035	94.63	22.63	-0.226	-0.100	-0.001
AM90063	45.13	45.13	-0.223	-0.077	-0.053
AM90066	22.63	45.13	-0.219	-0.063	-0.024
AM90067	94.63	45.13	-0.222	-0.063	-0.024
AM90068	37.75	45.13	-0.230	-0.062	-0.116
AM90069	51.50	45.13	-0.239	-0.046	-0.031
AM90070	125.82	22.63	-0.233	-0.124	-0.042
AM90071	104.29	22.63	-0.181	-0.107	-0.004
AM90072	99.00	22.63	-0.228	-0.048	-0.023
AM90073	45.13	80.00	-0.221	-0.096	-0.009
AM90074	316.79	1.6	-0.224	0.104	0.063
AM90023	39.32	39.32	-0.215	-0.063	-0.024
AM90026	35.69	35.69	-0.237	-0.081	-0.051
AM90034	31.38	31.38	-0.276	-0.045	0.023
AM90036	60.07	60.07	-0.216	-0.044	0.023
AM90038	44.07	44.07	-0.214	-0.063	-0.024
AM90064	43.75	43.75	-0.226	-0.052	-0.016

Table 5.1 The triazene substituent volumes and the partial charges used as training parameters for the neural network QSAR studies. *For the heterocyclic compounds the substituent volumes were calculated as half the heterocycle volume.

Inhibitor	actual RL/PC ratio	calculated RL/PC ratio	difference
AM90020	6.75	6.6	0.15
AM90022	0.92	1.3	-0.38
AM90024	0.47	0.1	0.37
AM90029	14.3	15.6	-1.3
AM90035	0.82	0.43	0.39
AM90063	7.7	8.3	-1.1
AM90066	10.3	10.1	0.2
AM90067	21	20.6	0.4
AM90068	0.37	1.6	-1.23
AM90069	2.48	0.86	1.62
AM90070	0.56	1.1	-0.54
AM90071	0.33	0.70	-0.37
AM90072	0.18	0.28	-0.10
AM90073	2.9	1.8	1.1
AM90074	0.28	0.32	0.04
AM90023	0.1	1.8	-1.7
AM90026	15.3	17.8	-2.5
AM90034	0.30	0.11	0.19
AM90036	0.4	1.1	-0.7
AM90038	0.44	1.4	-0.96
AM90064	1.5	1.3	0.2

Table 5.2 Experimental and calculated rat liver to *P. carinii* (RL/PC) DHFR IC₅₀ inhibitory activity ratio of the 3'-(3",3"-disubstitutedtriazene-1"-yl) analogues of pyrimethamine.

Chapter 6 Conformational Analysis of 2,4-Diamino-6-ethyl-5-(4'-chloro-3'-(3'',3''-dimethyl)triazene-1''-yl)phenylpyrimidine (AM90020) and 2,4-Diamino-6-ethyl-5-(4'-chloro-3'-(3''-hydroxyethyl-3''-benzyl)triazene-1''-yl)phenylpyrimidine (AM90067)

6.1 Introduction

The inhibitory activity of a ligand is dependant on its ability to retain a conformation that maximises the number of favourable interactions when docked in the active site of the target enzyme (Dolata *et al.* 1990). Although it is very probable that the ligand adopts a low energy conformation it may not be the crystal structure conformation. This can be highlighted by the comparisons of the crystal structure conformations of the DHFR inhibitors methotrexate and trimethoprim adopted when bound in the active site of *E. coli* and mouse DHFRs (Bolin *et al.* 1982, Stammers *et al.* 1987) respectively and in the free base crystal structures (Sutton & Cody 1986, Koetzle & Williams 1976). Thus, prior to docking studies, conformational analysis of the inhibitor is necessary to determine the possible low energy forms.

Computational chemistry calculations now have a major role in conformational analysis, because the more powerful computers are capable of rapidly generating and calculating the stability of a great number of molecular conformations. However, with an exponential increase in the number of conformations required to be computed with each increase in the number of rotatable bonds, there is limited use for more developed molecular modeling packages because of the demands these place on computing resources. Therefore the molecular mechanics concept (see 2.1) has remained the method of choice for the computer-assisted determination of thermodynamically stable conformations in flexible compounds. To reduce the number of conformers to be generated for the more flexible molecules, White and Kitson (1986) suggested that the rotatable bonds should be considered separately. Then, once individual stable conformations have been established, the minimum energy conformations should be evaluated for each torsion in turn with respect to each other. By considering separately the different moieties of a molecular structure that are unlikely to interfere with each other, then analysing the low energy conformers as an entire molecule allows further reduction in the computational demands (Chalmers *et al.* 1992). The success of

predicting low energy inhibitor conformations has been demonstrated by Folkers *et al.* 1991 who were able to propose rational explanations for the apparent selectivity of acyclovir for *Herpes simplex* virus kinase. Sansom *et al.* (1989), using the MM2 molecular mechanics force field to study the pyrimethamine analogue *meta*-azidopyrimethamine, discovered four minimum energy conformations. Subsequently, when docked in the active site of *E. coli* DHFR, the enhanced inhibitory activity of MZP could be explained by a favourable interaction with a particular residue when orientated into one of these minimum conformations. Because of the simplicity of the molecular mechanics concept, there has remained doubt about the accuracy of these calculations on more complex chemical systems. This has led to the limited use of semiempirical methods to determine and verify conformational data (Ruiz *et al.* 1993). Biological (Chan & Roth 1991) and NMR (Birdsall *et al.* 1990) experimental data have been used to obtain conformational information about an enzymatically bound ligand. Monitoring the effect a restricted ring system has on the biological activity compared to the non-restricted inhibitor can lead to suggestions of the binding conformation. However, unexpected interactions with the ring atoms in the active site may hide the implications of conformational data as demonstrated by Chan and Roth (1991). Another problem with using this type of data with DHFR inhibitors is that even a seemingly minor change in chemical structure can have a profound effect on biological activity. This was highlighted by the utilisation of the 6-ethyl group in 3'-chloropyrimethamine as a C6D-C2P dimethylene bridge (Rosowsky & 1989; Broughton and Queener 1991) (figure 5.1) which caused a hundred-fold increase in inhibitory activity.

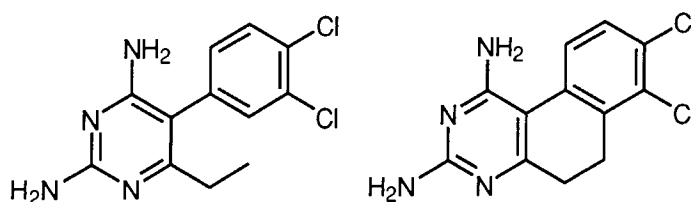


Figure 6.1 Structures of 3'-chloropyrimethamine (left) and its 6,2'-dimethylene bridged tricyclic analogue (right)

As mentioned previously, the biological activity of an organic compound is closely related to its conformation. Hence, the structural information concerning how an inhibitor orientates itself may be applied to its binding in the active site of a target enzyme. Because macromolecular systems significantly increase computational

demands, a reduction in the number of bound conformers that require evaluation will decrease the amount of central processing unit time necessary for the conformational energy calculations.

Once it has been established it becomes necessary to assume that some of these low energy conformations determined will be adopted in the active site. CADD visualisation techniques can provide rational explanations about the binding of structurally related compounds and can predict the suitability of any chemical modifications.

6.2 Materials and Methods

The aim of this investigation was to elucidate low energy conformations for AM90020 and AM90067 that could be used as starting coordinates to dock the compounds into the active site models of human and *P. carinii* DHFR. Thus, prior to all the calculations N1D (figure 6.2) was protonated and assigned a formal charge of 1.0.

Computationally, it was impractical and expensive to investigate every possible low energy torsion angle for a particular bond with respect to the other torsion angles. Because the hydroxyethyl and benzyl substituents of AM90067 are very flexible components a more systematic approach was used to determine thermodynamically stable conformations. Primarily, this involved regarding the triazene and its terminal substituents separately from the pyrimethamine fragment. For the triazene substituents, because the bonds investigated were saturated, a molecular mechanics treatment was used. However, the triazene moiety could be conjugated with the benzene ring. Therefore the more sophisticated molecular orbital methods were appropriate to investigate the C2P-C3P-N1-N2 and N1-N2-N3-C8 torsions. Then, in turn each torsion was orientated into these minimum energy conformations with respect to each other and optimised using the semi-empirical method suggested for phenyltriazenes from chapter 3. Because the pyrimidine and benzene rings of pyrimethamine and its analogues are known to be twisted enough with respect to each other such that conjugation is unimportant (Sansom *et al.* 1989), molecular mechanics methods were used to determine the low energy conformations for this moiety. Finally, the lowest energy conformations determined for the 3'-(3'',3''-disubstitutedtriazene-1''-yl) and 2,4-diamino-6-ethyl-5-(4'-chloro)phenylpyrimidine moieties of AM90020 and AM90067 were used as starting points to determine possible global minimum energy

conformations.

6.2.1 Sources of Chemicals

AM90020 (2,4-diamino-6-ethyl-5-(4'-chloro-3'-(3'',3''-dimethyl)triazene-1''-yl)phenylpyrimidine)

AM90067 (2,4-diamino-6-ethyl-5-(4'-chloro-3'-(3''-hydroxyethyl-3''-benzyl)triazene-1''-yl)phenylpyrimidine)

Aston Molecules Ltd., Gosta Green, Birmingham, West Midlands.

DMSO (D6)

Deuterium Oxide

Goss Scientific Instruments, Brentwood, Essex.

6.2.2 Computational Software

Conformational calculations using the molecular mechanics force field COSMIC (Morley *et al.* 1990) were accessed through the molecular modeling package NEMESIS (Oxford Molecular Ltd., Oxford) mounted on a Apple Macintosh IICx. The semi-empirical quantum mechanics calculations were completed using MOPAC (Stewart 1990a), accessed via molecular modeling package ChemX (Chemical Design Ltd., Chipping Norton) mounted on a VAX/VMS 8650 mainframe computer.

6.2.3 MOPAC Calculations

For these calculations the AM1 Hamiltonian and the keywords "CHARGE=1" and "GEO-OK" were specified. The calculations were then submitted to a batch queue. When the torsion angle N1-N2-N3-C7 was rotated, the restraints were removed from the bond angles N2-N3-C7, N2-N3-C8 and C7-N3-C8 (see figure 6.2 for atomic labelling).

6.2.4 MOPAC Optimisations

As well as using the AM1 parameters and the same keywords, the following bond distances, bond angles and torsion angles were restrained to their experimentally determined geometries using a weighting factor of 10 :-

bond distances - C3P-N1, N1-N2 and N2-N3;

bond angles - C2P-C3P-N1, C4P-C3P-N1, C3P-N1-N2, N1-N2-N3,
N2-N3-C7, N2-N3-C8 and C7-N3-C8;

torsion angles - C2P-C3P-N1-N2, C4P-C3P-N1-N2, C3P-N1-N2-N3,
N1-N2-N3-C7 and N1-N2-N3-C8.

When the torsion angle N1-N2-N3-C7 was rotated the restraints were removed from the bond angles N2-N3-C7, N2-N3-C8 and C7-N3-C8.

6.2.5 Construction of Protonated Model Compounds

For AM90020, in ChemX, the ethanesulphonate anion and the isopropyl alcohol molecule of crystallisation were removed from the crystal structure of AM90021 and the cation then saved as a .cssr file. These coordinates were used to construct AM90067, for which the atom names are shown in figure 6.2. From the ChemX fragment library methylamine and benzene were appended, with standard bond lengths, to C7 and C8 respectively of the model protonated AM90020 structure. The methylamine nitrogen atom was “retyped” as oxygen and one of the hydrogen atoms removed. The final coordinates were then saved to a .cssr file. The coordinates for the protonated AM90020 and protonated AM90067 that were used in subsequent conformational calculations were obtained by optimisation of these structures as described in section 6.2.4.

6.2.6 Conformation Generation and Energy Calculation

6.2.6.1 Rotation of the Torsions C4D-C5D-C1P-C2P and C5D-C6D-C61D-C62D

Under the “CONFORMATIONS” menu in NEMESIS the title torsion angles were selected, conformations generated in 20° increments with respect to each other and the Molecule Energy calculated. The results were then plotted on a contour map.

6.2.6.2 Rotation of the Torsion C2P-C3P-N1-N2

MOPAC calculations, as described in section 6.2.3, were used to establish the heat of formation (ΔH_f). Rotamers were generated by increasing the title torsion angle by increments of 15°, using the “ALTER GEOMETRY” facility in ChemX, for each

calculation. For each rotamer a full optimisation (6.2.4) was also completed.

6.2.6.3 Rotation of the Torsion N1-N2-N3-C7

As for 6.2.6.2

6.2.6.4 Rotation of the Torsions N2-N3-C7-C1H and N2-N3-C8-C1B

As for 6.2.6.1

6.2.6.5 Rotation of the Torsion N3-C7-C1H-O1H

Under the “CONFORMATIONS” menu in NEMESIS the title torsion angle was selected, conformations were generated in 5° increments and the ΔH_f calculated.

6.2.6.6 Rotation of the Torsion C7-C1H-O1H-H10H

As for 6.2.6.5

6.2.6.7 Rotation of the Torsion N3-C8-C1B-C2B

As for 6.2.6.5

6.2.7 MOPAC Optimisations of Low Energy Conformations

Using the “ALTER GEOMETRY” facility in ChemX, initial coordinates of AM90067 were generated that placed the torsions N1-N2-N3-C7, N2-N3-C7-C1H, N2-N3-C8-C1B and N3-C7-C1H-O1H, with respect to each other, in each of the low energy minima previously determined. The optimisation calculation was identical to that described in 6.2.4 with the additional keyword “PRECISE” specified.

6.2.8 Superimposing the Conformers of AM90020 and AM90067

In the “FITTING” option in the main menu in ChemX the atoms N21D, N41D, C61D and Cl1 from each conformer were matched with a weighting constant of 10. “RIGID FIT” was then selected to superimpose the two structures.

6.2.9 NMR Conformational Studies on AM90067

For NMR data at 293 K AM90067 was dissolved in DMSO (D6). The NMR spectra

were recorded on a Bruker 250-AC spectrometer at ^1H (250.1 MHz) and ^{13}C (62.9 MHz). ^1H and ^{13}C were referenced to trimethylsilane.

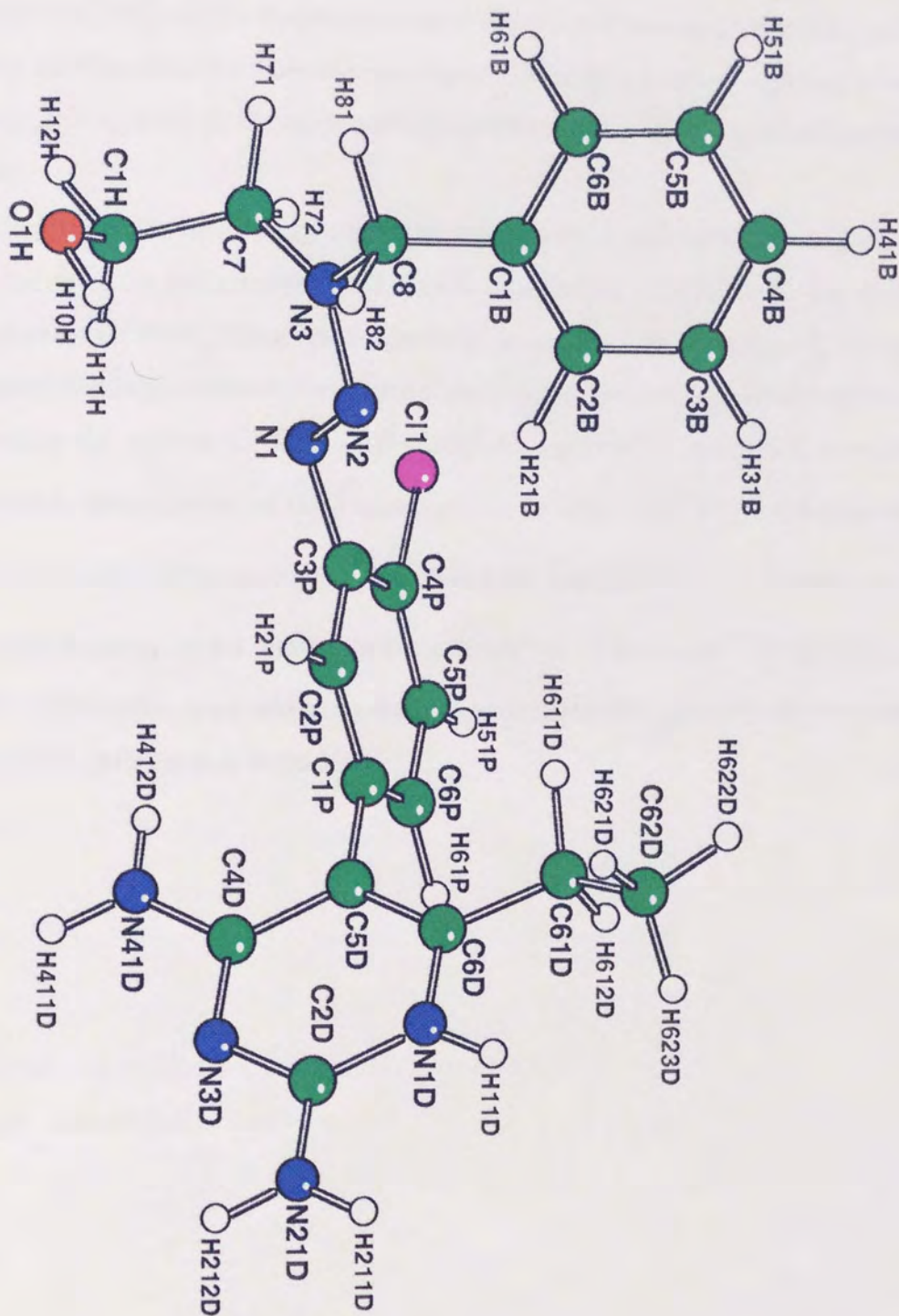


Figure 6.2 Atomic Numbering for Protonated AM90067

6.3 Results and Discussion

It must be emphasised that these conformations were calculated on isolated molecules *in vacuo*. However, the biological testing of these compounds occurred in aqueous solution at 37°C. At this temperature many of the bond torsions investigated can rotate freely and therefore the final calculated geometries should not be regarded as isolated points, but as some of the more probable conformations the compounds are likely to adopt.

To determine the low energy conformations from the rotation of the torsion angles exclusive to the hydroxyethyl and benzyl substituents of AM90067, the molecular mechanics COSMIC forcefield is the most practical. This is because it can rapidly generate the large number of conformations required to study these flexible moieties. Rotating the torsion C7-C1H-O1H-H10H through 360° (figure 6.3) produces, as expected, three minima of -18.3 kcal mol⁻¹, -17.5 kcal mol⁻¹ and -18.8 kcal mol⁻¹ at 60°, 180° and 300° respectively. The rotational barriers of 1.5 - 4 kcal mol⁻¹ and increased energy of the minima in the order of 0.1 - 1 kcal mol⁻¹ for the torsion C7-C1H-O1H-H10H, agree with accepted values determined experimentally for similar C-O torsions (Morrison & Boyd 1992)

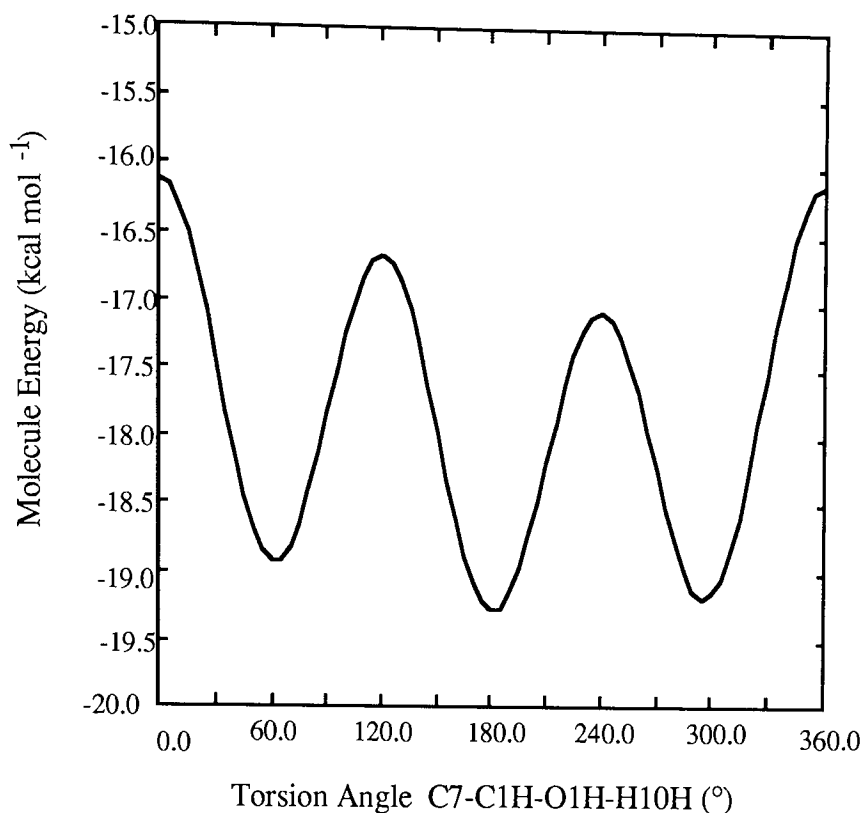


Figure 6.3 Change in the molecule energy of AM90067 as the torsion angle C7-C1H-O1H-H10H rotates through 360 °

For the torsion N3-C7-C1H-O1H, the very large Molecule Energy (1925 kcal mol⁻¹) is caused by O1H orientating to within 1.27 Å of N1 (figure 6.4a). This interatomic contact is responsible for the shift of the minimum energy conformation to 75° from the expected 60° (figure 6.4b). The three minima observed at 75°, 180° and 290° produce Molecule Energies of -19.0 kcal mol⁻¹, -19.4 kcal mol⁻¹ and -14.1 kcal mol⁻¹, as illustrated by their respective Newman projections (figure 6.5).

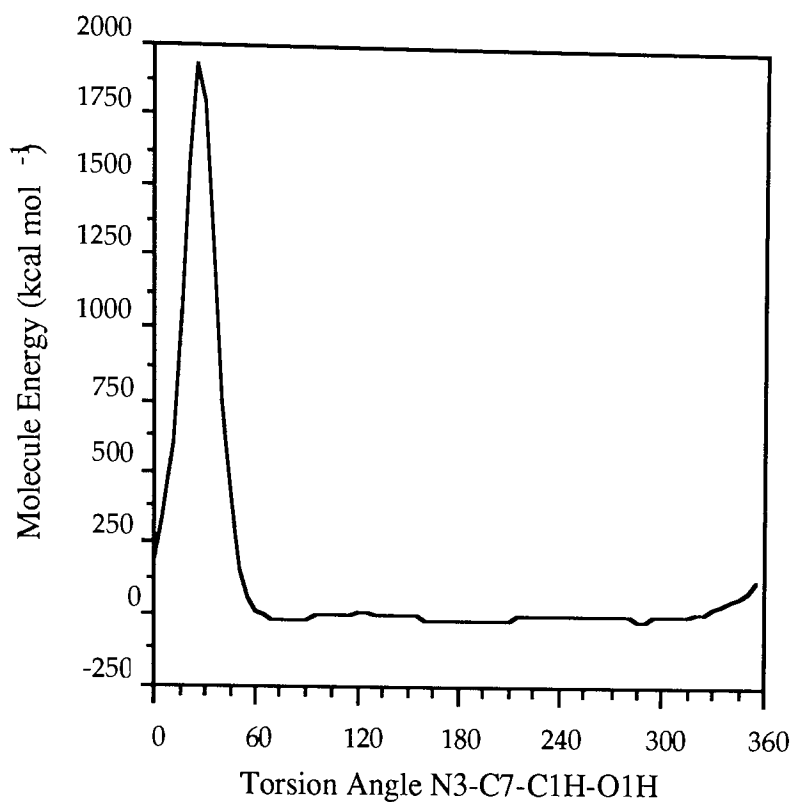


Figure 6.4a Change in the molecule energy of AM90067 as the torsion angle N3-C7-C1H-O1H rotates through 360°

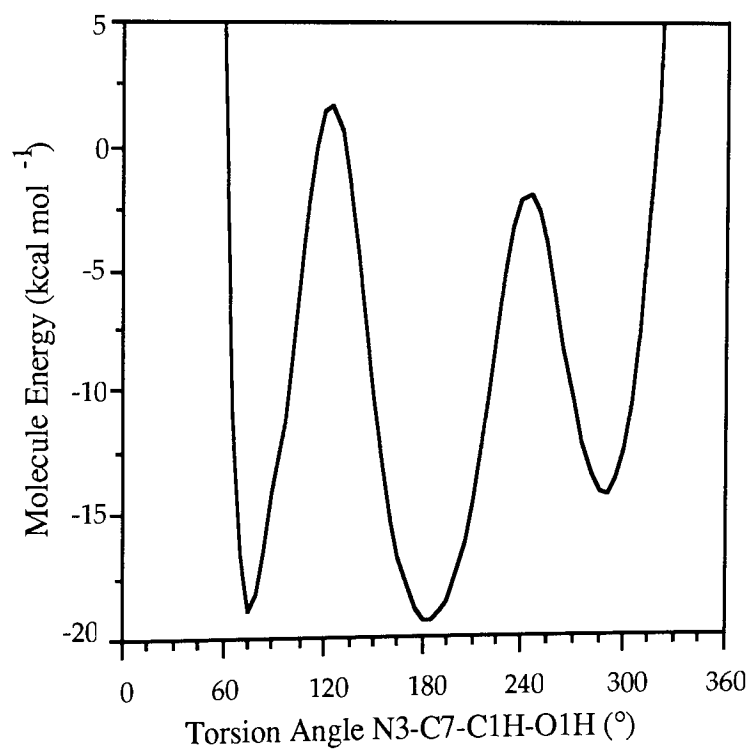


Figure 6.4b Change in the molecule energy, between -20 and 5 kcal mol⁻¹, of AM90067 as the torsion angle N3-C7-C1H-O1H rotates through 360°

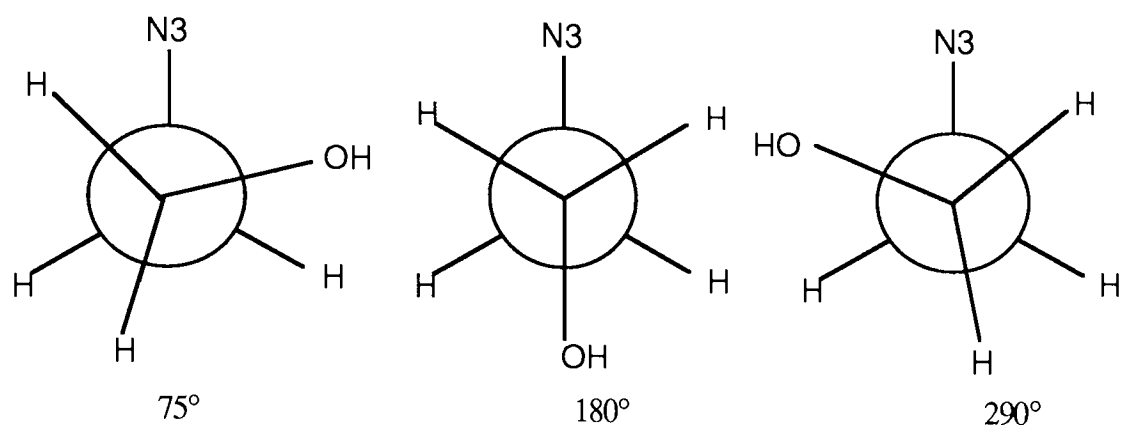


Figure 6.5 Newman projections of the low energy conformations as the torsion angle N3-C7-C1H-O1H rotates through 360°

Figure 6.6 shows that rotation of the torsion N3-C8-C1B-C2B through 360° produces two very distinct minima at 165° and 345°. Because of its shape and its planarity the benzene ring possesses 180° rotational symmetry. Therefore the two minima observed by rotating this torsion, which rotates the aromatic nucleus, correspond to the same conformation. The Newman projection in figure 6.7 illustrates this. The larger rotational barriers, 66.4 kcal mol⁻¹, strongly suggest that rotation, compared with the other torsions, is far more restricted, hence there is a greater probability for the torsion to remain in this conformation.

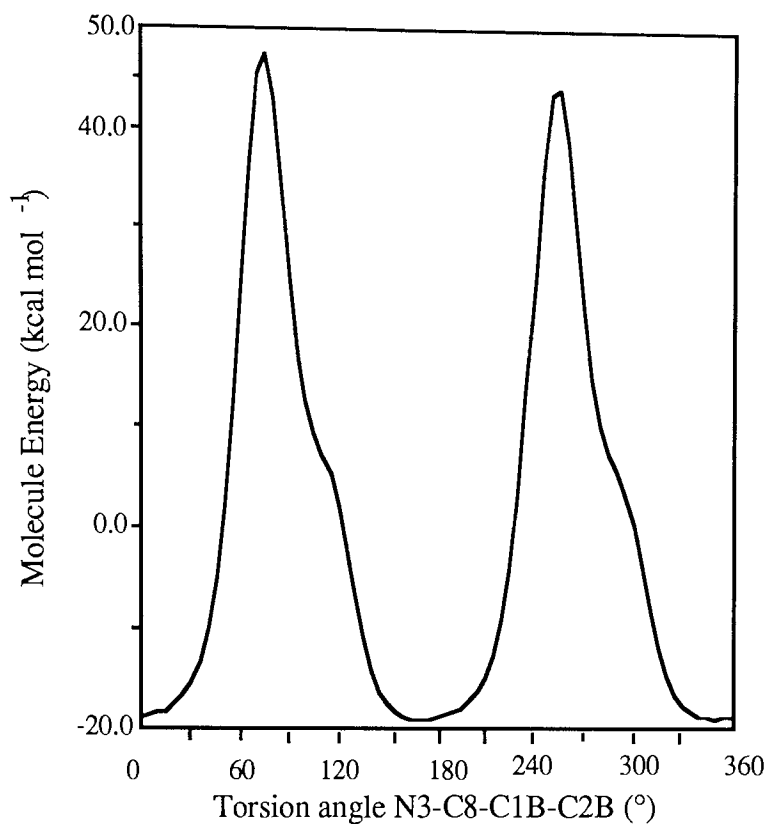


Figure 6.6 Change in the molecule energy of AM90067 as the torsion angle N3-C8-C1B-C2B rotates through 360°

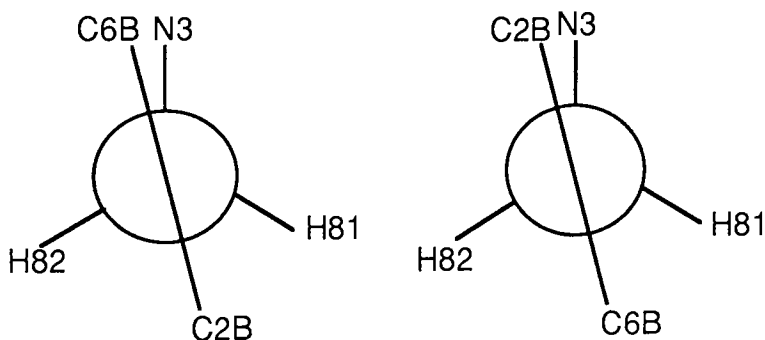


Figure 6.7 Newman projections along the bond C1B-C8 of the two low energy conformations

The two remaining torsions unique to AM90067, N2-N3-C7-C1H and N2-N3-C8-C1B, were rotated through 360° with respect to each other to consider the effect that the orientation of the triazene substituents had on the Molecule Energy. The contour plot (figure 6.8) shows four minima. It also shows areas of high Molecule Energy when the aforementioned torsion angle values are in the range of 90° - 270° and 165° - 280°

respectively, in particular when these torsion angle ranges overlap. This large region of high Molecule Energy is caused by certain atoms, in particular, C1H of the hydroxyethyl and C1B of the benzyl substituents, orientating in close proximity (to within 2.1 Å) to each other. The high Molecule Energy values observed when the torsions are approximately 180° and 60° is caused by a close contact between C1H and H81 (2.114 Å).

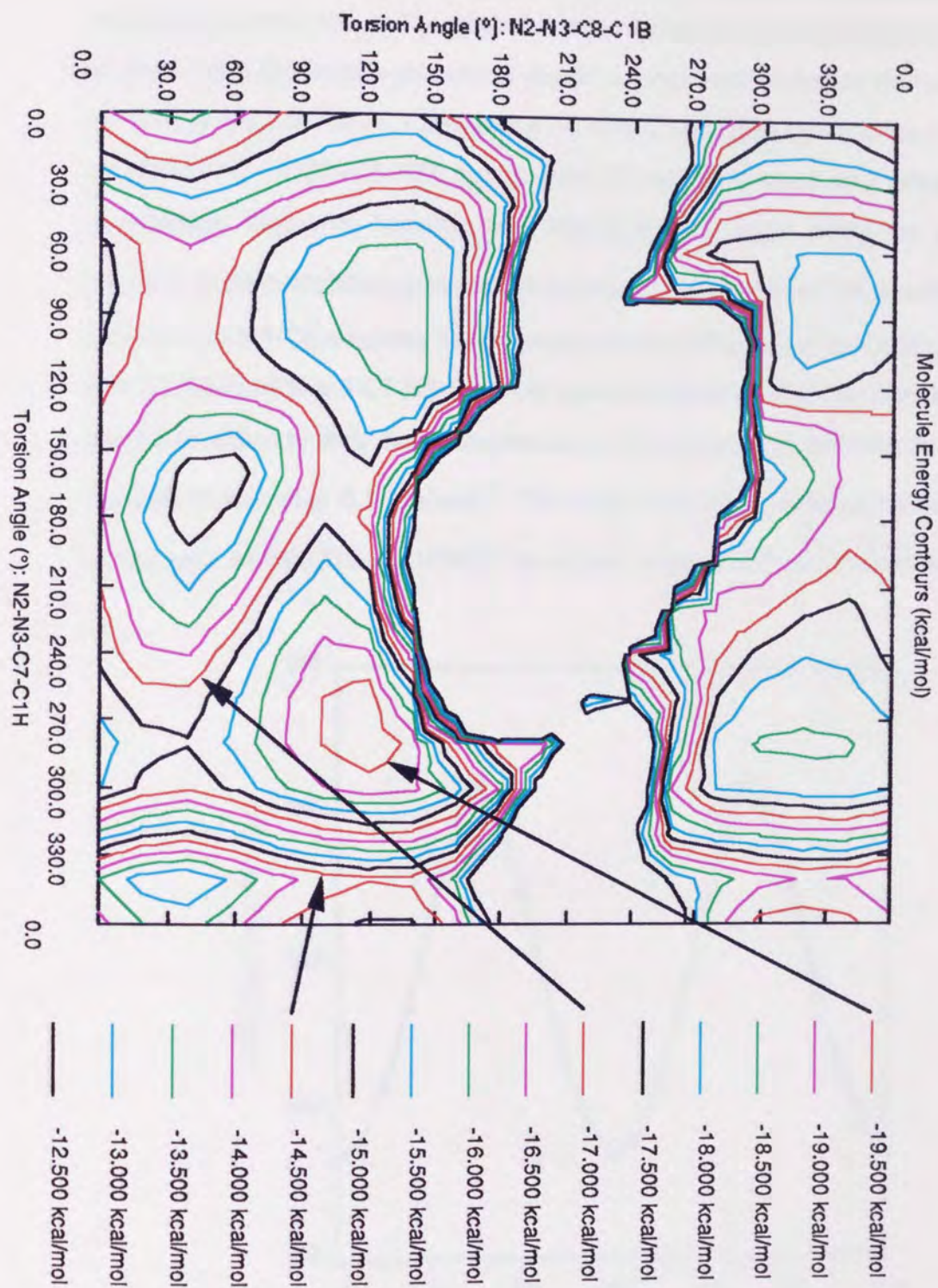


Figure 6.8 Energy contour plot showing the calculated Molecule Energy of AM90067 when the torsion angles N2-N3-C7-C1H and N2-N3-C8-C1B are rotated through 360° with respect to each other

Results from chapter 3 suggest that the AM1 parameters for semiempirical calculations and optimisations of phenyltriazenes are the most reliable. However, there are limitations in their use, and to minimise inconsistencies with experimental data it is recommended that certain geometries should be restrained, including the bond angles N2-N3-C7 and N2-N3-C8. Considering N2,N3,C7 and C8 to be coplanar, rotation of the torsion N1-N2-N3-C8 180° should cause C7 and C8 to adopt each other's nuclear coordinates. However, because these angles are not equal, using the suggested restraints in the calculation produces a discrepancy of 4.1 kcal mol⁻¹ between the 0° and 180° N1-N2-N3-C8 rotamers. To overcome this the restraints on the bond angles N2-N3-C7, N2-N3-C8 and C7-N3-C8, were removed which allowed the positions of C7 and C8 to adjust to their atomic environment. The discrepancy between the rotamers reduced to less than 0.1 kcalmol⁻¹. The results from this modified calculation for compounds AM90020 and AM90067 are shown in figures 6.9 and 6.10 respectively.

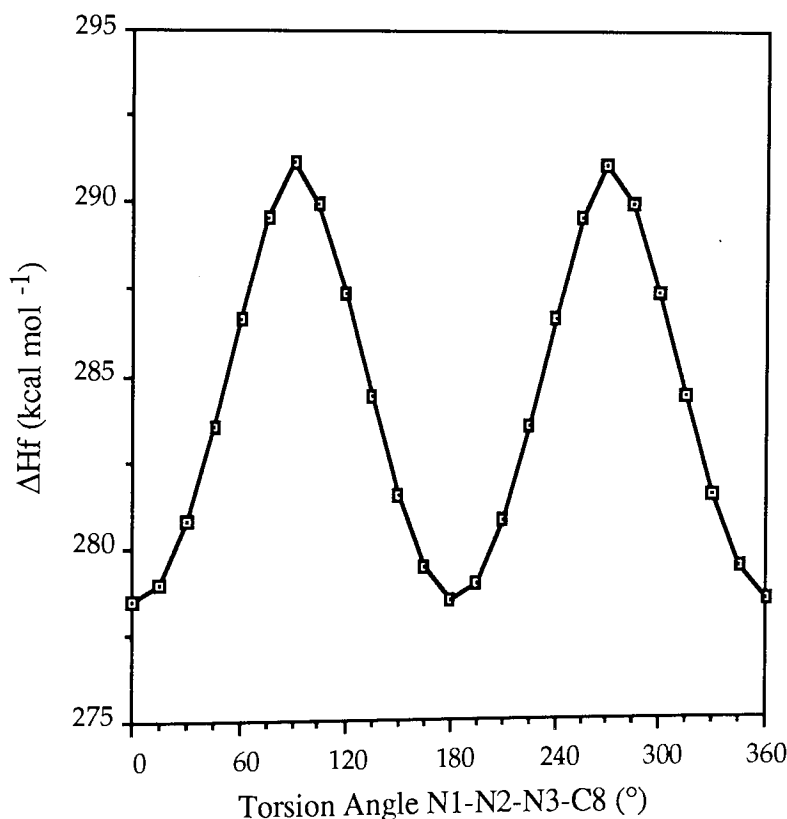


Figure 6.9 Changes in the calculated heat of formation (ΔH_f) in AM90020 as the torsion angle N1-N2-N3-C8 rotates through 360°

The two distinct minima observed for AM90020 at 0° and 180°, with a rotational barrier of 14.3 kcal mol⁻¹, are in good agreement with NMR data (Lippert *et al.* 1992) and MNDO calculations (Ramos & Periera 1986) for other phenyltriazenes. Rotation of the equivalent torsion in AM90067 produced one definitive minimum at 165° and a second, less obvious local minimum at 330°. Repetition of these calculations with the restraints on the following torsions removed; N2-N3-C7-C1H, N3-C7-C1H-O1H, C7-C1H-O1H-H10H, N2-N3-C8-C1B and N3-C8-C1B-C2B, did not cause any significant reduction in the calculated ΔH_f compared to the original values. Furthermore, comparison of the equivalent geometries between 285° and 360° did not reveal any structural alteration. Hence any doubt that intramolecular clashes are responsible for the larger than expected ΔH_f 's was eliminated. Therefore, it appears that physical differences between the terminal hydroxyethyl and benzyl groups on the triazene moiety of AM90067 are responsible for this preferred conformation with the hydroxyethyl group *cis* to N1.

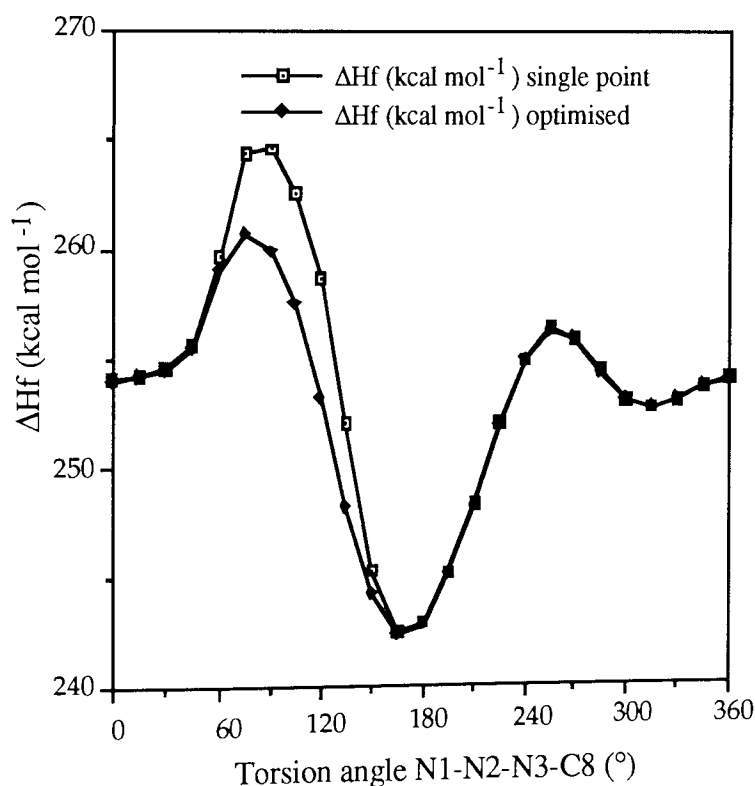


Figure 6.10 Change in the calculated heat of formation (ΔH_f) of AM90067 as the torsion angle N1-N2-N3-C8 rotates through 360°

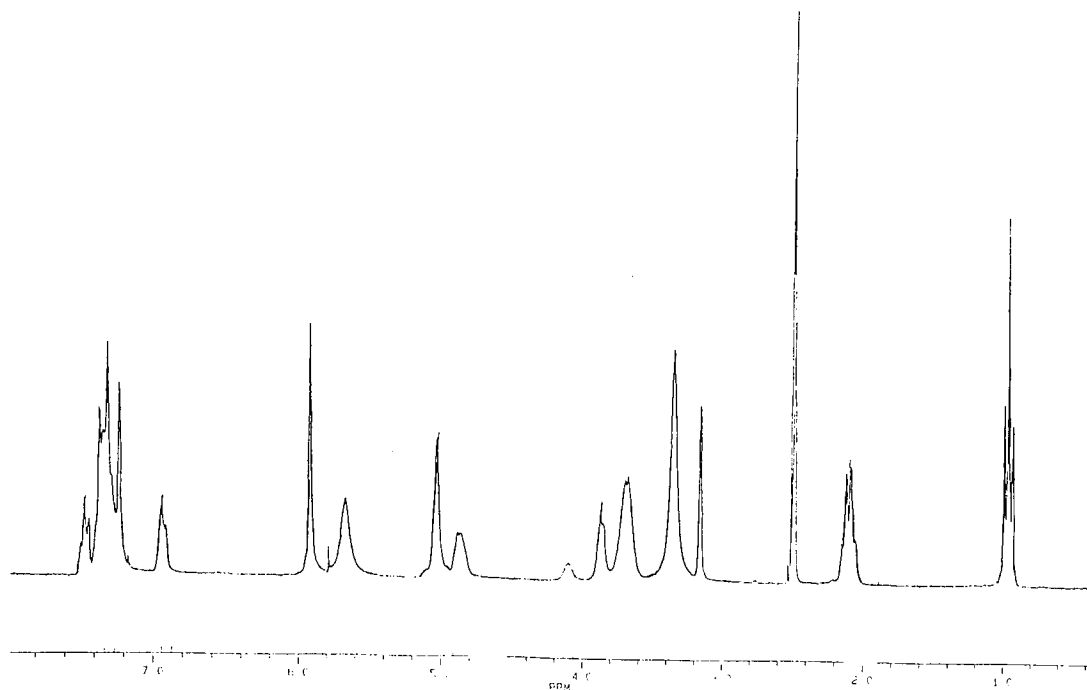
Removal of the restraints does reduce the rotational energy barrier from 22.7 to 18.2 kcal mol⁻¹. Table 6.1 shows that a possible reason for this reduction is the relief of steric clashes between O1H and N1 and O1H and C11 produced by rotation of the N2-N3-C7-C1H torsion from 69.4° to a maximum of 83.4°. The rotational energy barrier of 18.2 kcal mol⁻¹ agreed with NMR data (Lunazzi *et al.* 1978) that bulkier terminal triazene substituents did not have any profound effect on the barrier magnitude.

N1-N2-N3-C8 torsion angle (°)	single point N1 --- O1H distance (Å)	optimised N1 --- O1H distance (Å)	single point C11 --- O1H distance (Å)	optimised C11 --- O1H distance (Å)	N2-N3-C7-C1H torsion angle (°)
45	3.09	3.11	2.98	2.99	69.4
60	2.81	2.88	2.65	2.73	75.2
75	2.58	2.70	2.59	2.86	80.7
90	2.39	2.59	2.80	3.17	81.8
105	2.30	2.55	3.22	3.69	83.2
120	2.31	2.59	3.76	4.25	83.4
135	2.42	2.66	4.34	4.72	80.8
150	2.62	2.76	5.12	4.92	76.6
165	2.89	2.89	5.46	5.49	72.2

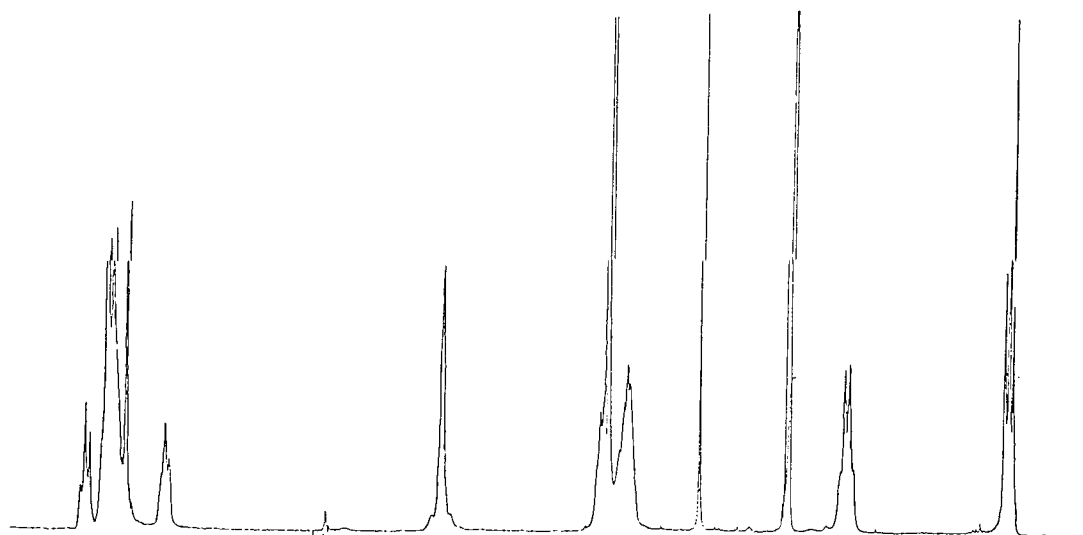
Table 6.1 Selected optimised and single point geometries of N1-N2-N3-C8 rotamers of AM90067

From the ¹H NMR spectrum (figure 6.11) measured in DMSO solution non-equivalent benzyl methylene protons, appearing as a double doublet and broad peaks corresponding to hydroxyethyl methylenes in the ratio of 3:1 instead of the more likely 1:1 ratio may be a result of conformational restriction. Exact assignment of the methylenes is prevented partly because of the relatively poor resolution from DMSO dissolution.

(a)



(b)



(c)

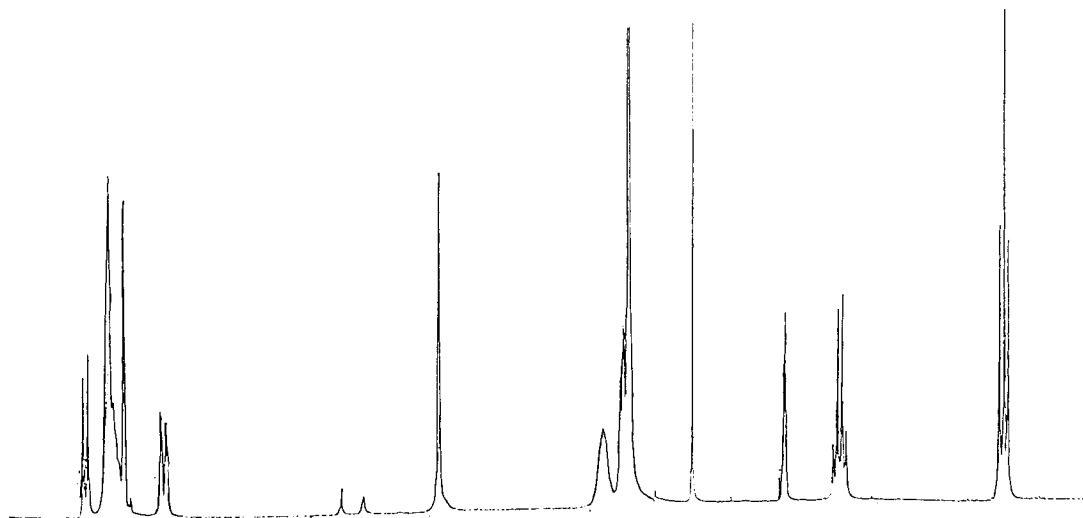


Figure 6.11 ^1H NMR spectra of AM90067 in DMSO at (a) 298 K and D_2O exchanged (b) 298 K and (c) 320 K.

Chemical Shift (δ /ppm)	Multiplicity	Integral	<i>J</i> (Hz)	Assignment
0.94 - 1.00	triplet	3	7.5	-CH ₂ -CH ₃
2.10 - 2.13	quartet	2	7.5	-CH ₂ -CH ₃
3.67 - 3.87	multiplet	4		-CH ₂ -CH ₂ -OH
4.87	singlet	1		-CH ₂ -OH
5.03	collapsed double doublet	2		-CH ₂ -C ₆ H ₅
5.66	broad singlet	2		aromatic-NH ₂
5.91	broad singlet	2		aromatic-NH ₂
6.9 - 7.5	multiplet	8		aromatic

Table 6.2 ¹H NMR data for AM90067 at 298 K in DMSO before D₂O exchange (figure 6.11a)

Further and more conclusive evidence of rotational hindrance around the N1-N2-N3-C7 torsion is inferred from the proton decoupled ¹³C NMR data (figure 6.12 and tables 6.3 and 6.4). The DEPT (Distortionless Enhancement of Polarisation Transfer) spectrum is able to assist in the inference by transforming the methylene carbon peaks to orientate down from the baseline and by removing the non-substituted aromatic carbon peaks. Firstly the aromatic peaks observed are far more complex than anticipated suggesting that some of the aromatic carbons are magnetically non-equivalent, reflecting more than one conformational isomer. Also, doubling of the peaks corresponding to the non-equivalent hydroxyethyl and benzyl methylene carbons is observed consistently with the major peak more downfield and approximately twice as large as the corresponding minor peak. This suggests that these atoms exist in two different magnetic environments. The presence of consistently larger and smaller peaks would suggest the presence of a major and a minor conformation.

Chemical Shift (δ /ppm)	Assignment	Chemical Shift (δ /ppm)	Assignment
14.11	CH_2CH_3	129.22	aromatic C
28.17	CH_2CH_3	129.60	aromatic C
51.40	$\text{CH}_2\text{CH}_2\text{OH}$	131.41	aromatic C
57.38	$\text{CH}_2\text{C}_6\text{H}_5$	135.74	aromatic CH
60.17	$\text{CH}_2\text{CH}_2\text{OH}$	136.96	aromatic CH
106.72	aromatic CH	147.58	aromatic CH
121.24	aromatic C	162.36	aromatic CH
127.99	aromatic C	162.83	aromatic CH
128.12	aromatic C	163.26	aromatic CH
128.81	aromatic C	163.36	aromatic CH
129.04	aromatic C	167.08	aromatic CH

Table 6.3 ^{13}C NMR data for AM90067 at 298 K

Identity	Chemical Shift (δ) at 313 K	Chemical Shift (δ) at 353 K
CH_3	12.97	12.66
CH_2	27.27	27.27
aromatic CH	105.61	105.91
aromatic C	120.43	120.63
aromatic C	126.97	127.05
aromatic C	-	127.99
aromatic C	128.37	128.25
aromatic C	130.21	130.17
aromatic CH	135.28	135.32
aromatic CH	136 and 137	136.5
aromatic CH	146.64	146.91
aromatic CH	161.80	-
aromatic CH	161.90	161.88
aromatic CH	166.07	166.34

Table 6.4 ^{13}C NMR chemical shift data for AM90067 at 313 K and 353 K

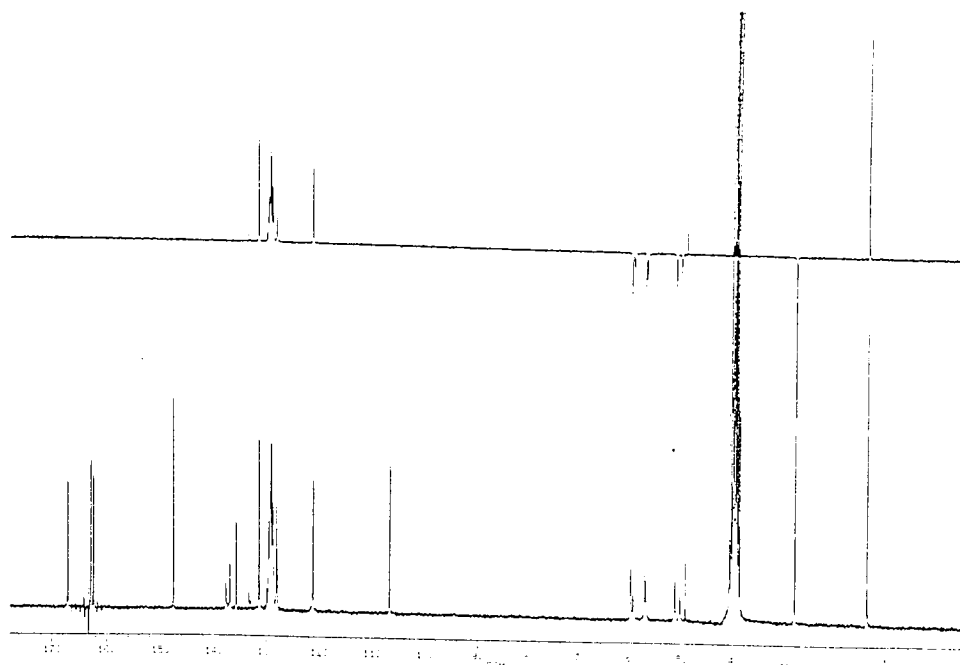
More conclusive evidence of conformational hindrance of the N1-N2-N3-C7 torsion of AM90067 was obtained from comparison of the ^{13}C NMR spectrum recorded at room temperature (293 K) with spectra recorded at 313 K and 353 K. Comparison of the ^1H NMR spectrum recorded at 293 K with the spectrum recorded at 320 K was also informative.

At 313 K and 353 K considerable simplification in the aromatic region of the ^{13}C NMR spectra was observed. The number of aromatic peaks observed between δ 128 and δ 130 at 293 K reduced at higher temperatures. The double peak observed at 293 K between δ 136 and 137 collapsed to a broad peak at 313 K and coalesced to give a single sharp peak at 353 K. Finally, the three peaks observed at δ 162.36, δ 162.83 and δ 163.26 collapsed to two peaks at 313 K. At 353 K this peak was detected as a singlet, though more detailed examination of the spectrum would suggest the presence of a second peak too close to determine accurately a chemical shifts. In the upfield aliphatic region, three methylene doublet peaks were observed at 293 K. At 313 K the benzylic CH_2 collapsed to a single peak and the other two were closer together. Only a broad single peak at δ 59 was observed at 353 K. Therefore it appears that by increasing the temperature it becomes possible for AM90067 to overcome rotational barriers. This free rotation released certain moieties from non-equivalent conformations, resulting in the simplified ^{13}C NMR spectra. The coalescing of the double methylene peaks strongly suggests that at 353 K there was now sufficient energy for the N1-N2-N3-C7 torsion to overcome the energy barrier and rotate freely.

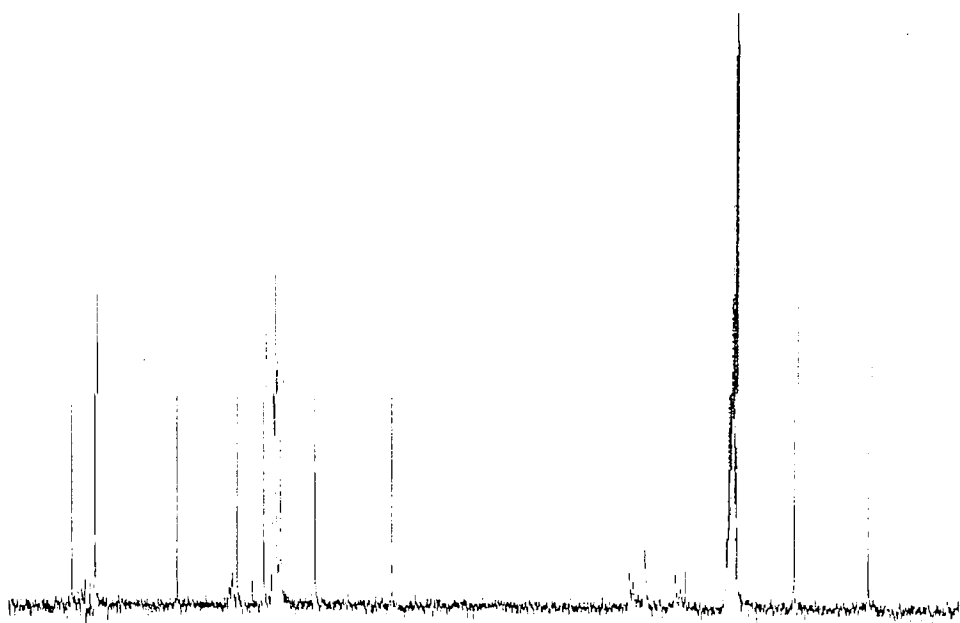
It is difficult to obtain entirely conclusive conformational data from this limited NMR study. However at 320 K the aromatic peaks at δ 6.90 and δ 7.46 in the D_2O -exchanged spectrum of AM90067 in DMSO were better resolved. The increased temperature can reduce the number of restrictive orientations of these protons. The fewer orientations caused the peaks to become sharper and better resolved. The benzylic methylene double doublet at 293 K collapsed to give a single sharp peak at 320 K. This may be because these protons were magnetically equivalent at the higher temperature because they were no longer restricted to two conformations.

It is probable that the singlet at δ 3.16 and the broad, D₂O exchangeable peak at δ 4.10 were due to methyl alcohol impurities in the sample. This would also account for the sharp positive peak at δ 49.38 in the ¹³C NMR (figure 6.12).

(a)



(b)



(c)

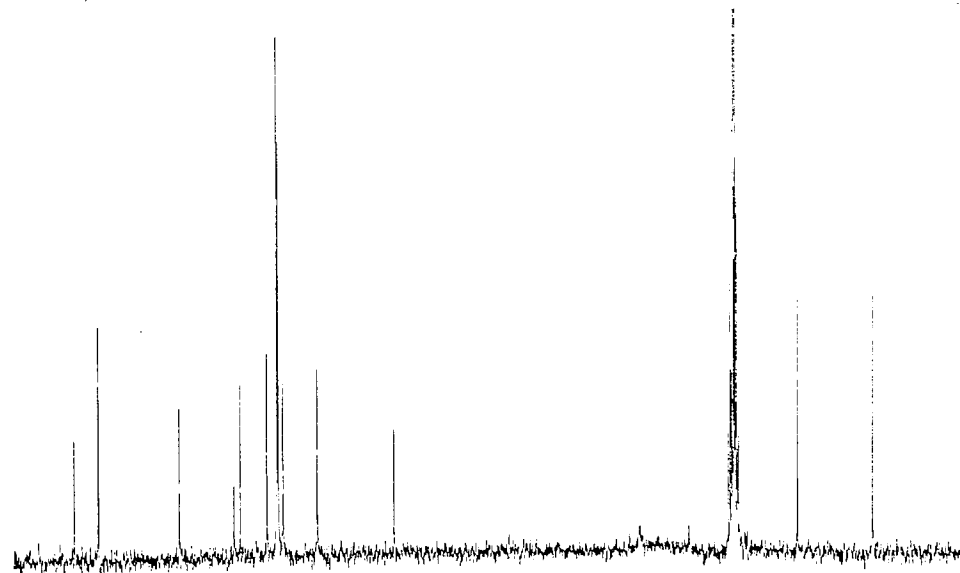


Figure 6.12 ^{13}C NMR spectra of AM90067 in DMSO at (a) 353 K, (b) 313 K and (c) 298 K. At 298 K the DEPT. spectrum is also displayed

The theoretical conformational data calculated for the terminal triazene substituents are summarised in table 6.5. These results are restricted to one or two rotating bonds only. To appreciate better the effect the orientation of these torsions has with respect to the entire molecular structure the protonated AM90067 model was positioned, in turn, into each of the determined minima and optimised as described in 6.2.7. To ensure the optimisation procedure did not direct the geometry toward the global minimum energy conformation the keyword "PRECISE" was specified. Only one of the three minima determined by the rotation of the C7-C1H-O1H-H10H was investigated because, once docked into the human and *P. carinii* DHFR active site models, this torsion would be rotated to determine if there are any potential hydrogen bond interactions. This could not be predicted from the optimisation. A C7-C1H-O1H-H10H torsion angle of 300° was selected because it produced the lowest energy conformation and it positioned the hydroxyl hydrogen away from the benzyl ring substituent when these two groups adopted orientation in close proximity to each other.

The optimised torsion angles and the ΔH_f 's are listed in table 6.5. Generally, the ΔH_f values for conformers 1 to 12 were lower than the equivalent conformers in 13 - 24. Conformers 8 and 20 had larger ΔH_f values, but this is caused by the close proximity of the hydroxyethyl and benzyl substituents in these conformations. Unexpected observations were made. When the torsion angle N3-C7-C1H-O1H had an initial geometry of 75° it orientated to between 57 and 59°. Orientation into this minimum energy conformation is permitted by the N2-N3-C7H-C1H torsion angle moving considerably from its initial geometry or because the initial geometry did not place O1H in close proximity to N1. Likewise, when the N3-C7-C1H-O1H torsion was 300° and therefore in some of the initial geometries O1H was in close proximity to N1, the N2-N3-C7-C1H torsion was observed to adjust accordingly. Except for conformers 20 and 24, the torsion angle N3-C8-C1B-C2B was consistently between 117.4° and 123.5°, therefore disagreeing with the COSMIC calculations which predicted a value of 165°. This may be a result of differences in the parameter sets between the AM1 and COSMIC programs and/or the fact that the entire molecule was now being considered in the calculation and enabled the determination of a lower minimum energy conformation. To account for these differences, once AM90067 was docked, the effect of rotating N3-C8-C1B-C2B on the energy of interaction would be investigated.

TORSION ANGLES (°)											
	ΔHf (kcal mol ⁻¹)	N2-N3-C7-C1H		N2-N3-C8-C1B		N3-C7-C1H-O1H		N3-C8-C1B-C2B		C7-C1H-O1H-H10H	
		starting	final	starting	final	starting	final	starting	final	starting	final
1	247.12	60	67.7	300	301.2	75	57.8	165	121.1	300	276.4
2	240.05	60	71.6	150	158.9	75	58.2	165	121.2	300	298.2
3	246.58	60	78.1	300	300.8	300	296.8	165	121.1	300	296.6
4	248.21	60	78.2	150	159.0	300	296.6	165	121.2	300	298.4
5	246.68	60	71.6	300	300.4	180	177.5	165	120.9	300	298.4
6	248.02	60	71.6	150	176.9	180	177.4	165	121.1	300	290.1
7	266.27	300	284.7	300	300.2	75	57.4	165	121.1	300	293.7
8	274.64	300	283.6	150	158.5	75	55.9	165	121.2	300	288.8
9	249.19	300	304.4	300	300.1	300	297.4	165	121.1	300	297.1
10	253.79	300	304.3	150	159.0	300	296.8	165	121.2	300	298.4
11	250.56	300	303.4	300	300.2	180	177.4	165	121.1	300	295.5
12	275.27	300	311.3	150	160.8	180	177.6	165	121.3	300	298.4
13	260.32	60	63.7	300	290.8	75	57.6	165	123.3	300	298.6
14	257.12	60	68.9	150	176.4	75	58.2	165	121.1	300	295.5
15	259.67	60	69.7	300	290.9	300	297.1	165	123.3	300	304.1
16	256.39	60	69.9	150	176.7	300	297.1	165	121.1	300	303.8
17	258.55	60	68.8	300	293.6	180	177.6	165	122.5	300	298.4
18	255.67	60	68.5	150	175.2	180	178.4	165	120.8	300	298.4
19	256.56	300	303.9	300	290.6	75	59.3	165	123.2	300	281.1
20	303.80	300	304.6	150	170.9	75	59.3	165	97.6	300	279.2
21	259.13	300	310.9	300	290.6	300	297.6	165	123.3	300	298.5
22	262.80	300	310.6	150	169.4	300	297.6	165	117.4	300	298.4
23	260.66	300	310.1	300	290.6	180	177.5	165	123.3	300	300.6
24	275.05	300	301.1	150	173.7	180	179.6	165	92.8	300	298.4

Table 6.5 The calculated heat of formation and the original and optimised torsion angles for the terminal triazene hydroxethyl and benzyl substituents of AM90067 when the torsion angle N1-N2-N3-C7 was 0° (1 - 12) and 180° (13 - 24)

Figures 6.13 and 6.14 show the change in the calculated AM1 heats of formation for AM90020 and AM90067 respectively when the bond torsion C2P-C3P-N1-N2 was rotated 360°. The large rotational energy barrier observed, 12 kcal mol⁻¹, was caused by the close proximity of N2 to Cl1, 2.2 Å at 180°. Only minor discrepancies between the heat of formation for single point and optimised calculations for AM90067 confirmed that no other steric interactions were responsible for this barrier. For both AM90020 and AM90067 only one minimum energy value, at approximately 315°, was observed.

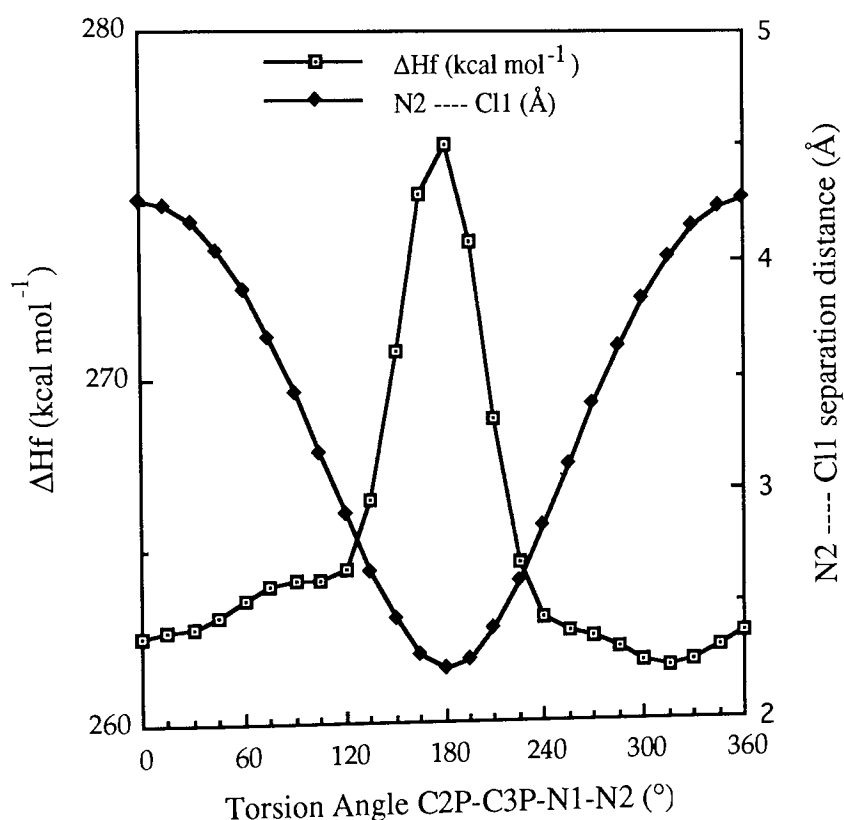


Figure 6.13 Changes in the calculated heat of formation in AM90020 and the N2 to Cl1 separation distance as the torsion angle C2P-C3P-N1-N2 rotates through 360°

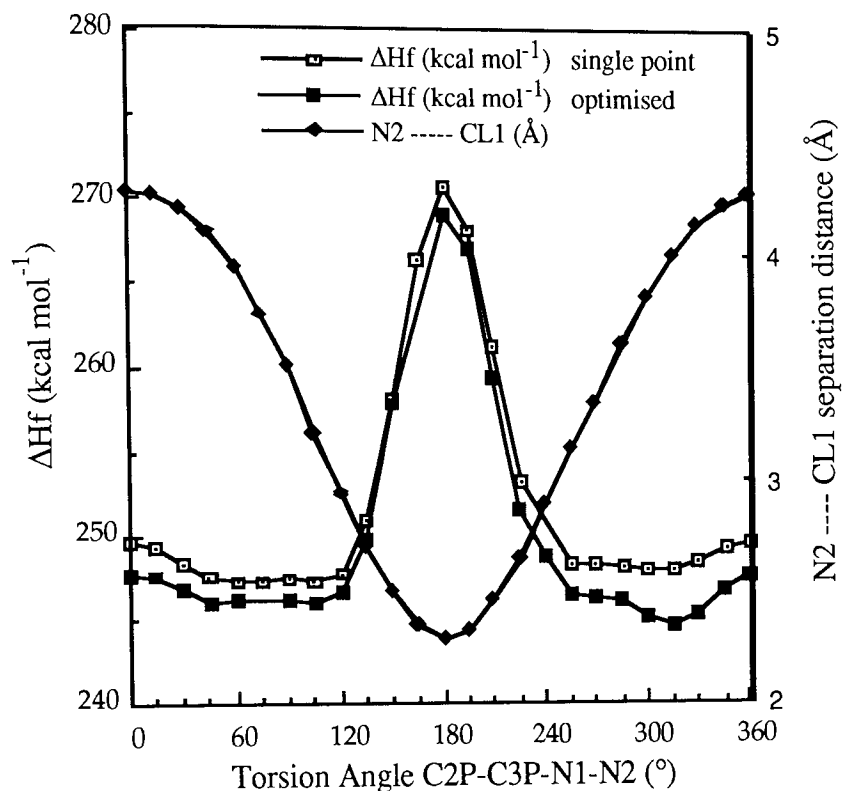


Figure 6.14 Change in the calculated heat of formation of AM90067 as the torsion angle C2P-C3P-N1-N2 rotates through 360°

The contour plots obtained by rotating the torsions C6D-C5D-C1P-C2P and C5D-C6D-C61D-C62D with respect to each other for AM90020 and AM90067 are shown on figures 6.15 and 6.16 respectively. For AM90020 four distinct low energy conformations occur when the torsion angle values are either +90° or -90°. This agrees very favourably with similar MNDO optimisation calculations for another pyrimethamine derivative, MZP (Sansom *et al.* 1989). Likewise, for AM90067, four low energy conformations are observed, though they are not as distinct. This is indicated by fewer contours separating the minima.

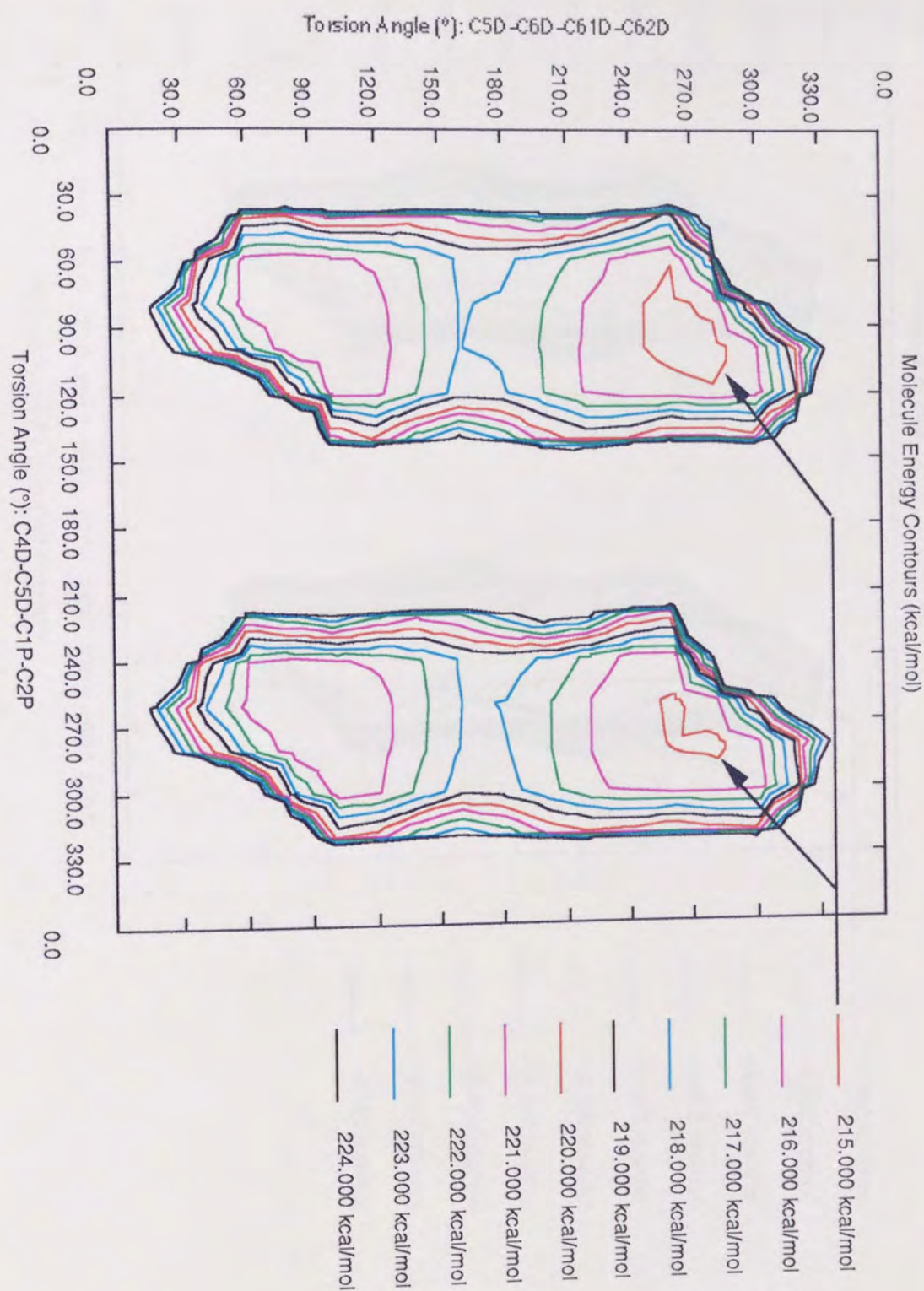


Figure 6.15 Energy contour plot showing the calculated Molecule Energy for AM90020 when the torsion angles C6D-C5D-C1P-C2P and C5D-C6D-C61D-C62D are rotated through 360° with respect to each other

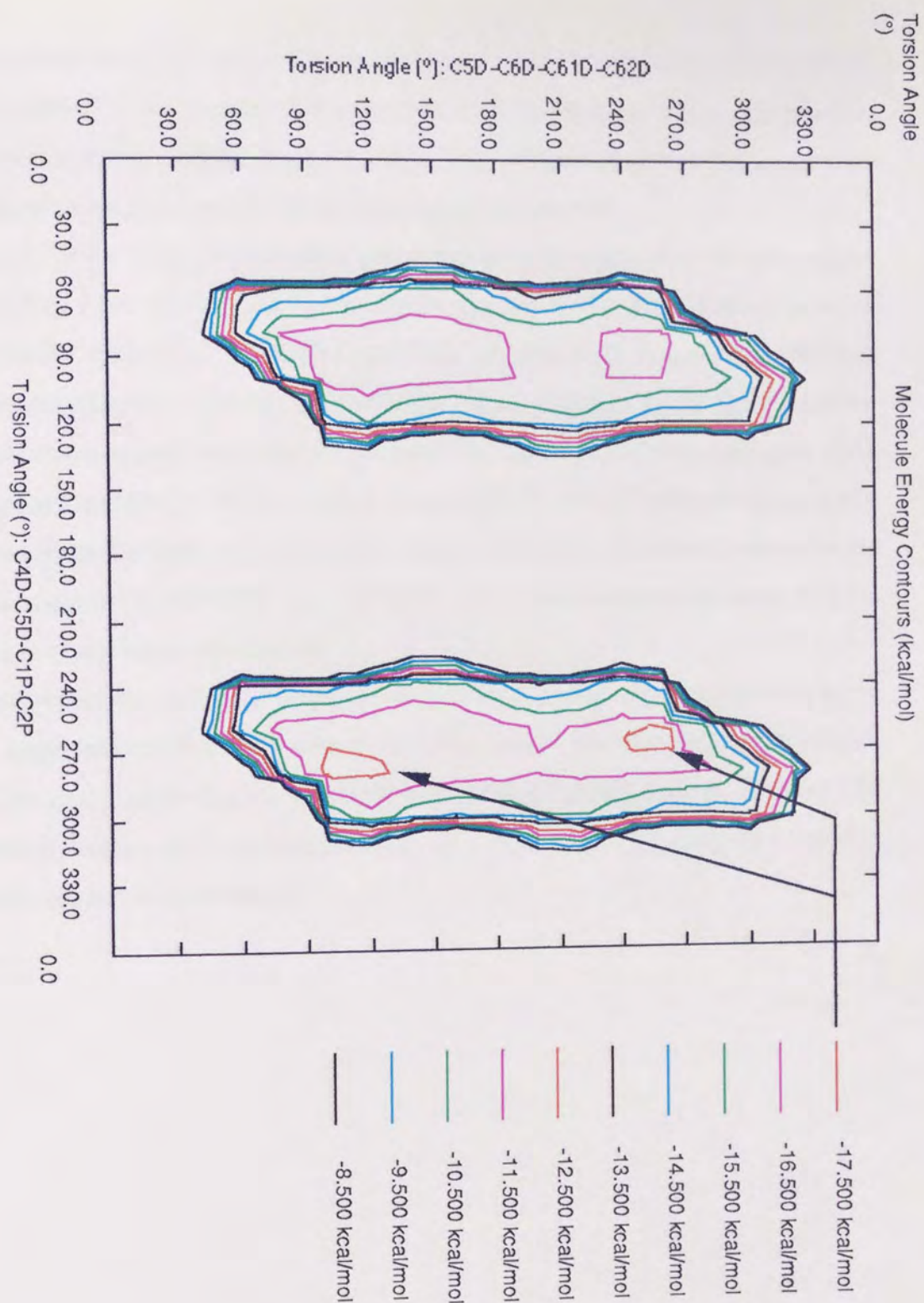


Figure 6.16 Energy contour plot showing the calculated Molecule Energy for AM90067 when the torsion angles C6D-C5D-C1P-C2P and C5D-C6D-C61D-C62D are rotated through 360° with respect to each other

To complete this conformational study, the model structures of protonated AM90020 and AM90067 were positioned in these four conformations, and then optimised as described in section 6.2.5. Also, for AM90067, the terminal triazene substituents were orientated into conformation 2 (table 6.5) prior to optimisation.

As expected the initial and optimised geometries of hydroxyethyl and benzyl moieties of AM90067 are very similar. However, for the AM90020 and AM90067 torsions C5D-C6D-C61D-C62D and C4D-C5D-C1P-C2P, the AM1 calculations produce geometries that are not exactly perpendicular (tables 6.6 and 6.7) and show a greater resemblance to crystal structures of pyrimethamine and pyrimethamine analogues. AM1 does produce C2P-C3P-N1-N2 torsion angles (300.5° - 305.4°) that are significantly different from the RMS crystal structure value (337.6°). The similar ΔH_f values for the conformations of AM90020 and AM90067 show that energetically there is little difference between conformations.

In conclusion, the molecular modeling conformational determination methods agree with experimental data and give a valuable insight into the three dimensional arrangements that the ligands AM90020 and AM90067 prefer to adopt. These can be applied to docking studies as biological activity is dependant on the ability of a ligand to maintain an active conformation.

ΔH_f (kcal mol ⁻¹)	<u>Torsion Angles(°)</u>					
	C5D-C6D-C61D-C62D		C4D-C5D-C1P-C2P		C2P-C3P-N1-N2	
	original	optimised	original	optimised	original	optimised
267.35	90	71.4	90	103.5	315	300.5
267.32	90	74.3	270	255.2	315	300.8
267.80	270	276.5	90	98.6	315	301.3
267.31	270	282.7	270	254.0	315	301.6

Table 6.6 The calculated heat of formation and selected original and optimised torsion angle values for AM90020 after optimisation as described in section 6.2.5

ΔH_f (kcal mol ⁻¹)	<u>Torsion Angles(°)</u>					
	C5D-C6D-C61D-C62D		C4D-C5D-C1P-C2P		C2P-C3P-N1-N2	
	original	optimised	original	optimised	original	optimised
245.34	90	93.0	90	78.9	315	303.0
244.44	90	94.3	270	278.9	315	303.3
245.66	270	264.6	90	278.3	315	303.2
244.51	270	266.9	270	81.5	315	305.4

Table 6.7 The calculated heat of formation and selected initial and optimised torsion angle values for AM90067, after optimisation as described in section 6.2.5

Chapter 7 Computer-Aided Active-Site Modeling of AM90067 in Human Liver and *Pneumocystis carinii* Dihydrofolate Reductase

7.1 Introduction

The advent of recombinant DNA base sequencing techniques and subsequent translation in the appropriate reading frame has caused an explosion in the number of determined protein primary structures. However the amino acid sequence alone reveals very little information concerning the three dimensional structures. X-ray crystallography is the most important method for obtaining accurate tertiary structural data. This is, however, an expensive and time consuming process, requiring crystalline protein samples. Relationships between the primary and secondary structures of proteins have been made by study of the available protein crystal structures (Garnier *et al.* 1978, Chou & Fasman 1974 a & b). However, predicting the tertiary structure directly from the primary sequence remains a distant target (Sali & Blundell 1993). Polypeptide regions showing the greatest structural conservation are the hydrophobic core and the residues implicated in the active site (Chothia & Lesk 1986). These regions also showed the greatest sequence homology (Havel & Snow 1991). It is this conservation of structure of the biologically most important regions of series of homologous proteins that has been exploited in the development of a knowledge-based approach to protein structure prediction (Blundell *et al.* 1988). Browne *et al.* (1969) were the first to construct an accurate model of the polypeptide α -lactalbumin by homology modeling using the crystal structure of lysozyme as a template. Since then, with the introduction of more sophisticated computational software and very high resolution graphics, this model building process has been made easier. Greer (1981) devised rational procedures to construct a model peptide backbone. Since then, improvements have made protein model building more complete and more uniform procedure, with important applications in drug design (Havel & Snow 1991). Advancements in molecular modeling software are directing homology model construction towards an entirely automated process, as outlined in figure 7.1 This is a graphical representation of the logical flow of procedures which could readily be incorporated into a computer molecular modeling package (Sali *et al.* 1990). However, in practice intuitive decisions in the design process are still required.

Content has been removed for copyright reasons

Figure 7.1 Scheme for the knowledge-based modeling of homologous proteins. The approach involves the derivation of rules from the comparison of the amino acid sequences and three dimensional structures and their use in the generation of a template to construct the three dimensional model (Sali *et al.* 1990).

7.1.1 Sequence Homology Alignment

The alignment and matching of equivalent residues in two or more amino acid sequences has been used to classify proteins, study evolutionary development at a molecular level (Doolittle 1989), and determine regions of structural conservation. Studies showed that these regions occurred in clusters (Doolittle 1981). Therefore the

development of algorithms capable of aligning sequences (Needleman and Wunsch 1970) became an effective method to determine these regions of structural homology. The most recent improvements in sequence homology alignment, still relying on the same fundamental principles of the Needleman and Wunsch algorithm, have introduced more sophisticated methods to accept "imperfect" matches (Argos 1987, Saqi & Sternberg 1991) and process more rapidly multiple alignments (Higgins & Sharp 1988). Multiple alignments of homologous proteins have been used to identify structurally conserved residues which can then be used as templates for the initial alignments of protein structures (Blundell *et al.* 1988). Three-dimensional data for these regions are used to form a structural template from which modeling of the regions of poor homology can be inserted.

7.1.2 Modeling Regions with Low Structural Homology

As previously stated the clusters of residues with greatest structural and sequence homology are located in the hydrophobic core. Thus the major structural deviations, the insertions and deletions between the clusters of sequence homology, tend to occur in the surface loops. Original methods relied on sequence homology between loops from different protein structures (Greer 1981). Although sequence homology may be a good indication of structural conservation, it has been established that identical oligopeptide sequences can exist in different conformations and secondary structures (Kabsch & Sander 1984). Since then there have been attempts to classify the peptide loops by analysing the conformation adopted of loops from highly resolved protein X-ray structures (Sibanda *et al.* 1989, Edwards *et al.* 1987) and relating this to distance separating the two terminal template residues and the amino acid sequence from the unknown structure (Singh & Thornton 1990). More recently the development of algorithms has enabled modeling of the protein loop regions without relying solely on crystallographic data (Zheung *et al.* 1993). Most of the classification has been completed for loop regions containing no greater than seven residues. Because of the immense conformational freedom that larger loops can have, without very good sequence homology it is impossible to predict accurately or rationally the conformation. Loop regions are usually conformationally very flexible. Hence, they are unlikely to be involved in the orientation of active site residues. Therefore, for studies involving the

binding of inhibitors known to bind in the active site, detailed analysis of the loop conformation is not necessary, as demonstrated by Folkers *et al.* (1991).

7.1.3 Prediction of Amino Acid Side Chain Conformations

The over-riding principle of homology-based modeling is to minimise the number of assumptions. Therefore, wherever possible, conformations for conserved residues are used. Investigation of side chain conformations in the hydrophobic core showed that the vast majority of residues appeared to exist in defined conformational orientations (Ponder & Richards 1987). This can provide starting conformations which can be evaluated in the completed model as to their suitability. Another knowledge-based approach was introduced by Sutcliffe *et al.* (1987) which superimposed the conformation of the residue from the known onto the unknown, using lower weighting restraints where the side chain differed. Subsequent developments have enabled accurate amino acid side chain conformation predictions from segment matching (Levitt 1992) and the utilisation of side chain rotamer libraries (Farid *et al.* 1991).

7.1.4 Docking and Calculation of Interaction

To compete with an enzyme substrate in the active site an inhibitor must make energetically-favourable non-covalent interactions to overcome steric repulsions and rotational freedom. These interactions include hydrogen bonds, ionic bonds and van der Waals forces. Where possible, docking an inhibitor extends the knowledge-based theme of utilising known data concerning the substrate binding and applying it to the inhibitor. Hence, similar or common chemical moieties are used to superimpose the ligand onto the docked substrate. Because of the immensity of the calculation, molecular mechanics based calculations are used to quantify the binding energy of the ligand.

Using this methodology produces a measure of the potential energy with regard to the defined atomic positions. It can be split into component van der Waals and electrostatic forces. Van der Waals forces are derived from the Lennard-Jones potential :-

$$E = c \left[\left(\frac{r_0}{r} \right)^{12} - 2 \left(\frac{r_0}{r} \right)^6 \right] \quad \text{Equation 7.1}$$

where E is the potential energy, c the constant that contains the force terms, r the

separation of two non-bonded atoms and r_0 the separation at the lowest potential energy. An electrostatic component of the interaction can be described as the attraction between two point charges, the energy being given by :-

$$E = \frac{q_i q_j}{Dr_{ij}}$$

Equation 7.2

where q_i and q_j are the partial charges on atoms i and j , r_{ij} is the distance between them and D is the dielectric constant for the medium in which interaction occurs.

For these complex calculations it is doubtful that the results are sufficiently reliable given the approximations and assumptions involved in determining the interacting forces. Their potential for detecting trends in the binding energy in various conformations has a great significance in deducing possible explanations for biological activity and the generation of new lead compounds.

7.2 Materials and Methods

7.2.1 Computational Software

Visualisation and manipulation of the protein structures was carried out using the molecular modeling package Quanta 3.2 (Molecular Simulations Inc., Burlington, MA 01803) mounted on a Silicon Graphics IRIS Indigo work station (Silicon Graphics, Inc., Mountain View CA94039-7311). The energy of interaction between the inhibitors and the enzyme was calculated using the CHARMM (Brooks *et al.* 1983) force field parameters available interfaced through Quanta 3.2.

7.2.2 DHFR Sequence Homology Studies

Twenty complete DHFR sequence files were taken from SWISSPROT, GenBank and EMBL protein and nucleic acid sequence databases (table 7.1) They were accessed via the SERC laboratories at Daresbury, U.K. using the University of Wisconsin Genetics Computers Group software package available under seqnet. The sequences were then edited and merged into one text file correctly formatted to be read by the multiple sequence alignment program CLUSTAL (Higgins & Sharp 1988). The alignments were made using the following parameters :-

Fixed Gap Penalty	10
Floating Gap Penalty	10
Protein Weight Matrix	PAM250 (Dayhoff <i>et al.</i> 1983)

7.2.3 Construction of a Human Liver DHFR Active Site Model

7.2.3.1 Secondary Structure Assignment

The two human DHFR structures, with the bound folate or 5-deazafolate (Davies *et al.* 1990), were taken from the Brookhaven Protein Database (Bernstein *et al.* 1977) available at the SERC laboratories, Daresbury, U.K. Both structures exist as dimers. Therefore, using a text editor, the atomic coordinates for each individual DHFR structure were written to separate Protein Database (.pdb) formatted files. The corresponding molecular structure files (.msf) were then generated using Quanta. "Analyse Secondary Structure" was chosen from the "Protein Design" function in the "Applications" menu of Quanta. For determining the secondary structure the

parameters were set as follows :-

With no Hydrogens - Maximum N-O distance	3.5 Å
With no Hydrogens - Minimum C-O-N angle	90°
Minimum α -carbon torsion for extended structure	-180°
Maximum α -carbon torsion for extended structure	-130°
Minimum α -carbon torsion for marginal extended structure	-130°
Maximum α -carbon torsion for marginal extended structure	-100°
Minimum α -carbon torsion for helical structure	30°
Maximum α -carbon structure for helical structure	70°

and the hydrogen bonds calculated.

7.2.3.2 Alignment and Superimposition of Protein Structures

“Align and Superpose” was selected from the “Protein Design” function. The following settings were used in “alignment weights” to align by secondary structures.

Sequence only	0.0
Secondary structure only	0.5
Distance only	0.5
Number of residues in window	1.0

The residues were then matched using this combined criteria and superimposition was restricted to α -carbons. The superimposition of the α -carbons was then repeated using the distance criteria only. For this the following parameters were used.

Weight to fix constrained and matched residues	20
Maximum distance score	10
Maximum distance cut-off	10

7.2.3.3 Preliminary Model Creation

Using the “Create Homology Model” facility in the “Protein Design” function an .msf file was created that was the average of the atomic coordinates from the four DHFR structures.

7.2.3.4 *meta*-Azidopyrimethamine and Water Inclusion

The α -carbons from the β -sheets of the mouse liver DHFR (Stammers *et al.* 1987) and the α -carbons from the active site model of *E. coli* DHFR with bound *meta*-azidopyrimethamine (Sansom *et al.* 1989) were superimposed onto the equivalent

atoms in the model human DHFR structure using the "Match Atoms" facility in the "**Molecular Similarity**" application. The atomic coordinates for all the atoms were saved to a .pdb file which was edited to retain the model atomic coordinates plus those for *meta*-azidopyrimethamine and the two water molecules that are important for the diaminopyrimidine active site binding (Denny *et al.* 1992).

7.2.4 Construction of a *P. carinii* DHFR Active Site Model

7.2.4.1 Creation of a Model of *P. carinii* DHFR

Under the "**Applications**" menu "Sequence Builder" was selected. From the file librarian "AMINO.RTF" was opened. National Biochemical Research Foundation (NBRF) single letter codes for amino acids for the *P. carinii* DHFR sequence (Kovacs *et al.* 1989) were entered and the file saved.

7.2.4.2 Creation of a Molecular Structure File of *P. carinii* DHFR from a Sequence File

From the "**Applications**" menu "Protein Design" was selected. From the *P. carinii* DHFR sequence file a molecular structure file (.msf) with undefined coordinates was generated.

7.2.4.3 Assignment and Alignment of the Secondary Structures of the Crystalline *Escherichia coli*, *Lactobacillus casei*, Mouse Liver and Chicken Liver DHFR Structures with the Model Human DHFR.

Molecular structure files (.msf) files were created for the corresponding *E. coli* (Bolin *et al.* 1982) *L. casei* (Bolin *et al.* 1982), mouse liver (Stammers *et al.* 1987) and chicken liver (Matthews *et al.* 1985) DHFR .pdb files. The secondary structures was assigned as described in section 7.2.3.1 and the structures superimposed upon the human DHFR structure according to section 7.2.3.2.

7.2.4.4 Alignment of the *P. carinii* DHFR Amino Acid Sequence to the Five Homologus Sequences

"Align and Superpose" was selected from the "Protein Design" function. The following

settings were used to align the *P. carinii* DHFR sequence to the five DHFR homologues :-

Sequence only	1.0
Secondary structure only	0.0
Distance only	0.0
Number of residues in window	10

7.2.4.5 Construction of the Peptide Backbone of the Active Site Model of *P. carinii* DHFR

Using the "Create Homology Model" in the "Protein Design" facility coordinates from the human DHFR model were copied to the .msf file created in 7.2.3.

7.2.4.6 Side Chain Conformation for *P. carinii* DHFR Active Site Residues with one or more Identical Matches with known DHFR Structures

For each DHFR structure the two adjacent residues and the matching residues were edited from their .pdb file and placed into a new .pdb file. After creating the equivalent .msf file, the tripeptide sequences were aligned in the "**Protein Design**" and the structures superimposed as described in sections 7.2.3.2. If necessary a larger gap penalty was used. In the "Create Homology Model" facility a tripeptide model was constructed from the model human DHFR peptide backbone using averaged coordinates from the side chain of the central residue. Where only one structure matched the *P. carinii* DHFR sequence then only these side chain coordinates were used. In the "**Molecular Similarity**" application the three α -carbons were used to superimpose the tripeptide model onto the equivalent residues in the model *P. carinii* DHFR. All the atoms were then saved to one .msf file and from the corresponding .pdb file the model tripeptide peptide atoms and the original side chain atoms in the model *P. carinii* DHFR were edited out. The .msf file created from this new .pdb file was then used for the insertion of the next set of side chain coordinates.

7.2.4.7 Side Chain Conformation of Residues Unique to *P. carinii* DHFR

Using "Create Homology Model" facility these residues in the *P. carinii* DHFR model

were mutated to their correct identity. The torsion angles in these new side chains were rotated to the median value of the same torsion angles in equivalent residues from the DHFR structures.

7.2.4.8 Exceptions to 7.2.4.6 and 7.2.4.7

7.2.4.8.1 Serine 24

The conformation for this residue was predicted after the inhibitor AM90067, in the first low energy conformations (table 5.1), had been docked into the active site. From the “**options**” menu “energy” and “continuous” were selected. The torsion N(ser24)-C α (ser24)-C β (ser24)-O γ (ser24) was rotated through 360° to determine the lowest energy conformation, which was retained for docking of the subsequent low energy conformations.

7.2.4.8.2 Arginine 59 and Lysine 60

For these two residues, despite other matches, only the mouse DHFR coordinates for the side chain atoms were used.

7.2.4.8.3 Valine 154

In all the DHFR structures the equivalent residue was threonine. Average atomic coordinates for threonine were inserted into the model *P. carinii* DHFR before the side chain oxygen was retyped to carbon.

7.2.4.9 meta-Azidopyrimethamine and Water Inclusion

As described in section 7.2.3.4 for human DHFR.

7.2.5 Docking of Pyrimethamine, AM90020 and AM90067 into the *P. carinii* and Human DHFR Active Site Models

In the “**Molecular Similarity**” application atoms C1D, N2D, C3D, N4D, C5D and C6D from AM90067 and the bound *meta*-azidopyrimethamine were matched with equal restraints. Using the “rigid fit” function AM90067 was superimposed onto *meta*-azidopyrimethamine. All the atomic positions were saved to one .msf file. The

equivalent .pdb file was created from which the *meta*-azidopyrimethamine atomic coordinates were removed. Under the “**Molecular Editor**” menu the calculated Gasteiger charges (Gasteiger 1980) were assigned to all atoms. Then, using the “Edit Atom” facility, AM1 partial charges were assigned to every atom calculated for each of the 24 low energy conformations predicted in chapter 5. The AM90067 torsion angles C5D-C6D-C61D-C62D and C4D-C5D-C1P-C2P were orientated to 90°.

7.2.6 Removal of Distant Residues

Under the “**Molecular Editor**” menu the “Active Atoms” facility was selected. All complete residues greater than 5 Å from the bound AM90067 were removed. This structure, with only the residues close to AM90067 was saved to a new .msf file.

7.2.7 Calculation of the Energy of Interaction

From the “**Options**” menu the facilities “Energy” and “Move Fragment” were selected to calculate an energy of interaction (ΔH_i). The following torsions; N(ser24)-C α (ser24)-C β (ser24)-O γ (ser24); N3-C8-C1B-C2B; and C7-C1H-O1H-H10H (see figure 5.1); were rotated separately, then orientated to the three lowest rotameric configurations to investigate the effect these had on the ΔH_i .

7.3 Results and Discussion

The residue numbers used in this section correspond to the *P. carinii* DHFR sequence only (Edman *et al.* 1989). Table 7.1 summarises the identity, calculated as the percentage ratio of identical matches to consensus length, between *P. carinii* and the other DHFR amino acid sequences. The reported range of identity values (20.7 - 36.1%) although very low, is similar to the identity expected when comparing DHFR sequences of non-related species (Roth & Stammers 1992). The *P. carinii* and *S. cerevisiae* DHFR sequences have a low identity between each other (27.3%). However, recent molecular biological evidence suggests that *P. carinii* is a fungus and that it is closely related to the yeast fungi including *S. cerevisiae* (Lungdren *et al.* 1992). The homology data for DHFR suggest otherwise. Despite the low sequence homology between the two sequences, the considerably higher similarity score (3302) (table 7.1) and comparatively short consensus length (216) (DHFR sequences of *P. carinii* and *S. cerevisiae* are 206 and 211 residues long respectively) suggests there are conserved matches, and hence a greater similarity between the sequences than the identity figures suggest. The generally higher similarity scores between the non-bacterial and *P. carinii* DHFR sequences is due to a greater number of scoring matches, which is a direct consequence of these DHFR sequences being longer.



Aston University

Content has been removed for copyright reasons

Table 7.1 Sequence homology data from aligning known dihydrofolate reductase sequences with the *Pneumocystis carinii* dihydrofolate reductase sequence using the multiple sequence alignment program CLUSTAL (Higgins & Sharp 1988) and the PAM250 matrix (Dayhoff *et al.* 1983)

The CLUSTAL multiple sequence alignment (table 7.2) identified the clusters of homologous residues between the sequence insertions and deletions described by Doolittle (1981). The *E. coli* residues reported to be involved in DHF substrate binding (Bolin *et al.* 1982) were interpolated to determine the equivalent residues in the other sequences in the aligned clusters. A number of these residues in the alignment table are either invariant or highly conserved in the other DHFR sequences. The equivalent residues to glutamate 32 and threonine 144 of *P. carinii* DHFR (in *E. coli* DHFR aspartate 27 and threonine 113) have been shown to form hydrogen bond linkages with the pteridine moiety of DHF. Arginine 75 (arginine 57 in *E. coli*) hydrogen bonds with the glutamate fragment of DHF and serine 64 and leucine 72 of *P. carinii* DHFR (serine 49 and leucine 54 in *E. coli*) interact favourably with the ^9C - ^{10}N linkage and the benzene ring of DHF. Many of the other identical or conserved residues identified by the alignment in table 7.2 are equivalent to those involved in the binding of the co-factor NADPH in the *L. casei* (Matthews *et al.* 1979) and *E. coli* (Adams *et al.* 1991) DHFR crystal structures. Thus, from this multiple sequence alignment many of the key residues in substrate and co-factor binding are identified.

Table 7.3 shows the sequence alignments from the four DHFR structures used in section 7.2.4.3 superimposed onto the human DHFR model (illustrated in figure 7.2). Removal of the non-homologous residues reveals eight regions of structural conservation, highlighted in bold type in table 7.3 and visualised in figure 7.3. Failure by CLUSTAL to identify certain structurally significant residues is a result of limitations of the program. CLUSTAL will not identify a conserved match if, in one sequence only the equivalent residue is considerably different from the others, or, if in any one sequence there is no equivalent residue. However, comparison of the CLUSTAL sequence alignments with the Quanta structural alignments reveals many of the key residues. Poor alignment of the bacterial sequences with the *P. carinii* sequence fails to identify residues 10, 11 and 12 as key residues. Within the fifth and seventh structurally conserved regions there are areas identified as having poor sequence homology using the CLUSTAL parameters. However, the fifth region does not contain any key residues, and in the seventh, residue 153 is identified, whereas 154 and 155 are not.

Table 7.3 also highlights the considerable amount of secondary structure conservation

between the DHFR structures and the residues that are used in the human and *P. carinii* DHFR active site models (*).

In conclusion, the multiple alignment of the DHFR sequences was able to identify conserved residues that are involved with substrate binding. Furthermore, these residues have been shown to be situated in regions common to all determined DHFR structures. Hence, the residues that should be studied further before they can be used to construct rationally the active site model of *P. carinii* DHFR have been identified.

	10	20	30	40	50
<i>P. carinii</i>	MN--QQKSLTLIVALTTTSYGIGRSNSLPW-KLKK E ISYFKRVTSFVPTFDSFESMNVVLM				
<i>S. cerevisiae</i>	MAGGKIPIVGIVACLPQEMGIGFRGGLPW-RLP S EMKYFRQVTSL---TKDPNKKNALIM				
HUMAN	----VGSLNCIVAVSQNMGIGKNGDLPWPPLR N EFRYFQRMST---TSSVEGKQNLVIM				
BOVINE	----VRPLNCIVAVSQNMGIGKNGDLPWPPLR N EFRYFQRMST---VSSVEGKQNLVIM				
PORCINE	----VRPLNCIVAVSQNMGIGKNGDLPWPPLR N EYKYFQRMST---TSSVEGKQNLVIM				
MURINE	----VRPLNCIVAVSQNMGIGKNGDRPWPPLR N EYKYFQRMST---TSSVEGKQNLVIM				
HAMSTER	----VRPLNCIVAVSQNHGIGKNGDFPWPMLR N EYKYFQRMST---TSSVEGKQNLVIM				
CHICKEN	----VRSLNSIVAVCQNMGIGKNGDLPWPPLR N EYKYFQRMST---TSHVEGKQNAVIM				
MOSQUITO	M----KKFSLIVAVCANGGIGIKGDLPW-RLRQ E LKYFSRMTK---KIQDSGKRNAIIM				
<i>H. saimiri</i>	M----VQALNCIVAVAQNMGIGKQGNLPWPRLM N DFKHFRQRMST---TSSVPDKQNLVIM				
<i>B. subtilis</i>	M-ISFIFAMD-----ANRL-IGKDNLDLPWH-LP N DLAYFKKITS-----GHSIIM				
<i>E. faecium</i>	MFISMW-AQD-----KNGL-IGKDGLLPWR-LP N DMRFFREHTM-----DKILVM				
<i>H. volcanii</i>	MELVSVAALA-----ENRV-IGRDGELPWPSP I PA D KKQYRSRIA-----DDPVVL				
<i>L. casei</i>	TAF-LW-AQD-----RDGL-IGKDGHLPWH-LP D D L HYFRAQTV-----GKIMVV				
<i>N. gonorrhoea</i>	MLKITIIAAC-----AENLCIGAGNAMPWH-IP E DAFFAFKVYTL-----GKPVIM				
<i>S. typhimurium</i>	MLISLIAALA-----HNNL-IGKDNLI P WH-LP A D L RHFKA V TL-----GKPVVM				
<i>Staph. aureus</i>	MTLSIIVAHD-----KQ R V-IGYQ N QLPWH-LP N D L KHIK Q LTT-----GNTLVM				
<i>E. coli</i>	M-ISLIAALA-----VDRV-IGMENAMPWN-LP A D L AGFKRNTL-----NKPVIM				
<i>C. freundii</i>	M-ISLIAALA-----VDRV-IGMENAMPWD-LP A D L AWFKRNTL-----NKPVVM				
<i>E. aerogenes</i>	M-ISLIAALA-----VDRV-IGMENAMPWN-LP A D L AWFKRNTL-----NKPVVM				
		**	**	.	.
	60	70	80	90	100
<i>P. carinii</i>	GRKTW S IPLQFR P L K GRINVVITRN-ESL-----DLGNGIHSAKSLDHALELLYRT				
<i>S. cerevisiae</i>	GRKTW S IPPKFR P L P NRMN V ISRSFKDD-----FVDHKERSIVQSNLANAIMNLESN				
HUMAN	GKKTW S IPEKNR P L K GRINVLVSRELKEP-----PQGAHFLS-RSLDDALKLTEQP				
BOVINE	GRKTW S IPEKNR P L K DRINIVLSRELKEP-----PKGAHFLA-KSLDDALELIQDP				
PORCINE	GRKTW S IPEKNR P L K DRINIVLSRELKEP-----PQGAHFLA-KSLDDALKLTEQP				
MURINE	GRKTW S IPEKNR P L K DRINIVLSRELKEP-----PRGAHFLA-KSLDDALRLIEQP				
HAMSTER	GRKTW S IPEKNR P L K DRINIVLSRELKEP-----PQGAHFLA-KSLDDALKLIEQP				
CHICKEN	GKKTW S IPEKNR P L K DRINIVLSRELKEA-----PKGAHYLS-KSLDDALALLDSP				
MOSQUITO	GRKTYF G VPESKR P L P ERLNIILTRDPSAN-----AYPSEVMVCTSMQEALKKLDEA				
<i>H. saimiri</i>	GKKTW S IPEKNR P L K GRINVLSKELKEL-----PHRAHFLA-KSLDDALKLTEQP				
<i>B. subtilis</i>	GRKTF S IG---R P L P NRKNIVVTSAPDSE-----FQCTVVSSLKDVLDIC				
<i>E. faecium</i>	GRKTY E G M G---KL S L P YRHIIVLTTQKD-----FKVEKNAEVLHSIDELLAYAK----				
<i>H. volcanii</i>	GRTTF S M---R D D L P G S AQIVMSRSERS-----FSDTAHRAAS-----VEEAVDIA				
<i>L. casei</i>	GRRTY S F P ---KR P L P ERTNVVLTHQED-----YQAQ-GAVVVHDVAAVFAYAK----				
<i>N. gonorrhoea</i>	GRKTW S LP---VK P L P GRRNIVISRQA-----DYCAAGAETVASLEVALALC				
<i>S. typhimurium</i>	GRRTF S IG---R P L P GRRNIVVSRNP-----QWQAEGVEVAPSLDAALALLTD----				
<i>Staph. aureus</i>	ARKTF S IG---R P L P GRRNIVLTNQAS-----FHHEGV D V-----INSLDEIKEIS				
<i>E. coli</i>	GRHTW S IG---R P L P GRKNIILSSQPGTDDRV T W-VKSVDEAIAACGD-----				
<i>C. freundii</i>	GRLTW S IG---R P L P GRKNIVISSKPGSDDRVQW-VKSVDEAIAACGDFSEFHDA-DAQ				
<i>E. aerogenes</i>	GRHTW S IG---R P L P GRKNIVISSKPGTDDRVQW-VKSVDEAIAACGD-----				
	.. *	*		

Bold typeface identifies key residues discussed

* indicates an identical match

. indicates a conserved match

NBRF single letter amino acid codes :-

A, alanine; C, cysteine; D, aspartate, E, glutamate; F, phenylalanine; G, glycine, H, histidine; I, isoleucine; K, lysine; L, Leucine; M, methionine; N, asparagine; P, proline; Q, glutamine; R, arginine; S, serine; T, threonine; V, valine; W, tryptophan; Y, tyrosine.

	110	120	130	140	150	160
<i>P. carinii</i>	YGSE-----SSVQINRIFVIGGAQLYKAAMDHPKLDRIMATIIYKDIHCDVFFP-LK--F					
<i>S. cerevisiae</i>	F-----KELHERIYVIGGGEVYSQIFSITDH--WLITKI-NPLDKNAT-PAMDTFL					
HUMAN	---E-----LANKVDMVWIVGGSSVYKEAMNHPGHLKLFVTRIMQDFESDTFFPEID--L					
BOVINE	---E-----LTNKVDVWVIVGGSSVYKEAMNHPGHVRLFVTRIMQEFESDAFFPEID--F					
PROCINE	---E-----LKDKVDMVWIVGGSSVYKEAMNKPGLHRLFVTRIMKEFESDTFFPEID--L					
MURINE	---E-----LASKVDMVWIVGGSSVYQEAAMNQPGHLRFLVTRIMQEFESDTFFPEID--L					
HAMSTER	---E-----LADKVDVWVIVGGSSVYKEAMNQPGHLRFLVTRIMQEFESDTFFPEID--L					
CHICKEN	---E-----LKSKVDMVWIVGCTAVYKAAMEKPINHRLFVTRILHEFESDTFFPEID--Y					
MOSQUITO	---P-----LVNEIENVWIVGGNAVYKEAMQSDRCHRIYLTETKETFECDAFFPEIT--					
<i>H. saimiri</i>	---E-----LANKVDMVWIVGGSSVYKEAMSYPCDLKLFVTRIMQDFECDTFFPEFD--L					
<i>B. subtilis</i>	SGPEECF-----VIGGAQLYTDLF--PYADRLYMTKIHHEFEGDRHFPE---F					
<i>E. faecium</i>	-----D-IPEDIYVSGGSRIQALL--PETKIIWRTLIDAEFEGDTFIGE---I					
<i>H. volcanii</i>	ASLD-----AETAYVIGGAIIYALFQ--PHLDRMVLSPVPGYEGDTYPE---W					
<i>L. casei</i>	-----QHLDQELVIAGGAQIFTAFK--DDVDTLVLVTRLAGSFEGDTKMIP---L					
<i>N. gonorrhoea</i>	AG-----AEEAVIMCGAQIYQAM--PLATDLRITTEVDLSVEGDAFFPE---I					
<i>S. typhimurium</i>	-----C-----EEAMIIGGGQLYAEAL--PRADRLYLTYIDAQLNGDTHFPD---Y					
<i>Staph. aureus</i>	G-----HVFIFGGQTLYEAMI--DQVDDMYITVIDGKFQCDTFFPP-----					
<i>E. coli</i>	-----VPEIMVIGGGRVYEQFL--PKAQKLYLTHIDAEEVGDTHFPD---Y					
<i>C. freundii</i>	NSHSYCFEILERRABEIMVIGGGRVYEQFL--PKAHKLYLTHIDAEEVGDTHFPD---Y					
<i>E. aerogenes</i>	-----EPEIMVIGGGRVYEQFL--PKAQKLYLTHIDAEEVGDTHFPD---Y					
	
	170	180	190	200		
<i>P. carinii</i>	RDKEWSSVWKKEKHSDESWSVGTQVPHGKI-----NEDGFYEYEFEMWTR-DL-----					
<i>S. cerevisiae</i>	DAKKLEEVFSEQDPAQLKEFLPKVELPETDCDQRYSSLEEKGYCFEFTLYNR--K-----					
HUMAN	-----EKYKLLPEYPG--VLSDVQ-----EEKGIKYKFEVYEK-ND-----					
BOVINE	-----EKYKLLPEYPG--VPLDVQ-----EEKGIKYKFEVYEK-NN-----					
PORCINE	-----EKYKLLSECSG--VPSDVQ-----EEKGIKYKFEVYEK-NN-----					
MURINE	-----GKYKLLPEYPG--VLSEVQ-----EEKGIKYKFEVYEK-KD-----					
HAMSTER	-----EKYKLLPEYPG--VLSEVQ-----EEKGIKYKFEVYEK-KG-----					
CHICKEN	-----KDFKLLTEYPG--VPADIQ-----EEDGIQYKFEVYQK-SV-----					
MOSQUITO	-----SDFQLVKNDDD--VPEDIQ-----EENGIQYQYRIYEKVPK-----					
<i>H. saimiri</i>	-----EKYKLLIEYPS--VLSNVQ-----EEKSIKYKFEVYEK-NH-----					
<i>B. subtilis</i>	DESNWKLVSSEQGQTKD-----EKNPYDYEFLLMYEKKNSSKVGGF					
<i>E. faecium</i>	DFTSFELVEEHEGIVN-----QENQYPHRFQKWQKM--SKVV--					
<i>H. volcanii</i>	DAAEWELDAETDHE-----GFTLQEWVRSASSR---					
<i>L. casei</i>	NWDDFTKVSSRT-VED-----TNPALTHTYEVWQKK--A----					
<i>N. gonorrhoea</i>	DRTHWREAERTERRVS-----SKGV-AYTFVHYLGK-----					
<i>S. typhimurium</i>	LSLGWQELERSTHPAD-----DKNSYACEFVTLRSQR-----					
<i>Staph. aerus</i>	-----TTFEQLD-----EKNTPHTFLHLVRRKGK-----					
<i>E. coli</i>	EPDDWESVFSEFHDAD-----AQNSHSYCFQILER-R-----					
<i>C. freundii</i>	DPDEWESV-----					
<i>E. aerogenes</i>	EPDDWESVFSEFHDAD-----AQNSHSYCFEILER-R-----					

Table 7.2 CLUSTAL (Higgins & Sharp 1988) multiple sequence alignment of twenty DHFR sequences: *Pneumocystis carinii* (Kovacs *et al.* 1989); *Saccharomyces cerevisiae* (Lagosky *et al.* 1987); human liver (Masters *et al.* 1983); bovine liver (Lai *et al.* 1982); porcine liver (Smith *et al.* 1979); murine liver (Simonsen & Levinson 1983); hamster liver (Mellera *et al.* 1984); chicken liver (Kumar *et al.* 1980); day forest mosquito (Shotkoski & Fallon 1991); *Herpesvirus saimiri* (Trimble *et al.* 1988); *Bacillus subtilis* (Iwakura *et al.* 1988); *Enterococcus faecium* (Peterson *et al.* 1975); *Halobacterium volcanii* (Zusman *et al.* 1989); *Lactobacillus casei* (Batley & Morris 1977); *Salmonella typhimurium* (Joyner *et al.* 1984); *Neisseria gonorrhoea* (Baccanari *et al.* 1984); *Staphylococcus aureus* (Rouch *et al.* 1989); *Escherichia coli* (Bolin *et al.* 1982); *Citrobacter freundii* (Garvin & Hardies 1991); *Enterobacter aerogenes* (Garvin & Hardies 1991)

	Ba	cb	βb	αc
	***	***	*****	*****
<i>P. carinii</i>	MNÖKS LTLIALTTTSYGIGRSNLSLPM K- LKKEISYFKVT SEVPTDSFESM NVYLAGKRTWESI PLÖF			
Human Liver	---VGS LNCIYAVBSÖNMGIGGNEDLPM PP LKNEFKYFÖGGE TT--SSVEGK--Ö NLVINGKRTWFSI PEKN			
Mouse Liver	---VGP LNCIYAVBSÖNMGIGGNEDLPM PP LKNEFKYFÖGGE TT--SSVEGK--Ö NLVINGKRTWFSI PEKN			
Chicken Liver	---VRS LMSIYAVCÖNMGIGGNEDLPM PP LKNEFKYFÖGGE ST--SHVEGK--Ö NVALMGKRTWFSI PEKN			
<i>E. coli</i>	-----M ISLIAALAVDRIVGGENAMPM N- LPADLIAMFKRMT -----LD----- KPVINGKRTWESI --G			
<i>L. casei</i>	-----TAPLMAÖDRDGLIGDGLPM H- LPDDLIHTFRAÖT V-----G KIMVVGRRRTYESF --GK			
	βc	βd	αe	βe
	**	**		*****
<i>P. carinii</i>	RPLKGRINVTIT RNESLDLNGI HSAKSLDHALLELYR TYGSESSVQINRI FVIGGAÖLYKAM DHPKIDR			
Human Liver	RPLKGRINVTIS RELKEPPÖGAH FLRSLLDDALKTETÖ PELANK---VDMV NIVGSSSVYKEM NHPGHLK			
Mouse Liver	RPLKGRINVTIS RELKEPPÖGAH FLRSLLDDALKTETÖ PELASK---VDMV NIVGSSSVYÖEAM NÖPGHLK			
Chicken liver	RPLKGRINVTIS RELKEAPKGAH YLSKSLDDALALIDS PELISK---VDMV NIVGGLAVYKAM EKPINHR			
<i>E. coli</i>	RPLPGRKNILIS --SQPGTDDRV TWKSVDEALIAACGD ---VPEI----- NVIGGGRVTEÖFL PKA--ÖK			
<i>L. casei</i>	RPLPERTNVILT HÖEDYÖ-AÖGA VVVDVAAVFAAYAKÖ H---PDÖ---EL VIAGGAÖITFAFK DDV--DT			
	βf	βg		βh
	*	***		*
<i>P. carinii</i>	IMATIIYKDICHCVFFFLKFRDKEM SSVWKKEKHSDLSEWGTQVPHGKINEDGF EYEFEMNTRD L			
Human Liver	LFVTRIMÖDFESDTFFPEIDLEKXK LL-----PEYPGVLSDVÖEKG-----I KYKFEVYÖKS V			
Mouse Liver	LFVTRIMÖDFESDTFFPEIDLEKXK LL-----PEYPGVLSDVÖEKG-----I KYKFEVYÖKS D			
Chicken Liver	LFVTRILHEFESDTFFPEIDYKFDK LL-----TEYPGVPADIQEEDG-----I ÖYKFEVYÖKN D			
<i>E. coli</i>	LYLTHIDAEVEGDTHFPDYEPPDME SVFSEFHNAQAÖNSH-----SYCFKILERR -			
<i>L. casei</i>	LLVTRLAGSFEGDTKMTPLANDPT KVSSTRTVEDTNPALT-----NHTTEVWÖKK A			

Table 7.3 *P. carinii* DHFR sequence aligned with determined DHFR structures.
Code :- bold type structural homology; red, α-helix; blue, β-sheet, green, β-turn; *, residues used in the *P. carinii* DHFR active site model.



Figure 7.2 Four dihydrofolate reductase structures, as described in section 7.2.4.3: red, mouse liver; blue, chicken liver; yellow, *E. coli*; green, *L. casei*; superimposed onto the model human dihydrofolate reductase (white)



Figure 7.3 Superimposed structurally conserved regions, as highlighted in table 7.3, for the five dihydrofolate reductase structures: white, human liver; red, mouse liver; blue, chicken liver; yellow, *E. coli*; green, *L. casei*:

There are four sets of crystallographic coordinates available for human DHFR. However construction of a model of human DHFR from these determined structures instead of using only one of the structures was preferred. This is because the reported resolution of these four DHFR structures (2.3 Å) could be translated to a 0.4 Å error on the atomic coordinates. The model structure consisting of the averaged coordinates reduces the impact of any anomalies in any one of the DHFR structures. Furthermore, as discussed in chapter 4, enzyme kinetic evidence indicates that AM90067 interacts with both the substrate and cofactor sites in DHFR. It has been shown that the binding of NADPH can alter the conformation of the peptide backbone in certain regions (Champness *et al.* 1986). Because the human DHFR crystal structures have been solved without the co-factor present, the peptide backbone is already in the conformation it adopts in the absence of NADPH. This is more applicable for the molecular modeling studies investigating the interaction of AM90067 docked in the active site of human and *P. carinii* DHFR.

A cautious approach was taken in the construction of the *P. carinii* DHFR active site model. Only the model human DHFR coordinates were used to construct the peptide backbone of the *P. carinii* DHFR model. As far as possible the averaged atomic coordinates from identical residue matches from known DHFR structures were used to predict the side chain conformation of *P. carinii* DHFR. Where only one DHFR structure provided an identical match to the equivalent *P. carinii* DHFR residue, then only these side chain coordinates were introduced into the model structure. In the cases where the residues were unique to *P. carinii* DHFR the procedure for model building side chains in equivalent residues described by Sutcliffe *et al.* (1987) was employed. This involved determining the spatial location of the side chain atoms in the known DHFR structures and applying these to the *P. carinii* model as far as the correlation is observed. There were exceptions to this. In the DHFR crystal structures there was no consistency in the torsion angle N(ser24)-C α (ser24)-C β (ser24)-C γ (ser24) of serine 24. Thus, using the suggested methods it was not possible to arrive at a rationally predicted conformation for serine 24. It has been shown that the side chains of α -helices and β -sheets in the hydrophobic core of a protein are restricted in the conformations they can adopt (Piela *et al.* 1987, Levitt 1992). Ponder and Richards (1986) have shown that the serine side-chain is limited to three possible torsion angle ranges in this region. Table

7.3 shows that this residue in the DHFR structures is located in a β -sheet, hence forming part of the hydrophobic core. It is the peptide backbone and not the side chain that is involved with NADPH binding. Therefore, with respect to AM90067 binding in the NADPH active site, the side chain conformation of this residue is of little consequence. Thus, once a low energy side chain conformation was determined it was retained. Only the conformations from the mouse DHFR were used to predict the side chain conformations of residues arginine 54 and lysine 55. In these instances, from superimposing the equivalent residues from the DHFR structures it was obvious that these residues did not adopt similar side chain conformations. Hence, there was no logical basis in using averaged side chain coordinates. Because the mouse DHFR had an identical sequence to *P. carinii* in this region (residues 57 to 62) it was decided that only these coordinates should be used. This would provide a better basis for subsequent exploitation in rational drug design. For valine 154, because of the obvious structural similarities between valine and threonine, it was assumed that C γ 1 of valine in *P. carinii* DHFR would adopt the same position as O γ 1 of threonine in the known DHFR structures. Therefore O γ 1 in the model structure was directly retyped to a carbon atom. This very conservative approach to construction of the *P. carinii* DHFR active site model introduced changes from other DHFR structures only when compelled to by the sequence differences. It was reasoned that any apparent structural influence on the greater differential activity of AM90067 towards *P. carinii* DHFR would be genuine, and might be amplified in experimental analysis of any potential inhibitor substituent modification derived from the docking studies. The disadvantage of this approach is the possibility that true interactive effects may be masked.

The human DHFR structures have either DHF or the folate analogue, 5-deazafolate bound in the active site. Previous studies have shown that the pteridine ring in methotrexate is inverted relative to DHF when bound to the human DHFR (Oefner *et al.* 1988). This causes significant differences between the hydrogen bonding of DHF and methotrexate with the active site residues (Fontecilla-Camps *et al.* 1979, Charlton *et al.* 1979). Crystallographic studies have shown that the 2,4-diaminopyrimidines bind in the same way as methotrexate (Baker *et al.* 1983). Thus, using the model of the *meta*-azidopyrimethamine bound in the active site of *E. coli* DHFR (Sansom *et al.* 1989) suitable coordinates were obtained to dock AM90067 that were in agreement

with the expected pattern of hydrogen bonding for this class of DHFR inhibitors (figure 4.3b).

There is no information available in the deposited human DHFR structures concerning the position of solvent molecules. Two water molecules were included that are known to be highly conserved in the bacterial (*E. coli* and *L. casei*), chicken liver and mouse liver enzymes. These are known to be important in the binding of 2,4-diaminopyrimidines (Matthews *et al.* 1985, Stammers *et al.* 1987). It was judged that this simplistic approach, plus the other approximations, would suffice for the comparison of conformations of the same compound in the two models.

Removing the residues greater than 5 Å from each conformation of AM90067 docked in the active sites of the model human and *P. carinii* DHFR reduced the number of residues involved in the interaction energy calculation. Because hydrogen bond contacts involve bond lengths of between 2.5 and 3.5 Å (Taylor & Kennard 1984) and at a 5 Å interatomic separation the van der Waals interaction forces are minimal (Gilbert & Hyde 1992) it is unlikely that residues greater than 5 Å from AM90067 when docked in DHFR will have a significant effect on its binding energy.

To minimise the number of conformational variables the side chains were assumed to be stationary points. As the main interest of this study is to evaluate the effect that the 3''-hydroxyethyl-3''-benzyltriazen-1''-yl substituent has on the binding of AM90067 in the human and *P. carinii* DHFR models, the torsion C5D-C6D-C61D-C62D (figure 5.1) has little significance. Therefore, throughout the series of conformations docked in both active site models this torsion was maintained at 90°. Results from chapter 2 show the C4D-C5D-C1P-C2P (figure 5.1) torsion angle in single crystal structures of pyrimethamine and its derivatives is not exactly perpendicular, but varies from 60 to 120°. However, it has been shown that when pyrimethamine is bound to *E. coli* DHFR this torsion is 90° (Baker *et al.* 1983). By continuing the knowledge-based approach to the docking of AM90067, because of the 3'-substituent, this torsion angle should be fixed at 90° or 270°. Although this appears a simplistic assumption, the results from chapter 5 and the preliminary docking of pyrimethamine support it. The calculations for the docking of pyrimethamine agree with the protein crystallographic data suggesting two minima; one at 90°, the other at 270°. Presumably the level of sophistication of the CHARMM force field used to determine the energy of interaction is not sufficient to

predict non-perpendicularity in this torsion.

The interaction of AM90020 in the active site of human and *P. carinii* DHFR with the torsion C4D-C5D-C1P-C2P set to 270° and the 3'-(3'',3''-dimethyltriazene-1''-yl) moiety in its crystallographic configuration revealed that the two methyl carbons, C7 and C8, produce steric contacts with residues threonine 61 and leucine 72 (figure 7.4) respectively in human and *P. carinii* DHFR models. Rotation of the C2P-C3P-N1-N2 (figure 5.1) torsion was necessary to relieve these contacts. However this procedure orientates N2 close to the 4'-chloro substituent. Conformational studies from the previous chapter show a considerable increase in the molecular energy when these two atoms locate in close proximity to each other as the C2P-C3P-N1-N2 torsion rotates. Thus, energetically, it would be unfavourable for the 3'-triazenyl analogues of pyrimethamine to align in the active site with the C4-C5D-C1P-C2P torsion at 270°. By contrast there were no steric contacts observed when the C4D-C5D-C1P-C2P was set at 90°. In this conformation the 3'-triazenyl substituent locates in the nicotinamide region of the NADPH binding site (figure 7.5), which coincides with the results from chapter 4 showing AM90067 to be a competitive inhibitor of NADPH at low inhibitory concentrations. Therefore, in the docking of the conformations of AM90067 the torsion angle C4D-C5D-C1P-C2P was restricted to 90°.

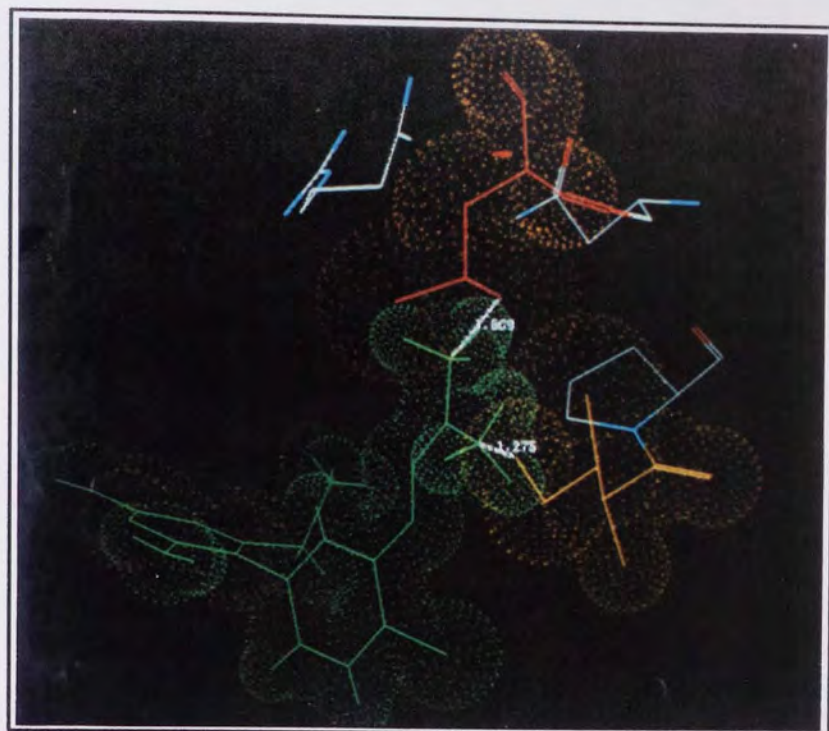


Figure 7.4 AM90020 (green) in the active site model of *P. carinii* DHFR with the torsion C4D-C5D-C1P-C2P at 270° and the 3'-(3'',3'')-dimethyltriazen-1''-yl) substituent in its crystallographic geometry. Close contact distances with residues 61 (yellow) and 72 (red) are highlighted. For clarity, most residues have been omitted.

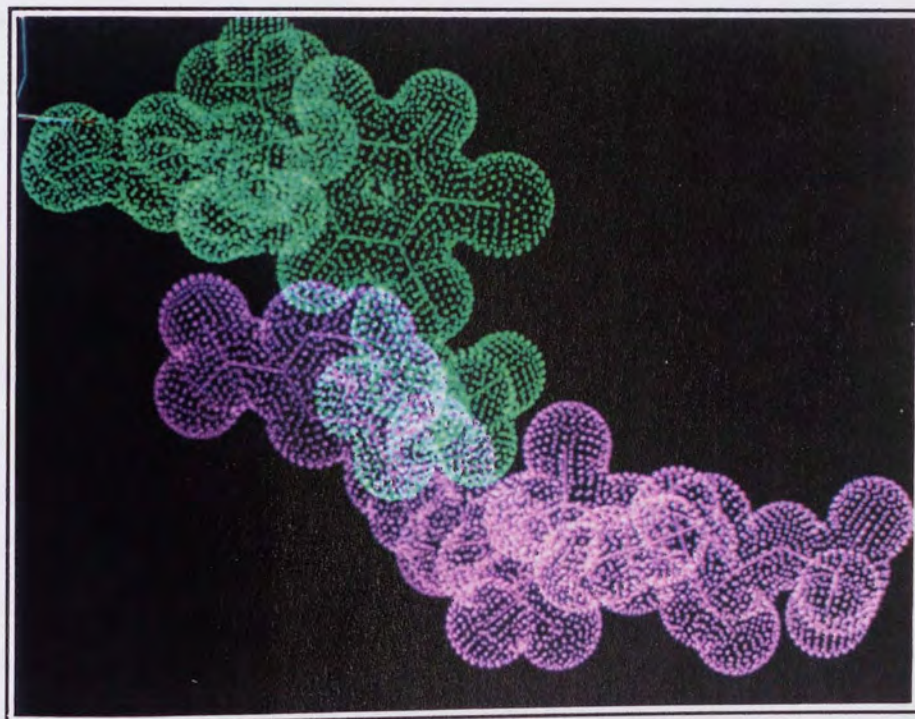


Figure 7.5 The 3'-(3'',3'')-dimethyltriazen-1''-yl) substituent of AM90020 (green) occupying the nicotinamide region of the NADPH binding site are highlighted. Homology modelling was used to introduce the NADPH (pink) structure from the mouse liver DHFR

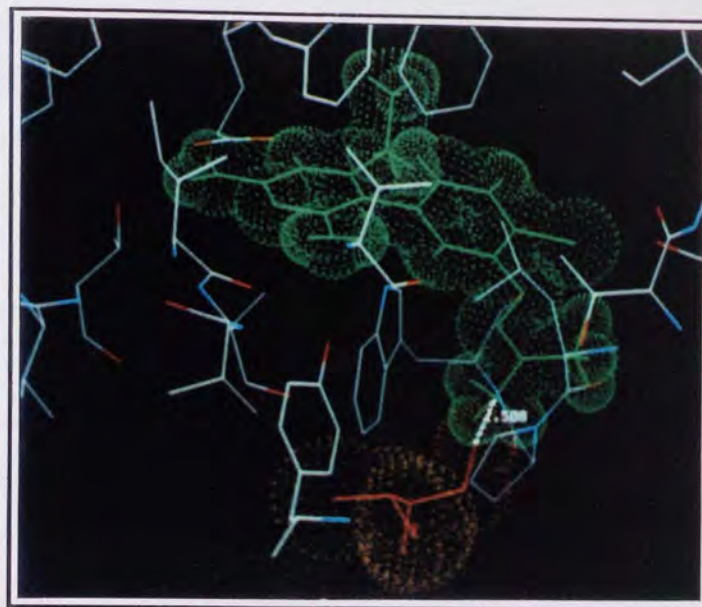


Figure 7.6 AM90020 (green) docked in the active site model of *P. carinii* DHFR with the C4D-C5D-C1P-C2P torsion set to 90° . Steric contact with residue 16 (red) when the C2P-C3P-N1-N2 torsion was 20° .

Conformational studies on the C2P-C3P-N1-N2 torsion in AM90020 and AM90067 in the previous chapter predict a minimum energy at approximately 315° . The final AM1 optimisations of AM90067, in its low energy conformations, consistently predicted a C2P-C3P-N1-N2 torsion in the range of 300° to 310° . This differs from the range of 335° to 20° for phenyltriazene structures determined experimentally by X-ray diffraction. It would appear that limitations in the AM1 parameter set cause the overprediction of the magnitude of this torsion. From the docking of AM90020, it was observed that when the C2P-C3P-N1-N2 torsion was 20° , C8 was 1.508 \AA from the side chain of isoleucine 18 (figure 7.6). In its crystallographic conformation this distance was 2.45 \AA . Therefore, for the docking of AM90067 it was assumed that the C2P-C3P-N1-N2 torsion would adopt a similar geometry to the AM90020 crystallographic conformation when docked in the active site DHFR models.

The interaction energy monitored is based on the electrostatic and van der Waals forces between the enzyme and AM90067. Because of the approximations involved in the active site model construction and the docking of AM90067, the energy of interaction values quoted in tables 7.4 and 7.5 should not be regarded as absolute. Furthermore,

when bound in the active site, the flexible nature of AM90067 is not taken into consideration, but it is regarded as a rigid structure. Therefore, at best, these values produce qualitative and not quantitative data that could possibly explain the selective activity of this compound toward *P. carinii* DHFR. Comparison of the values of the calculated ΔH_i was unable to detect quantifiably any difference between the binding of AM90067 in the active site models of *P. carinii* and human DHFR. When the torsions N1-N2-N3-C7 and N2-N3-C8-C1B were 180° and 300° respectively, the resulting large ΔH_i values are a consequence of the benzyl group occupying the same space as serine 64. The considerable reduction of ΔH_i in conformation 24 in both models is caused by the hydroxyl moiety of AM90067 being rotated away from the carbonyl oxygen of residue 20. When the torsions N1-N2-N3-C7 and N2-N3-C8-C1B were 0° and 159° low ΔH_i values are observed. In this conformation the benzyl moiety of AM90067 was orientated to within 3.0 \AA of one of the γ -carbons in valine 154 in the *P. carinii* model (figure 7.7). In the human DHFR model the benzyl moiety orientated to 3.0 \AA of the hydroxyl oxygen of the equivalent threonine residue. Although there is no experimental data to confirm these as potential binding conformations, they do provide an explanation for the greater inhibitory activity for *P. carinii* DHFR exhibited by AM90067. Furthering this, for drug design purposes, analogues of AM90067 with a benzyl substituent retained on one of the terminal triazene nitrogens provide a rational basis for the synthesis of the next generation of inhibitors of *P. carinii* DHFR.



Figure 7.7 AM90067 (green) docked in the active site of *P. carinii* DHFR. Contact between the 3''-benzyl moiety and residue 154 (red) is highlighted. Residues have been omitted for clarity.

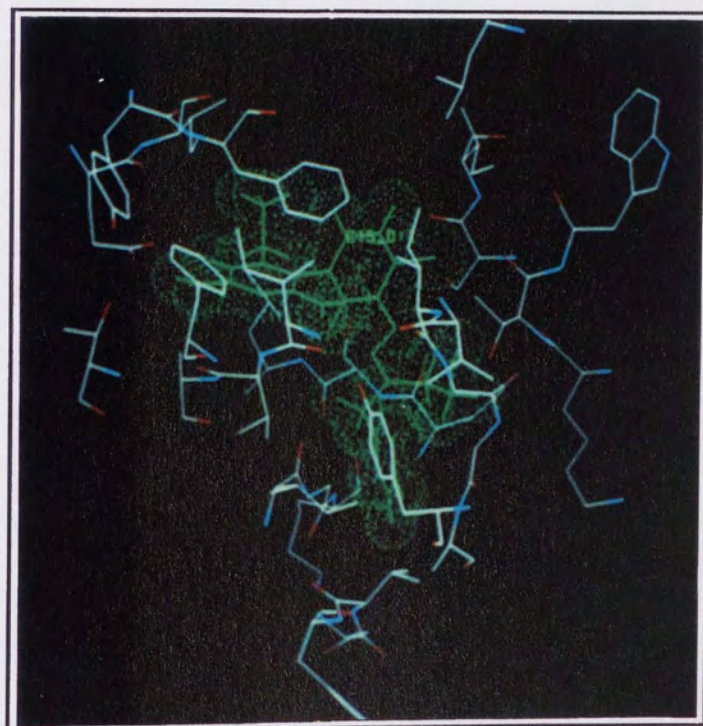
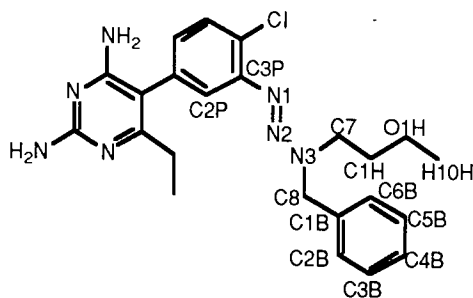


Figure 7.8 AM90067 (green) docked in the active site model of *P. carinii* DHFR.

Docking Position	ΔH_i (rigid fit before rotation) (kcal mol ⁻¹)	C7-C1H-O1H-H10H		N2-N3-C8-C1B	
		torsion (°)	ΔH_i (after rotation) (kcal mol ⁻¹)	torsion (°)	ΔH_i (after rotation) (kcal mol ⁻¹)
1	1430	298	558	40	507
2	1360	270	31.0	279	24.0
3	2500	298	1840	70	393
4	1920	264	1880	89	5.0
5	1710	301	585	298	585
6	301	266	24.3	294	21
7	1760	301	611	294	605
8	61.2	266	18.7	298	18.0
9	1560	301	538	62	532
10	236	275	45.2	61	38.2
11	1900	299	742	123	616
12	485	270	207	112	13.8
13	44x10 ⁷	9	5.2x10 ⁷	-	5.2x10 ⁷
14	17.0	269	-0.8	118	-11.9
15	70x10 ⁷	7	7.9x10 ⁷	-	7.9x10 ⁷
16	65.9	271	46.1	176	-8.74
17	3.9x10 ⁷	106	3.2x10 ⁷	-	3.2x10 ⁷
18	318	272	291	205	34.6
19	73x10 ⁷	7	8.2x10 ⁷	-	8.2x10 ⁷
20	16.8	121	16.8	38	5.4
21	86x10 ⁷	7	9.7x10 ⁷	-	9.7x10 ⁷
22	33.2	121	33.0	54	8.6
23	62x10 ⁷	7	7.0x10 ⁷	-	7.0x10 ⁷
24	0.29x10 ⁶	269	0.28x10 ⁶	58	20.0



Docking Position	ΔH_i (rigid fit before rotation) (kcal mol ⁻¹)	C7-C1H-O1H-H10H torsion (°)	ΔH_i (after rotation) (kcal mol ⁻¹)	N2-N3-C8-C1B torsion (°)	ΔH_i (after rotation) (kcal mol ⁻¹)
1	1330	298	605	279	549
2	90.0	268	23.0	274	14.9
3	5630	298	4680	90	550
4	5650	268	5580	88	19.6
5	1560	299	577	130	575
6	248	268	31.5	255	28.3
7	2490	301	848	298	844
8	89.3	269	22.4	291	21.9
9	2140	300	789	65	783
10	87.3	271	29.8	71	23.0
11	2600	301	917	90	838
12	369	270	143	91	17.4
13	174x10 ⁷	350	11.5x10 ⁷	-	11.5x10 ⁷
14	186	269	-0.6	118	-16.8
15	273x10 ⁷	350	18.6x10 ⁷	-	18.6x10 ⁷
16	28	269	73.1	0	-5.2
17	11.8x10 ⁷	299	3.76x10 ⁷	-	3.76x10 ⁷
18	838	268	662	200	41.3
19	282x10 ⁷	348	19.4x10 ⁷	60	13.9x10 ⁷
20	7.0	284	6.3	33	4.5
21	328x10 ⁷	347	23.0x10 ⁷	-	23.0x10 ⁷
22	27.8	119	25.5	52	-3.3
23	241x10 ⁷	350	16.3x10 ⁷	-	16.3x10 ⁷
24	0.1x10 ⁷	270	0.1x10 ⁷	64	23.5

Table 7.5 Energies of interaction (ΔH_i) for the original low energy conformations of AM90067 bound in the human dihydrofolate reductase active site model and after the C7-C1H-O1H-H10H and then the N3-C8-C1B-C2B torsions have been rotated to the minimum active site interaction energy conformations.

Chapter 8 Treatment of Experimental *Pneumocystis carinii* Pneumonia using AM90067

8.1 Introduction

Access to a reliable model of PCP that can be used as a source of large numbers of organisms for laboratory studies and as a tool for evaluating the efficacy of new drugs is paramount. *In vitro* drug testing has frequently used feeder cell layers and relied on demonstrating the differences in numbers of organisms in untreated and drug-treated culture cells. However, *P. carinii* cell culture has met with only limited success, caused by the lack of reproducibility between laboratories and the inability to cultivate the organism long term. Thus animal studies have been employed for drug assessments. Historically, most animal studies used to study *P. carinii* relied on steroid-induced immunosuppression to activate latent *P. carinii* infections (Frenkel *et al.* 1966). The presence of the latent infection in many breeding rat colonies resulted in this rodent becoming the animal of choice for many investigators (Boylan & Current 1992). The utility of this animal model is limited because of the long time required for the development of heavy infections (8 to 12 weeks) and the acquisition of secondary, in particular fungal, infections.

Several processes of evaluation of antipneumocystis activity have been developed. The most common method takes a histological section of the infected lung and after microscopic examination, a score between 0 (no infection) and 4 (heavy infection) is assigned (Hughes *et al.* 1974). This technique requires a trained eye and is very subjective (Durkin *et al.* 1992). However, it does give the researcher a better opportunity to assess the status of the infection and host tissue changes in the lungs (Kim *et al.* 1987). The other techniques produce a count of the total number of cysts in the lung. This does not give any indication as to the extent of infection throughout the tissue. Included in these techniques are enzyme-linked immunosorbent assays (ELISA) and manual counting of lung tissue homogenates. Despite many published ELISA methods to quantify the number of *P. carinii* cysts in lung tissue (Lane *et al.* 1991, Durkin *et al.* 1992) the reliability and reproducibility of these methods are uncertain. Manual counting involves separating the cysts and trophozoites from the lung tissue. Primarily this is achieved by homogenisation, filtration of the lung tissue prior to

staining, and quantifying the number of cysts microscopically. Thus, although it is time consuming, this technique accurately and reliably expedites the evaluation of the drug treatments with minimal subjectivity and is preferred for drug screening programs (Kim *et al.* 1987).

8.2 Materials and Methods

8.2.1 Sources of Chemicals

The following reagents were obtained or purchased from the sources indicated

Aston Molecules Ltd., Gosta Green, West Midlands.

AM90067

BDH Chemicals Ltd., Atherstone, Warwick.

Formaldehyde Solution (38 - 41%)

Sodium chloride

Sodium dihydrogen orthophosphate

Disodium hydrogen orthophosphate

Fisons, Loughborough, Leicester.

Ammonium chloride

Hopkins and Williams, Chadwell Heath, Essex.

Leishmann's Stain

Northumbria Diagnostics Ltd., Northumbria

Pneumocystis carinii direct immunofluorescence diagnostic kit

Sigma Chemical Company, Poole, Dorset.

Dexamethasone-21-phosphate sodium

Histo-Paque 1052

Pentamidine isethionate

Tetracycline hydrochloride

8.2.2 Animal Protocol

For each drug tested fifty male Wistar rats (200 - 250g) (Aston University, Department of Pharmaceutical Sciences in-house colony) were immunosuppressed with the addition of dexamethasone (2 mg l⁻¹) to the drinking water for a period up to six weeks. As a

cover to bacterial infections tetracycline (500 mg l⁻¹) was also added to the drinking water. All animals were maintained on a low protein (8%) diet. Either after six weeks or at the onset of symptoms of PCP (diagnosed as weight loss, with rapid breathing or blooded noses) the rats were divided into dosing groups of up to 10 rats. Each group received daily, for the next ten days, intramuscular (i.m.) injections of pentamidine isethionate in 0.9% saline or AM90067 in a 2:1 mixture of DMSO and 0.9% saline. Free solutions of 0.9% saline and the aforementioned mixture were administered to a group each as controls. Treatment was discontinued if the symptomatic weight loss became greater than 10% of the original body weight. On termination of treatment the animals were anaesthetised with halothane before being sacrificed and the chest cavity was opened to expose the heart and lungs. A 20-gauge needle was inserted into the right atrium and gently exsanguinated until no further blood could be removed. The lungs were removed and placed in phosphate-buffered saline (PBS) (NaH₂PO₄, 2.5g; Na₂HPO₄, 2.523g; NaCl, 8.2g; H₂O to 1000 ml) and stored at -20°C prior to the evaluation of the cysts.

8.2.3 Quantitative Determination of *P. carinii* Cysts in Rat Lung

The lungs were sectioned into small pieces and then, with 10 ml of PBS, placed in a stomacher bag (Seward Medical) and homogenised for 15 minutes. The contents were then filtered through a 63mm stainless steel sieve. The filtrate was placed in a centrifuge tube with 3 ml of Histo-paque 1052 and then centrifuged at 3000g for 10 minutes in a 8 x 50 ml angle rotor (Beckman J2-21). The layers at the interface were collected, resuspended in PBS and incubated at 37°C for 5 minutes in 0.85% ammonium chloride to lyse red blood cells. The suspension was centrifuged at 3000g for 10 minutes, washed twice with 2ml of PBS and then made up to 2 ml with PBS and resuspended. 5µl aliquots from 1 in 3, 1 in 5, 1 in 10 and 1 in 20 dilutions were placed on a 6mm diameter teflon coated multispot slide (Henley-Essex), allowed to dry and then stained with Leishmann's Stain (Mackie & McCartney 1953). Cysts were counted from at least forty fields at x1000 magnification using a Zeiss Axioskop microscope with bright field illumination.

8.2.4 Direct Immunofluorescence for *P. carinii*

A non-stained slide was fixed in cold acetone for 10 minutes and then allowed to evaporate to dryness. Reconstituted enzyme preparation (20 µl and 5 µl), provided with the Northampton Diagnostics *Pneumocystis carinii* direct immunofluorescence kit, was overlaid onto the test and control slides respectively and incubated in a humidified chamber at 37 °C for 30 minutes. The enzyme preparation was employed to improve access of the antibody to the cysts. The slides were then washed in distilled water for 10 minutes. The excess water was removed with blotting paper and then allowed to air dry. Antipneumocystis fluorescein isothiocyanate (FITC)-labelled antibody preparation (15 µl and 5 µl) was applied to the test and control slides, incubated for a further 10 minutes, then washed and dried. The FITC-labelled goat antipneumocystis antibody was applied to the test and control slides, incubated in a humidified chamber at 37°C, washed and dried. A drop of mountant (PBS, pH8.5 and 20% glycerol) was placed onto each test and control well, the appropriate coverslip applied before any excess mountant was removed. Both the control and the test slides were examined for bright apple green cysts using an epifluorescence objective with 520 nm filter and UV illuminant.

8.3 Results and Discussion

Pentamidine was chosen as a positive control in this experiment because it is used clinically as a first line therapy for PCP in patients with AIDS. Therefore it could provide an ideal comparison for determining the *in vivo* activity of AM90067 for this condition. It was necessary to select an effective route of administration suitable for both pentamidine and AM90067. Oral and subcutaneous administration were discounted because pentamidine has poor gastrointestinal tract absorption (Tidwell *et al.* 1991) and causes tissue necrosis at the injection site (Jones *et al.* 1992). Intramuscular (i.m.) delivery was preferred to intravenous (i.v.) because both pentamidine (Cheung *et al.* 1993) and pyrimethamine (Worth & Werbel 1984) have been shown to be efficacious using this route. Therefore, assuming the free bases of pyrimethamine and AM90067 have similar pharmacokinetic profiles then an i.m. route of administration of AM90067 would appear to be suitable. Furthermore the expense and expertise of anaesthetising the animals prior to i.v. administration was avoided.

Daily i.m. doses between 1 and 10 mgkg⁻¹ (body weight) of pentamidine (Jones *et al.* 1990) have been shown to be efficacious for PCP treatment in rats. Similarly, 3 - 9 mgkg⁻¹ doses of pyrimethamine are efficacious (Walzer *et al.* 1992) in rats when used with a sulphonamide. The lower doses were chosen to investigate the possibility of AM90067 possessing antipneumocystis activity greater than that of pentamidine or pyrimethamine without co-administration of another antimicrobial agent. Although the human clinical dose regimen for PCP chemotherapy lasts for 14 - 21 days, Hughes *et al.* (1990) showed that 10 days' treatment is sufficient to demonstrate antipneumocystis activity in rats.

The following evidence confirms the presence of *P. carinii* cysts. Firstly figure 8.1 shows the presence of polynuclear cysts, approximately 6 µm in diameter stained with Leishmann's stain. *P. carinii* cysts have been reported to be between 5 and 8 µm in diameter. Secondly the direct immunofluorescence technique confirmed the presence of *P. carinii* cysts (figure 8.2). However, not all the polynuclear cysts revealed by the Leishmann's stain were detected by the diagnostic kit. This may be due to the kit being developed to detect specifically *P. carinii* from human lung tissue and the current literature reports poor antigenic cross-reactivity between *P. carinii* from different animal sources (Gigliotti *et al.* 1993). Finally, there is a dose-response relationship between increasing pentamidine doses and the cyst count (figure 8.1) over an expected range. Thus by considering these separate pieces of evidence together it may be concluded with reasonable confidence that the purple stained cellular bodies counted are *P. carinii* cysts.

Examination of the lungs of rats withdrawn treatment early because of excessive weight loss enables an assessment of the severity of the infection to be made. Table 8.1 lists the doses of pentamidine and AM90067, the duration of treatment and the number of cysts per lung for the animals that were withdrawn from treatment early. The range of cyst counts reported, especially for pentamidine at 6.4 mgkg⁻¹, suggests that a severe form of PCP was not responsible for sacrificing the animals before the treatment finished. Also the animals were not observed to be suffering from any toxic reactions. Because so few rats were withdrawn from the treatment (7/95) it is difficult to make any definitive conclusions, though one explanation may be attributed to effects of continued

corticosteroid immunosuppression.

Drug	Concentration (mg kg ⁻¹)	Cyst Count (x10 ⁶)
control	-	117.20
Pentamidine	6.4	1.30
Pentamidine	6.4	70.98
Pentamidine	6.4	16.61
AM90067	0.4	96.30
AM90067	0.4	74.70
AM90067	6.4	56.95

Table 8.1 Dosing groups and cyst counts of the animals prematurely withdrawn from treatment

Pages 162

Missing page(s) from the bound copy

The maximum number of cysts detected in any one lung was in the order of 10^8 . This is less than the 10^9 achieved by other researchers and may be due to the experiment being subject to Home Office regulations under which no more than 10% weight loss was permitted before sacrifice. The lowest cyst burden detected was in the range of 10^5 , a sensitivity that agrees with other studies (Lane *et al.* 1991, Kim *et al.* 1987).

Figures 8.3 and 8.4 show the cyst counts for all the rats in the control and in each test dosing group for pentamidine and AM90067 respectively. For statistical analysis the cyst counts for the test dosing groups were expressed as a percentage of the mean cyst count of the control group. Applying the one way analysis of variance statistical method to the modified pentamidine data, variation in the entire data is observed ($P = 0.000$). Using the multiple range Turkey-B test, the cyst count for pentamidine at 0.1 mgkg^{-1} is shown to be significantly different from the cyst count for the 0.4 mgkg^{-1} group, which itself is significantly different from the 1.6 and 6.4 mgkg^{-1} groups ($P = 0.05$). Similarly, for AM90067 one way analysis of variance shows significant variation at $P = 0.0011$. Using the LSD multiple range test the data from the 6.4 mgkg^{-1} group of AM90067 differs significantly from the 0.1 and 0.4 mgkg^{-1} dosing groups. Linear regression analysis shows a very strong log[dose]-response relationship in the reduction of cysts with increasing concentrations of pentamidine (Pearson $R = 0.667$) and AM90067 (Pearson $R = 0.527$). Examination of these relationships reveals for pentamidine and AM90067 gradients of $-34.89 (\pm 7.11)$ and $-22.08 (\pm 6.20)$ respectively (standard error in parentheses). The gradient for pentamidine is considerably greater than that for AM90067, thereby suggesting, that compared to AM90067, pentamidine at a similar concentration is more effective at reducing the incidence of *P. carinii* cysts in immunosuppressed rats.

Thus, in summary, AM90067 clearly exhibits a modest degree of activity in reducing the number of *P. carinii* cysts per rat lung, though, in direct comparison with pentamidine, it does not perform as well as had been anticipated.

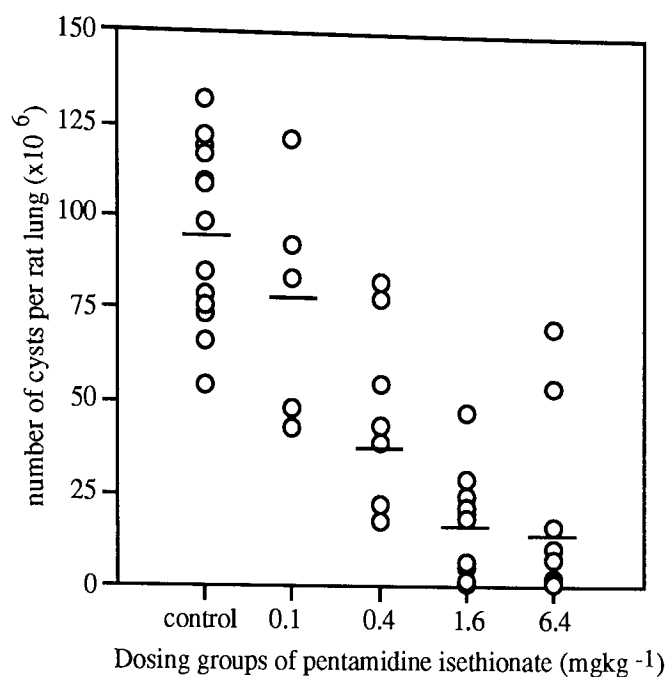


Figure 8.3 Scatter diagram of data points for the control and pentamidine isethionate dosing groups against the number of cysts per rat lung

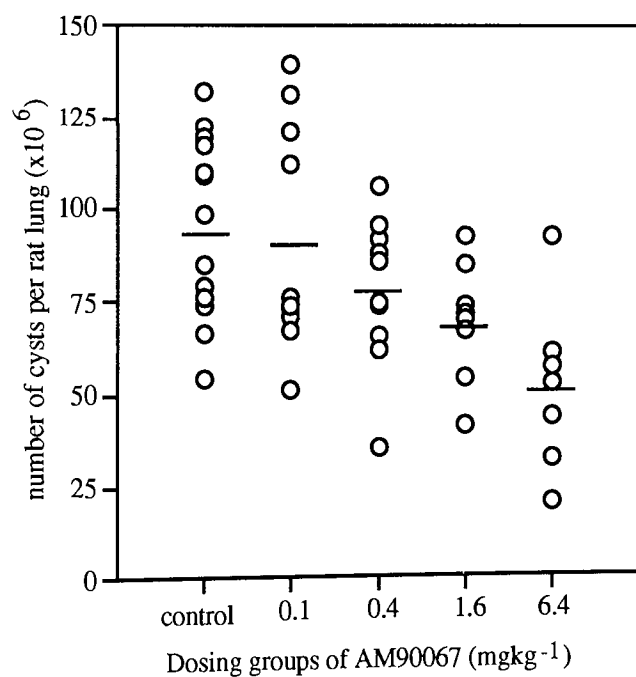


Figure 8.4 Scatter diagram for the control and AM90067 dosing groups against the number of cysts per lung

Chapter 9 Concluding Remarks

P. carinii infections remain life-threatening to the immunocompromised population, especially to those who develop the condition as a direct consequence of contracting AIDS, because of the lack of suitable antipneumocystis chemotherapy. Because of this there is an urgent need to develop a clinically acceptable antipneumocystis agent. A major objective of this project was to utilise computer-aided molecular modeling techniques to explain why the 3'-(3'',3''-dimethyltriazene-1''-yl) moiety in a pyrimethamine analogue is responsible for the greater selective inhibition of *P. carinii* DHFR. From this, the rational design of a second generation of selective inhibitors of *P. carinii* DHFR could be pursued.

In order to carry out this investigation, methods for accurately determining the steric and electronic parameters of the triazene moiety were necessary so that the 3'-(3'',3''-disubstitutedtriazene-1''-yl) pyrimethamine analogues could be studied. The experimentally determined AM90021 atomic coordinates provided an ideal starting geometry for the construction of the AM90020 analogues. These could then be subjected to quantum mechanical and molecular mechanical calculations to provide electronic distribution data. However, because of the molecular size of AM90020 it was not possible to complete any *ab initio* quantum mechanical calculations on the entire compound. Comparison studies on the model 2'-chloro-1'-(*N,N*-dimethyl) phenyltriazene established that the AM1 parameter set with the triazene moiety subjected to certain geometrical constraints gave the closest results to the *ab initio* calculations and therefore was most appropriate for the study of electronic properties of the phenyltriazene function. Comparison of the triazene moiety with the chemically related formamidine moiety revealed steric and electronic differences. However it remains to be confirmed if this is responsible for their entirely different DHFR inhibitory activity profiles.

The work presented in chapter 4 shows the DHFR inhibitory activity of the different pyrimethamine analogues. Considering the range of activities and the fact that these compounds only differ in the identity of the terminal triazene substituents, these substituents appear to have a considerable influence on the analogues' DHFR inhibitory activities. However, state of the art computational techniques were unable to establish any correlation between the biological activities and the calculated steric and electronic

parameters of the pyrimethamine analogues. Several reasons are suggested for this, including the limited number of compounds available with biological data. However, it was clear from this chapter as that the 3'-(3''-hydroxyethyl-3''-benzyltriazene-1''-yl) pyrimethamine analogue, AM90067, demonstrated the greatest inhibitory activity and the largest RL/PC inhibitory activity ratio and warranted further investigation.

CADD techniques have been developed to determine how a ligand is able to align in the active site of the target receptor. From this an assessment of the interaction between the ligand and the receptor can be made, which can be related directly to the biological activity of the ligand. Such studies require structural data of the target active site. As an extensively studied enzyme there is considerable structural data available for many species' DHFR. The results from chapter 7 outline the approaches used for the construction of models of *P. carinii* and human DHFR active sites. Although a very conservative approach was adopted for the construction of the *P. carinii* DHFR model, it should be emphasised that with this strategy any structural or electronic property that can be correlated with the biological activities may be amplified if incorporated into second generation compounds.

These active site models were used to investigate how AM90067 interacts with *P. carinii* and human DHFR enzymes. However, due to the flexible nature of AM90067 a combination of semi-empirical and molecular mechanical calculations were used in the conformational analysis techniques that identified minimum energy conformations. Results from chapter 6 present 24 low energy conformations of the 3'-(3''-hydroxyethyl-3''-benzyltriazene-1''-yl) substituent. One surprising result was that the rotation of the N1-N2-N3-C7 torsion (figure 5.1) produced two minima, one of which had a considerably lower energy and was therefore thermodynamically more favourable. Although NMR data confirmed the presence of two conformations and suggested that one was in greater abundance, it was unable to distinguish which of the two minima identified from the theoretical calculations was preferred. More detailed experimental conformational work would be necessary to identify which of the two conformations has the lower energy and compare it with the theoretical analysis. Further reductions in the number of low energy conformations that were required to be investigated were derived from crystallographic data for pyrimethamine binding and from chapter 4. This suggested that AM90067 was able to compete for occupancy of both the substrate DHF

and NADPH cofactor binding sites in the mammalian DHFR. By assuming that the equivalent atoms in AM90067 adopt a similar position to those of pyrimethamine when bound it was observed that the triazene group could be accommodated in the NADPH binding groove.

Although the energies of interaction calculated for the equivalent conformations of AM90067 did not differ significantly between the two active site models it was still possible to propose reasons for the greater inhibitory activity of AM90067 for *P. carinii* DHFR. The orientation of the AM90067 benzyl substituent to within 3 Å of a valine residue in *P. carinii* DHFR was sufficiently close that the substituent and amino acid side chain could form a hydrophobic van der Waals contact. The equivalent residue in human DHFR was threonine, more hydrophilic in nature, and therefore less likely to form a hydrophobic contact. These results indicate that the 3''-benzyl group is crucial for antipneumocystis activity. More recently the acetate ester of AM90067, AM92008 (figure 9.1), has also been shown to possess a even greater inhibitory activity of *P. carinii* over mammalian DHFR (personal communication from Dr. S.F. Queener). This further supports the importance of the group benzyl. Furthermore the results from chapter 4 showed that the 3''-methyl-3''-benzyl analogue, AM90035, exhibited a greater *P. carinii* DHFR activity, though not to the same extent as the other two 3''-benzyl analogues. It could be that other structural features may be involved in the ligand binding that neither model has been able to detect.

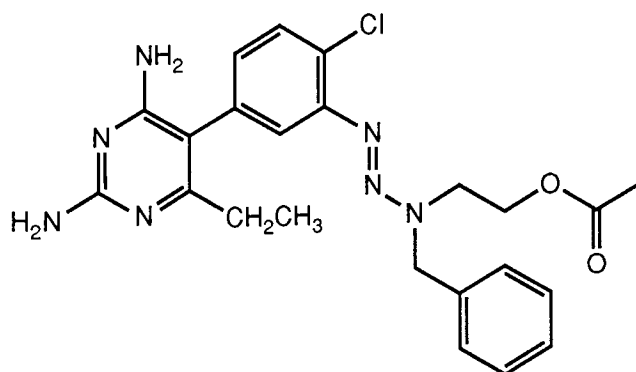


Figure 9.1 The structure of AM92008

Although the models have been able to establish a possible reason for the selective inhibitory activity of AM90067 for *P. carinii* DHFR they have not been able to rationalise the importance of the triazene moiety. Future work to confirm the

incorporation of the triazene function could involve investigating the effect that structurally similar moieties to the triazene have an inhibitory activity. To continue this drug design project second generation analogues of AM90067 should be synthesised to investigate the biological effect substitution of the benzyl moiety has.

The results from the *in vivo* evaluation in chapter 8 showed that AM90067 reduces the *P. carinii* cyst load in infected immunosuppressed rat lungs. The results also showed that AM90067 did not perform as well as pentamidine, one of the treatments of choice for PCP. However, this *in vivo* study was not designed to determine the mode of action of AM90067. Therefore it remains unclear whether the observed reduction in the *P. carinii* load is a direct consequence of selective inhibition of *P. carinii* DHFR by AM90067, or is due to some indirect action in the rats.

In conclusion, the initial objective to determine why the triazene substituent in the pyrimethamine analogue AM90020 is responsible for the increase in inhibitory activity against *P. carinii* DHFR has not been achieved. This is because the *in vitro* screening of other 3'',3''-disubstituted triazen-1''-yl analogues showed that the inclusion of the 3'-triazene moiety did not automatically confer *P. carinii* DHFR inhibitory activity. Although AM90067 was able to demonstrate antipneumocystis activity *in vivo*, the results would suggest that it would not be a suitable chemotherapeutic agent for the treatment of PCP because it is not as active as pentamidine, one of the treatments of choice. However, the molecular modeling established that there were structural differences between the active sites of *P. carinii* and human DHFR which could explain the *in vitro* activity of AM90067. Furthermore, as previously stated, the *in vivo* studies showed AM90067 to have some antipneumocystis activity. Therefore, for drug design purposes, AM90067 could serve as a new lead compound for the development of second generation selective inhibitors of *P. carinii* DHFR.

References

- Adams J.A., Fierke C.A. & Benkovic S.J.** (1991) The Function of the Amino Acid Residues Contacting the Nicotinamide Ring of NADPH in Dihydrofolate Reductase from *Escherichia coli*. *Biochemistry*. 30: 11046 - 11054
- Akhtar M.H., McDaniel R.S., Feser M. & Oehlschlager A.C.** (1968) NMR Study of Hindered Rotation in 1-Aryl-3,3-dimethyltriazenes. *Tetrahedron*. 24: 3899 - 3906
- Alberti-Flor J.J., Hernandez M.E., Ferref J.P., Howell S. & Jeffers L.** (1989) Fulimant Liver Failure and Pancreatitis Associated with the Use of Sulphamethoxazole-Trimethoprim. *American Journal of Gastroenterology*. 84: 1577 - 1579
- Allegra C.J., Kovacs J.A., Drake J., Swan J.C., Swan J.C., Chabner B.A. & Masur H.** (1987a) Activity of Antifolates Against *Pneumocystis carinii* Dihydrofolate Reductase and Identification of a Potent New Agent. *Journal of Experimental Medicine*. 165: 926 - 931
- Allegra J.C., Chabner B.A., Tuazon C.U., Ogata-Arakaki D., Baird B., Drake J.C., Simmons J.T., Lack E.E., Shelhamer J.H., Balis F., Walker R., Kovacs J.A., Lane H.C. & Masur H.** (1987b) Trimetrexate for the Treatment of *Pneumocystis carinii* Pneumonia in Patients with the Acquired Immunodeficiency Syndrome. *New England Journal of Medicine*. 317: 978 - 986
- Allen F.H., Davies J.E., Galloy J.J., Johnson O., Kennard O., MacCrae C.F., Mitchell E.M., Smith J.M. & Watson D.M.** (1991) The Development of Version 3 and Version 4 of the Cambridge Structural Database System. *Journal of Chemical Information and Computer Science*. 31: 187 - 204
- Andrea T.A. & Kalayeh H.** (1991) Applications of Neural Networks in Quantitative Structure-Activity Relationships of Dihydrofolate Reductase. *Journal of Medicinal Chemistry*. 34: 2824 - 2836
- Andrews P.R.** (1986) Functional Groups, Drug Receptor Interactions and Drug Design. *Trends in Pharmacological Sciences*. 7: 148 - 151
- Andrews P.R., Craig D.J. & Martin J.L.** (1984) Functional Group Contributions of Drug Receptor Interactions. *Journal of Medicinal Chemistry*. 27: 1648 - 1657
- Andrews P.R., Sadek M., Spark M.J. & Winkler D.A.** (1986) Conformational Energy Calculations and Electrostatic Potential of Dihydrofolate Reductase Ligands: Relevance to Mode of Binding and Species Specificity. *Journal of Medicinal Chemistry*. 29: 698 - 708
- Aoyama T. & Ichikawa H.** (1991a) Basic Operating Characteristics of Neural Networks When Applied to Structure-Activity Studies. *Chemical and Pharmaceutical Bulletin*. 39: 358 - 366
- Aoyama T. & Ichikawa H.** (1991b) Obtaining the Correlation Indices Between Drug Activity and Structural Parameters Using a Neural Network. *Chemical and Pharmaceutical Bulletin*. 39: 372 - 378

Argos P. (1987) A Sensitive Procedure to Compare Amino Acid Sequences. *Journal of Molecular Biology*. 193: 385 - 396

Baccanari D.P., Tansik R.L., Paterson S.J. & Stone D. (1984) Characterisation and Amino Acid Sequence of *Neisseria gonorrhoea* Dihydrofolate Reductase. *Journal of Biological Chemistry*. 259: 12291 - 12298

Baker D.J., Beddell C.R., Champness J.N., Goodford P.J., Norrington F.E., Roth B & Stammers D.K. (1983) *Chemistry and Biology of Pteridines*: 545 - 549 Editor Blair J. A. Published by Walter de Gruyter & Co., Berlin.

Baker N.J., Griffin R.J., Irwin W.J., Reichert U. & Schmidt R. (1991) Uptake, Cytotoxicity and Metabolism of *meta*-Azidopyrimethamine and Related Lipophilic Antifolates in SV-K14 Human Keratinocytes. *Biochemical Pharmacology*. 42: S124

Bartizal K., Abruzzo G., Trainor C., Krupa D., Nollstadt K., Schmartz D., Hammond M., Balkovec J. & Van Middlesworth F. (1992) In Vitro Antifungal Activities and In Vivo Efficacies of 1,3- β -D-Glucan Synthesis Inhibitors L-671,329, L-646,991, Tetrahydroechinocandin B and L687,781, a Papulocandin. *Antimicrobial Agents and Chemotherapy*. 36: 1648 - 1657

Bartlett M.S., Edlind T.D., Durkin M.M., Shaw M.M., Queener S.F. & Smith J.W. (1992) Antimicrotubule Benzimidazole Inhibits the In Vitro Growth of *Pneumocystis carinii*. *Antimicrobial Agents and Chemotherapy*. 36: 779 - 782

Bartlett M.S., Queener S.F., Tidwell R.R., Milhous W.K., Beman J.D., Ellis W.Y. & Smith J.W. (1991) 8-Aminoquinolines from the Walter Reed Army Institute for Research for the Treatment and Prophylaxis of *Pneumocystis carinii* Pneumonia in Rats. *Antimicrobial Agents and Chemotherapy*. 35: 277 - 282

Batley K.E. & Morris H.R. (1977) Dihydrofolate Reductase from *Lactobacillus casei*: N-terminal Sequence and Comparison with the Substrate Binding Region of Other Reductases. *Biochemical and Biophysical Research Communications*. 75: 1010 - 1014

Berkowitz C.D. (1985) AIDS and Parasitic Infections, Including *Pneumocystis carinii* and Cryptosporidiosis. *Pediatric Clinics of North America*. 32: 933 - 952

Bernard E.M., Sepkowitz K.A., Telzak E.E. & Armstrong D. (1992) Pneumocystosis. *Medical Clinician of North America*. 76: 107 - 119

Bernstein F.C., Koetzle T.F., Williams G.J.B., Meyer E.F., jr., Brice M.D., Rogers J.R., Kennard O., Shimaouchi T. & Tasumi M. (1977) The Protein Databank: a Computer-based Archival File for Macromolecular Structures. *Journal of Molecular Biology*. 112: 535 - 542

Bertino J. & Fisher M. (1964) Techniques for the Study of Resistance to Folic Acid Antagonists. *Methods in Medical Research*. 10: 297 - 307

Birdsall B., Tendler S.J.B., Arnold J.R.P., Feeney J., Griffin R.J., Carr M.D., Thomas J.A., Roberts G.C.K. & Stevens M.F.G. (1990) NMR Studies of Multiple Conformations in Complexes with *Lactobacillus casei* Dihydrofolate Reductase with Analogues of Pyrimethamine. *Biochemistry*. 29: 9660 - 9667

Blaney J.M., Hansch C., Silipo C. & Vittoria A. (1984) Structure-Activity Relationships of Dihydrofolate Reductase Inhibitors. *Chemical Reviews*. 84: 333 - 407

Bliss E.A., Griffin R.J. & Stevens M.F.G. (1987) Structural Studies on Bioactive Compounds. 5. Synthesis and Properties of 2,4-Diaminopyrimidine Dihydrofolate Reductase Inhibitors Bearing Lipophilic Groups. *Journal of the Chemical Society, Perkin Transactions 1*. 2217 - 2228

Blundell T.L., Carney D., Gardner S., Hayes F., Howlin B., Hubbard T., Overington J., Singh D.A., Sibanda B.L. & Sutcliffe M. (1988) Sir Hans Krebs Lecture: Knowledge-based Protein Modeling and Design. *European Journal of Biochemistry*. 172: 513 - 520

Bolin J.T., Filman D.J., Matthews D.J., Hamlin R.C. & Kraut J. (1982) Crystal Structures of *Escherichia coli* and *Lactobacillus casei* Dihydrofolate Reductase Refined at 1.7 Å Resolution. 1. General Features and Binding of Methotrexate. *Journal of Biological Chemistry*. 257: 13650 - 13662

Bonatti L., Cosentino E., Frascini G.M. & Pitea D. (1992) Molecular Electrostatic Potential of Substituted Aromatic Compounds: Factors Affecting the Differences Between *Ab Initio* and Semi-empirical Results. *Journal of Computational Chemistry*. 13: 842 - 850

Bondii (1964) van der Waals Volumes and Radii. *Journal of Physical Chemistry*. 68: 441 - 457

Boylan C.J. & Current W.L. (1992) Improved Rat Model of *Pneumocystis carinii* Pneumonia-Induced Laboratory Infections in *Pneumocystis*-Free Animals. *Infection and Immunity*. 60: 1589 - 1597

Brooks B.R., Bruccoleri R.E., Olafson B.D., States D.J., Swaminathan S. & Karplus M. (1983) CHARMM - A Program for Macromolecular Energy, Minimisation and Dynamics. *Journal of Computational Chemistry*. 4: 187 - 217

Broughton M.C. & Queener S.F. (1991) *Pneumocystis carinii* Dihydrofolate Reductase Used to Screen Potential Antipneumocystis Drugs. *Antimicrobial Agents and Chemotherapy*. 35: 1348 - 1355

Browne W.J., North A.C.T., Phillips D.C., Brew K., Vanaman T.C. & Hill T.C. (1969) Possible Structure of α -Lactalbumin. *Journal of Molecular Biology*. 42: 65 - 86

Bryant P.K., Wing K.P., Colby J., Schwalbe C.H., Stevens M.F.G., Griffin R.J. & Bliss E.A. (1986) *Chemistry and Biology of Pteridines*. 1005 - 1008. Editors Cooper B.A. & Whitehead V.M. Published by Walter de Gruyter & Co., Berlin.

Burchall J.J. & Hitchins G.H. (1965) Inhibitor Binding Analysis of Dihydrofolate Reductase from Various Species. *Molecular Pharmacology*. 1: 126 - 136

Burt C., Richards W.G. & Huxley P. (1990) The Application of Molecular Similarity Calculations. *Journal of Computational Chemistry*. 11: 1139 - 1146

Carbo R., Leyda L. & Arnau M. (1980) An Electron Density Measure of the Similarity Between Two Compounds. *International Journal of Quantum Chemistry*. 17: 1185 - 1190

Chagas C. (1909) Nova Tripanomiazaea Humana. Ueber eine neue Trypanosomiasis de Menschen. *Mem. Inst. Oswaldo Cruz.* 1: 159

Chalmers D.K., Munro S.L.A., Iskander M.N. & Craik D.J. (1992) Models for the Binding of Amiodarone to the Thyroid Model Receptor. *Journal of Computer-Aided Molecular Design.* 6: 19 - 31

Champness J.N., Stammers D.K. & Beddell C.R. (1986) Crystallographic Investigation of the Cooperative Interaction Between Trimethoprim, Reduced Cofactor and Dihydrofolate Reductase. *FEBS.* 199: 61 - 67

Chan J.H. & Roth B. (1991) 2,4-Diamino-5-benzylpyrimidines as Antibacterial Agents. 14. 2,3-Dihydro-1-(2,4-diamino-5-pyrimidyl)-1*H*-indenes as Conformationally Restricted Analogues of Trimethoprim. *Journal of Medicinal Chemistry.* 34: 550 - 555

Charlton P.A., Young D., Birdsall B., Feeney J. & Roberts G.C.K. (1979) Stereochemistry of Reduction of Folic Acid Using Dihydrofolate Reductase. *Journal of the Chemical Society, Chemistry Communications.* : 922 - 924

Chen M.-J., Chi C.-S. & Chen Q.-Y. (1990) Reaction of Amines or Amides with Phthaloylchloride. Synthesis of Flourine-Containing *N,N*-dimethyl-*N'*-Substituted Amidines and Related Compounds. *Journal of Flourine Chemistry.* 49: 99 - 106

Cheung T.W., Matta R., Neibart E., Hammer G., Chusid E., Sacks H.S., Szabo S. & Rose D. (1993) Intramuscular Pentamidine for the Prevention of *Pneumocystis carinii* Pneumonia in Patients Infected with Human Immunodeficiency Virus. *Clinical Infectious Diseases.* 16: 22 - 27

Chopra G. & Schwalbe C.H. (1991) Structural Studies on Bioactive Compounds. Part 18. Crystal Structure and Molecular Modeling of Methyl-4-(3-Ethyl-3-Hydroxymethyltriazene-1-yl)benzoate. *Anti-Cancer Drug Design.* 6: 495 - 500

Chothia C. & Lesk A.M. (1986) The Relation Between the Divergence of Sequence and Structures in Proteins. *EMBO Journal.* 5: 823 - 826

Chou P.Y. & Fasman G.D. (1974) Conformational Parameters for Amino Acids in Helical, Beta-sheet and Random coil Regions Calculated from Proteins. *Biochemistry.* 13: 211 - 222

Chou P.Y. & Fasman G.D. (1974) Prediction of Protein Conformation. *Biochemistry.* 13: 222 - 245

Ciszak E., Gdaniec M., Jackoski M., Kostarkiewicz Z., Owsianski J. & Tykarska E. (1989) The Structures of 3 *N*₂,*N*₁-Dimethyl-*N*₂-*para* - nitrophenylamidine, Formpropanamidine, Acetopropanamidine and 2,2-Dimethylpropamidine. *Acta Crystallographica.* C45: 433 - 438

Clark T. (1985) *A Handbook of Computational Chemistry. A Practical Guide to Chemical Structure and Energy Calculations.* Publishers, Jon Wiley & Sons, Inc.

Clotet B., Sirera G., Romeu J., Gimeno J.M., Condom M.J., Tor J. & Foz M. (1991) Twice Weekly Dapsone-Pyrimethamine for Preventing PCP and Cerebral Toxoplasmosis. *AIDS.* 5: 601 - 602

Cody V., Luft J.R., Ciszak E., Kalman T.I. & Freisheim J.H. (1992) Crystal Structure Determination at 2.3 Å Resolution of Recombinant Human Dihydrofolate Reductase Ternary Complex with NADPH and Methotrexate-amm-tetrazole. *Anti-Cancer Drug Design*. 7: 483 - 491

Cody V. & Zakrzewski S.F. (1982) Molecular Studies of 2,4-Diaminopyrimidines Antifolates with Neoplastic Activity. *Journal of Medicinal Chemistry*. 25: 427 - 430

Coker R.J., Nieman R., McBride M., Mitchell D.M., Harris J.R.W. & Weber J.N. (1992) Co-trimoxazole versus Dapsone-Pyrimethamine for the Prevention of *Pneumocystis carinii* Pneumonia. *Lancet*. 340: 1099 - 1102

Cushion M.T., Stringer J.R. & Walzer P.D. (1991) Cellular and Molecular Biology of *Pneumocystis carinii*. *International Reviews in Cytology*. 131: 59 - 107

Dannemann B., McCutchan J.A., Israelski D., Antoniskis D., Leport C., Luft B., Nussbaum J., Clumeck N., Morlat P., Chiu J., Vilde J.M., Orellana J.M., Feigal D., Barlok A., Heseltine P., Leedon J. & Remington J. (1992) Treatment of Toxoplasmic Encephalitis in Patients with AIDS. A Randomised Trial Comparing Pyrimethamine and Sulfadiazine with Pyrimethamine and Clindamycin. *Annals of Internal Medicine*. 116: 33 - 43

Davey R.T. & Masur H. (1990) Recent Advances in the Diagnosis, Treatment, and Prevention of *Pneumocystis carinii* Pneumonia. *Antimicrobial Agents and Chemotherapy*. 34: 499 - 504

Davies J.F., Delcamp T.J., Prendergast N.J., Ashford V.A., Freisheim J.H. & Kraut J. (1990) Crystal Structures of Recombinant Human Dihydrofolate Reductase Complexed with Folate and 5-Deazafolate. *Biochemistry*. 29: 9467 - 9479

Dayhoff M.O., Barker W.C. & Hunt L.T. (1983) Establishing Homologies in Protein Sequences. *Methods in Enzymology*. 91: 524 - 545

Dei-Cas E., Cailliez J.C., Palluault F., Aliouat E.M., Mazars E., Soulez B., Suppin J. & Camus D. (1992) Is *Pneumocystis carinii* a Deep Mycosis-like Agent? *European Journal of Epidemiology*. 8: 460 - 470

Delanoe P. & Delanoe M. (1910) Sur les Rapports der Kystes de Carinii du Pulmon Rats avec *Trypanosoma lewisii*. *C.R. Acad. Sci. Paris*. 155: 658

Denny B.J., Ringan N.S., Schwalbe C.H., Lambert P.A., Meek M.A., Griffin R.J. & Stevens M.F.G. (1992) Structural Studies on Bioactive Compounds. Part 20. Molecular Modeling and Crystallographic Studies on Methylbenzoprim, a Potent Inhibitor of Dihydrofolate Reductase. *Journal of Medicinal Chemistry*. 35: 2315 - 2320

Dewar M.J.S., Healey E.F., Holder A.J. & Yuan Y.-C. (1990) Comments on a Comparison of AM1 with the Recently Developed PM3 Method. *Journal of Computational Chemistry*. 11: 541 - 542

Dewar M.J.S. & Theil W. (1977) Ground States of Molecules. 39. MNDO Results for Molecules Containing Hydrogen, Carbon, Nitrogen and Oxygen. *Journal of the American Chemical Society*. 99: 4907 - 4922

- Dewar M.J.S., Zoebisch E.G., Healy E.F. & Stewart J.J.P.** (1985) AM1: A New General Purpose Quantum Mechanical Molecular Model. *Journal of the American Chemical Society*. 107: 3902 - 3909
- Dohn M.N., Frame P.T., Baughman R.P., Lafon S.W., Smulian A.G., Caldwell P. & Rogers M.D.** (1991) Open-label Efficacy and Safety Trial of 42 Days of 566C80 for *Pneumocystis carinii* Pneumonia in AIDS Patients. *Journal of Protozoology*. 38:S141 - S142
- Dolata D.P., Leach A.R. & Prout K.** (1990) Molecular Modeling by Symbolic Logic. *Computer-Aided Molecular Design*. Edited by Richards W.G. Publishers IBC Technical Services, London.
- Domenighini M., Montecucco C., Ripka W.C. & Rappuoli R.** (1991) Computer Modeling of the NAD Binding Site of ADP-Ribosylating Toxins: Active-Site Structure and Mechanism of Binding. *Molecular Microbiology*. 5: 23 - 31
- Doolittle R.F.** (1989) Similar Amino Acid Sequences Revisited. *Trends in Biochemical Sciences*. 14: 244 - 245
- Doolittle R.F.** (1981) Similar Amino Acid Sequences - Chance or Common Ancestry. *Science*. 214: 149 - 159
- Dunn W.J., III** (1990) Quantitative Structure Activity Relationships. *Drug Discovery Technologies*. Edited by Clark C.R. & Moos W.H. Publishers Ellis Horwood Ltd., Chichester.
- Dunn W.J., Lins C., Kumar G., Manimaram T., Grigoras S., Edlund U. & Wold S.** (1983) Substituent Effects on ^{13}C and ^{15}N Chemical Shifts in Triazines Studied by Principal Components Multivariate Data Analysis. *Organic Magnetic Resonance*. 21: 450 - 456
- Durkin M.M., Bartlett M.S., Queener S.F., Shaw M.M., Lee C.H. & Smith J.W.** (1992) An Enzyme-Linked Immunosorbent Assay for Enumeration of *Pneumocystis carinii* In Vitro and In vivo. *Journal of Clinical Microbiology*. 30: 3258 - 3262
- Edman J.C., Edman U., Cao M., Lundgren B., Kovacs J. & Santi D.V.** (1989a) Isolation and Expression of the *Pneumocystis carinii* Dihydrofolate Reductase Gene. *Proceedings of the National Academy of Sciences U.S.A.* 86: 8625 - 8629
- Edma J.C., Kovacs J.A., Masur H., Santi D.V., Elwood H.J. & Sogin M.L.** (1988) Ribosomal RNA Sequence Shows *Pneumocystis carinii* to be a Member of the Fungi. *Nature*. 334: 519 - 522
- Edman U., Edman J.C., Lundgren B. & Santi D.V.** (1989b) Isolation and Expression of the *Pneumocystis carinii* Thymidylate Synthase Gene. *Proceedings of the National Academy of Sciences of the U.S.A.* 86: 6503 - 6507
- Edwards M.S., Sternberg M.J.E. & Thornton J.M.** (1987) Stuctural and Sequence Patterns Between Loops of $\beta\alpha\beta$ Units. *Protein Engineering*. 1: 173 - 181
- Edwards S.L., Chapius G., Templeton D.H. & Zalkin A.** (1977) *Acta Crystallographica*. B33: 276 - 278

- Eeftinck Schaffenkirk J.K., Lange M., van Steenwijk R.P. & Danner S.A.** (1990) Can the Course of High Dose Cotrimoxazole for *Pneumocystis carinii* Pneumonia in AIDS be Shorter? A Possible Solution to the Problem of Cotrimoxazole Toxicity. *Journal of Internal Medicine*. 227: 359 - 362
- Falloon J., Kovacs J.A., Hughes W.T., O'Neill D., Polis M., Davey R.T., jr., Rogers M., Lafon S., Feuerstein I., Lancaster D., Land M., Tuazon C., Dohn M.N., Greenberg S., Lane H.C. & Masur H.** (1991) A Preliminary Evaluation of 566C80 for the Treatment of *Pneumocystis carinii* Pneumonia in Patients with the Acquired Immunodeficiency Syndrome. *New England Journal of Medicine*. 325: 1534 - 1538
- Falloon J., Allegra C., Kovacs J., O'Neill D., Ogataaraki D., Feuerstein I., Poplis M., Davey R., Lane H.C., Lafon S., Rogers M., Zurich K., Zurlo J., Tuazon C., Parenti D., Simon G. & Masur H.** (1990) Piritrexim with Leucovorin for the Treatment of *Pneumocystis* Pneumonia (PCP) in AIDS Patients. *Clinical Research*. 38:S2014
- Farid H., Shenkin P.S., Greene J. & Fetrow J.** (1991) Prediction of Side Chain Conformations in Protein Cores and Loops from Rotamer Libraries. *FASEB Journal*. 6: 2014
- Ferenczy G.G. & Garrett M.M.** (1989) The Active Site of Cytochrome P-450 Nifedipine Oxidase: a Model Building Study. *Journal of Molecular Graphics*. 7: 206 - 211
- Filman D.J., Bolin J.T., Matthews D.J. & Kraut J.** (1982) Crystal Structures of *Escherichia coli* and *Lactobacillus casei* Dihydrofolate Reductase Refined at 1.7 Å Resolution. 2. Environment of Bound NADPH and Implications for Catalysis. *Journal of Biological Chemistry*. 257: 13663 - 13672
- Folkers G., Trumppkallmeyer S., Gutbrod O., Krickl S., Fetzer J. & Keil G.M.** (1991) Computer-Aided Active-Site-Directed Modeling of the *Herpes simplex* Virus-1 and Human Thymidine Kinase. *Journal of Computer-Aided Molecular Design*. 5: 385 - 404
- Fonticella-Camps J.C., Bugg C.E., Temple C., Rose J.D., Montgomery J.A. & Kisliuk R.L.** (1979) Absolute Configuration of Biological Tetrahydrofolates. A Crystallographic Determination. *Journal of the American Chemical Society*. 101: 6114 - 6115
- Frenkel J.K., Good J.T., Schultz J.A.** (1966) A Latent *Pneumocystis* Infection in Rats, Relapse and Chemotherapy. *Laboratory Investigation*. 15: 1559 - 1585
- Fronczek F.R., Hansch C. & Watkins S.F.** (1988) P-(3,3-Dimethyl-1-Triazeno)Benzonitrile. *Acta Crystallographica*. C44: 1650 - 1651
- Gallant J.E., Hoehn-Saric E. & Smith M.D.** (1991) Respiratory Insufficiency from Dapsone Induced Methemoglobinaemia. *AIDS*. 5: 1392 - 1393
- Garnier J., Osguthorpe D.J. & Robson B.** (1978) Analysis of the Accuracy and Implications of Simple Methods for Predicting the Secondary Structure of Globular Proteins. *Journal of Molecular Biology*. 120: 97 - 120

- Garvin L.D. & Hardies S.C.** (1991) Temporal and Topological Clusterings of Diverged Residues Among Enterobacterial Dihydrofolate Reductase Genes. *Molecular Biology and Evolution*. 8: 654 -668
- Gasteiger J. & Marsili M.** (1980) Iterative Partial Equalisation of Orbital Electronegativity - A Rapid Access to Atomic Charges. *Tetrahedron*. 36: 3219 - 3228
- Gazzard B.** (1989) *Pneumocystis carinii* Pneumonia and its Treatment in Patients with AIDS. *Journal of Antimicrobial Chemotherapy*. 23: S67 - S75
- Gifkins M.R. & Jacobson R.A.** (1980) *N'*-(4-chloro-*o*-tolyl)-*N,N*-dimethylformamidine. *Crystal Structure Communications*. 9: 959 - 963
- Gigliotti F., Haidaris P.J., Haidaris C.G., Wright T.W. & van der Meid K.R.** (1993) Further Evidence of Host Species-Specific Variations in Antigens of *Pneumocystis carinii* Using the Polymerase Chain Reaction. *Journal of Infectious Diseases*. 168: 191 - 194
- Gilbert A.S. & Hyde R.M.** (1992) Drug Discovery and Invention: Some Approaches Compared. *The Design of Drugs to Macromolecular Targets*. Chapter 8: 249 - 261. Edited by Beddell C.R. Published by John Wiley & Sons Ltd.
- Good A.C., So S.S. & Richards W.G.** (1993) Structure-Activity Relationships from Molecular Similarity Matrices. *Journal of Medicinal Chemistry*. 36: 433 - 438
- Good A.C., Peterson S.J. & Richards W.G.** (1993) QSAR's from Similarity Matrices. Technique Validation and Application in the Comparison of Different Similarity Evaluation Methods. *Journal of Medicinal Chemistry*. 36: 2929 - 2937
- Gordin F., Simon G., Wofsy C. & Mills J.** (1984) Adverse Reactions to Trimethoprim-Sulphamethoxazole in Patients with the Acquired Immunodeficiency Syndrome. *Annals of Internal Medicine*. 100: 495 - 499
- Graziano M.J., Kim S.M., MacDonald J.R. & Watkins J.R.** (1992) Toxicity of the Anticancer Folate Antagonist in Rats. *Fundamental and Applied Toxicology*. 10: 115 - 125
- Greenwood D.** (1989) *Antimicrobial Chemotherapy*. 66 - 74. Publishers Oxford Medical Press.
- Greer J.** (1981) Comparative Model-building of the Serine Proteases. *Journal of Molecular Biology*. 153: 385 - 396
- Griffin R.J. & Lowe P.R.** (1992) Acetylated pyrimethamine analogues *Journal of the Chemical Society, Perkin Transactions 2*. 1811 - 1819
- Griffin R.J., Meek M.A., Schwalbe C.H. & Stevens M.F.G.** (1989) Synthesis, Crystal Structure and Biological Properties of a New Series of 2,4-Diamino-5-aryl-6-ethylpyrimidines. Dihydrofolate Reductase Inhibitors with *in vivo* Activity Against a Methotrexate Resistant Tumour Cell Line. *Journal of Medicinal Chemistry*. 32: 2468 - 2474
- Grumant R., Washtien W.L., Caput D. & Santi D.V.** (1986) Bifunctional Thymidylate Synthase-Dihydrofolate Reductase from *Leishmania tropica*. Sequence Homology with the Corresponding Monofunctional Proteins. *Proceedings of the National Academy of Sciences of the U.S.A.* 83: 5387 - 5391

- Ham E.K., Greenberg S.D., Reynolds R.C. & Singer D.B.** (1971) Ultra Structure of *Pneumocystis carinii*. *Experimental Molecular Pathology*. 14: 362 - 372
- Hansch C., Clayton J., Fujita T. & Muir R.** (1962) Correlation of the Biological Activity of Phenoxyacetic Acids with Hammett Substituent Constants and Partition Co-efficients. *Nature*. 194: 178 - 180
- Havel T.F. & Snow M.E.** (1991) A New Method for Building Protein Conformations from Sequence Alignments with Homologues of Known Structure. *Journal of Molecular Biology*. 217: 1 - 7
- Heald A, Flepp M., Chave J.P., Malinverni R., Ruttimann S., Gabriel V, Renold C., Sugar A. & Hirschel B.** (1991) Treatment of Cerebral Toxoplasmosis Protects Against *Pneumocystis carinii* Pneumonia in Patients with AIDS. The Swiss Cohort Study. *Annals of Internal Medicine*. 115: 760 - 763
- Hendley J.O. & Weller T.H.** (1971) Activation and Transmission in Rats of Infection with *Pneumocystis*. *Proceedings of the Society of Experimental and Biological Medicine*. 137: 1401 - 1404
- Higgins D.G. & Sharp P.M.** (1988) Clustal: A Package for Performing Multiple Sequence Alignments. *Gene*. 73: 237 - 244
- Hitchings G.H. & Burchall J.J.** (1965) Inhibition of Folate Biosynthesis and Function as a Basis for Chemotherapy. *Advanced Enzymology*. 27: 417 - 468
- Hodgkin E.E. & Richards W.G.** (1987) Molecular Similarity Based on Electrostatic Potential and Electric Field. *International Journal of Quantum Chemistry. Quantum Biology Symposium*. S14: 105 - 110
- Hopfinger A.J.** (1980) A QSAR Investigation of DHFR Inhibition by Bakers Triazines Based Upon Molecular Shape Analysis. *Journal of the American Chemical Society*. 102: 7196 - 7206.
- Howe W.J. & Moon J.B.** (1993) A Computational Method for the *De Novo* Design of Organic Peptidomimetic Ligands. *Journal of Cellular Biology*. S17C: P224
- Hughes W.T., McNabb P.C., Makres T.D. & Feldman S.** (1974) Efficacy of Trimethoprim and Sulphamethoxazole in the Prevention and Treatment of *Pneumocystis carinii* Pneumonia. *Antimicrobial Agents and Chemotherapy*. 5: 289 - 293
- Hughes W.T.** (1976) Treatment of *Pneumocystis carinii* Pneumonitis. *New England Journal of Medicine*. 295: 726 - 727
- Hughes W.T., Kuhn S., Chaudbray S., Feldman S., Verzosa M., Aur R.J.A., Pratt C. & George S.C.** (1977) Successful Chemoprophylaxis for *Pneumocystis* Pneumonitis. *New England Journal of Medicine*. 297: 1419 - 1426
- Hughes W.T.** (1982) Natural Mode of Acquisition for *de novo* Infection with *Pneumocystis carinii*. *Journal of Infectious Diseases*. 145: 842 - 847
- Hughes W.T.** (1987) *Pneumocystis carinii* Pneumonia. 4 - 20. CRC Press, Boca Raton, Florida.

Hughes W.T. (1991) Prevention and Treatment of *Pneumocystis carinii* Pneumonia. *Annual Reviews in Medicine*. 42: 287 - 295

Hughes W.T., Bartley D.L. & Smith B.M. (1983) A Natural Source of Infection due to *Pneumocystis carinii*. *Journal of Infectious Diseases*. 147: 595

Hughes W.T., Gray V.L., Gutteridge W.E. Latter V.S. & Pudney M. (1990) Efficacy of a Hydroxynaphthoquinone, 566C80, in Experimental *Pneumocystis carinii* Pneumonitis. *Antimicrobial Agents and Chemotherapy*. 34: 225 - 228

Hughes W.T., Jacobus D.P., Canfield C. & Killmar J (1993) Anti-*Pneumocystis carinii* Activity of PS-15, a New Bi-guanide Folate Antagonist. *Antimicrobial Agents and Chemotherapy*. 37: 1417 - 1419

Ivaldy G. & Paldy L. (1958) Ein neues Behandlungsverfahren der Interstitiellen Plasmazelligen Pneumonie Frühgeborener mit Fünfwertigen Stibium und Aromatischen Diamidinen. *Monatsschr. Kinderheilkd.* 106: 10

Iwakura M., Kawata M., Tsuda K. & Tanaka T. (1988) Nucleotide Sequence of the Thymidylate Synthase-B and Dihydrofolate Reductase Genes contained in one *Bacillus subtilis* operon. *Gene*. 64: 9 - 20

Jackson H.C., Colthrust D., Hancock V., Marriott M.S. & Tuite M.F. (1991) No Detection of Characteristic Fungal Protein Elongation Factor EF-3 in *Pneumocystis carinii*. *Journal of Infectious Diseases*. 163: 675 - 677

Johnson C.K. (1976) ORTEP. Report ORNL-5138, Oak Ridge National Laboratory, Tennessee, U.S.A.

Johnson M.P., Goodwin S.D. & Shands J.W., jr. (1990) Trimethoprim-Sulphamethoxazole Anaphylactoid Reaction in Patients with AIDS: Case Reports and Literature Reviews. *Pharmacotherapy*. 10: 413 - 416

Jones R.S., jr., Collier-Brown C. & Suh B. (1992) Localized Subcutaneous Reaction to Intravenous Pentamidine. *Clinical Infectious Diseases*. 15: 561 - 562

Jones S.K., Hall J.E., Allen M.A., Morrison S.D., Ohemeng K.A., Reddy V.V., Geratz J.D. & Tidwell R.R. (1990) Novel Pentamidine Analogues in the Treatment of Experimental *Pneumocystis carinii* Pneumonia. *Antimicrobial Agents and Chemotherapy*. 34: 1026 - 1030

Joyner S.S., Fling M.E., Stone D. & Baccanari D.P. (1984) Characterisation of an R-Plasmid Dihydrofolate Reductase with a Monomeric Structure. *Journal of Biological Chemistry*. 259: 5851 - 5856

Kabsch W. & Sander C. (1984) On the Use of Sequence Homologies to Predict Protein Structure: Identical Pentapeptide Sequences Can Have Completely Different Conformations. *Proceedings of the National Academy of Sciences of the U.S.A.* 81: 1075 - 1078

Kim S.-H. & Cech T. (1987) Three-Dimensional Model of the Active Site of the Self-Splicing rRNA Precursor of *Tetrahymena*. *Proceedings of the National Academy of Science U.S.A.* 84: 8788 - 8792

- Kim C.K., Foy J.M., Cushion M.T., Stanforth D., Linke M.J., Hendrix H.L. & Walzer P.D.** (1987) Comparison of the Histologic and Quantitative Techniques in Evaluation of Therapy for Experimental *Pneumocystis carinii* Pneumonia. *Antimicrobial Agents and Chemotherapy*. 31: 197 - 201
- Koetzle J.S. & Williams M.** (1976) The Crystal and Molecular Structure of Trimethoprim 2,4-diamino-5-(3,4,5-trimethoxybenzyl)pyrimidine. *Journal of the American Chemical Society*. 98: 2074 - 2078
- Kolaskar A.S. & Kulkarni-Kale U.** (1992) Sequence Alignment Approach to Pick Up Conformationally Similar Protein Fragments. *Journal of Molecular Biology*. 223: 1053 - 1061
- Kondrashev Y.D.** (1974) Crystal Structure of *p*-nitrodiazoniumaminobenzene. *Zh. Strukt. Khim.* 15: 441-446
- Kovacs J.A., Allegra C.J., Chabner B.A., Swan J.C., Drake J., Lundle M., Pamillo J.E. & Masur H.** (1987) Potent Effect of Trimetrexate, a Lipid Soluble Antifolate, on *Toxoplasma gondii*. *Journal of Infectious Diseases*. 155: 1027 - 1032
- Kovacs J.A., Allegra C.J., Beaver J., Boarman D., Lewis M., Parrillo J.E., Chabner B.A. & Masur H.** (1989) Characterisation of *de novo* Folate Synthesis in *Pneumocystis carinii* and *Toxoplasma gondii*. Potential Utilisation for Screening Therapeutic Agents. *Journal of Infectious Disease*. 160: 312 - 320
- Kovacs J.A., Allegra C.J. & Masur H.** (1990) Characterisation of Dihydrofolate Reductase from *Pneumocystis carinii* and *Toxoplasma gondii*. *Experimental Parasitology*. 71: 60 - 68
- Krajewski J.W., Urbanczik-Lipowska Z., Gluzinski P., Busko-Oszczapowicz I., Oszczapowicz J., Bleidelis Y.A. & Kemme A.** (1981) Amidines, 8. X-Ray Investigation on Conformation of Amidine Group in N1-(hexylene-1,6)-N2-*para*-Nitrophenylformamidine. *Polish Journal of Chemistry*. 55: 1015 - 1024
- Kumar A.A., Blankenship D.T. Kaufman B.T. & Freisheim J.H.** (1980) Primary Structure of Chicken Liver Dihydrofolate Reductase. *Biochemistry*. 19: 667 - 678
- Lagosky P.A., Taylor G.R. & Haynes R.H.** (1987) Nucleotide Sequence of the Dihydrofolate Reductase Gene of *Saccharomyces cerevisiae*. *Nucleic Acid Research*. 15: 10355 - 10355
- Lai P.-H., Pan Y.C.E., Gleisner J.M., Peterson D.L., Williams K.R. & Blackley R.L.** (1982) Structure of Dihydrofolate Reductase - Primary Sequence of the Bovine Liver Enzyme. *Biochemistry*. 21: 3284 - 3294
- Lane A., Hissey P.H. & Jackson H.C.** (1991) Detection and Quantification of *Pneumocystis carinii* Using a Sandwich ELISA. *Journal of Protozoology*. 38: S124 - S216
- Latajka Z., Ratajczak H., Scheiner S. & Barycki J.** (1991) AM1 and *ab initio* Studies of Aminomethylphosphonic Acid. *Journal of Molecular Structure (Theochem)*. 235: 417 - 422

- Leban I.** (1974) *N,N*-dimethyl-*N'*-(6-chloropyridiazin-3-yl)formamidine. *Crystal Structure Communications*. 3: 249 - 252
- Lee B.L., Medina I., Benowitz N.L., Jacob P., III, Wofsy C.B. & Mills J.** (1989) Dapsone, Trimethoprim and Sulfamethoxazole Plasma Levels during Treatment of *Pneumocystis* Pneumonia in Patients with Acquired Immunodeficiency Syndrome. *Annals of Internal Medicine*. 110: 606 - 611
- Leighton R.R.** (1992) The Aspirin/MIGRAINES V6.0 Neural Network, the MITRE Corporation.
- Lesk A. M. & Chothia C.** (1980) How Different Amino Acid Sequences Determine Similar Protein Structures: The Structure and Evolutionary Dynamics of the Globins. *Journal of Molecular Biology*. 136: 225 - 270
- Lewis D.F.V. & Moereels H.** (1992) The Sequence Homologies of Cytochromes P-450 and Active-Site Geometries. *Journal of Computer-Aided Drug Design*. 6: 235 - 252
- Levitt M.** (1992) Accurate Modelling of Protein Conformation by Automatic Segment Matching. *Journal of Molecular Biology*. 226: 507 - 533
- Lin J.T. & Bertino J.R.** (1991) Update on Trimetrexate, a Folate Antagonist with Antineoplastic and Antiprotozoal Properties. *Cancer Investigation*. 9: 159 - 172
- Lippert T., Wokaun A., Dauth J. & Nuyken O.** (1992) NMR Studies of Hindered Rotation and Thermal Decomposition of Novel 1-Aryl-3,3-Dialkyltriazenes. *Magnetic Resonance in Chemistry*. 30: 1178 - 1185.
- Lloyd E.J. & Andrews P.R.** (1986) A Common Structural Model for Central Nervous System Drugs and Their Receptors. *Journal of Medicinal Chemistry*. 29: 453 - 462
- Loftus P.** (1990) Computer-assisted Drug Design. *Drug Discovery Technologies*. 1 - 22 Editors Clark C.R. & Moos W.H., Publishers Ellis Horwood, Chichester, U.K.
- Lunazzi G., Ceriori G., Foresti E. & Marciantelli D.** (1978) Conformational Studies by Dynamic Nuclear Magnetic Resonance. Part X. Stereodynamics and Conformations of Hindered Triazenes. *Journal of the Chemical Society, Perkin Transactions 2*. 686 - 691
- Lundgren B., Kovacs J.A., Nelson N.N., Stock F., Martinez A. & Gill V.J.** (1992) *Pneumocystis carinii* and Specific Fungi have a Common Epitope Identified by a Monoclonal Antibody. *Journal of Clinical Microbiology*. 30: 391 - 395
- MacKay D., McIntyre D.D. & Taylor N.J.** (1982) Structure of 1,6-Bis(*para*-Chlorophenyl)-3,4-Diacetyl-1,5-hexazadiene - A Compound with a Highly Electrophilic *N*-Acetyl Group. *Journal of Organic Chemistry*. 47: 532 - 535
- Mackie T.J. & MacCartney J.E.** (1953) Handbook of Practical Bacteriology: A guide to Bacteriology Laboratory Work. Publishers E.&S. Livingstone Ltd., Edinburgh.

- Martin M.A., Cox P.H., Beck K., Styer C.M. & Beall G.N.** (1992) A Comparison of the Effectiveness of Three Regimens in the Prevention of *Pneumocystis carinii* Pneumonia in Human Immunodeficiency Virus-Infected Patients. *Annals of Internal Medicine*. 152: 523 - 528
- Masters J.N. & Attardi G.** (1983) Studies of Amino Acid Sequence in Dihydrofolate Reductase from Human Methotrexate Resistant Cell Line KB/6B - Structural and Kinetic Comparison with Mouse L1210 Enzyme. *Gene*. 21: 59 - 63
- Masur H.** (1993) Drug Therapy - Prevention and Treatment of *Pneumocystis carinii* Pneumonia. *New England Journal of Medicine*. 328: 1136 - 1146
- Matthews D.A., Alden R.A., Freer S.T., Xuong N.-H. & Kraut J.** (1979) Dihydrofolate Reductase from *Lactobacillus casei*. Stereochemistry of NADPH Binding. *Journal of Biological Chemistry*. 254: 4144 - 4151
- Matthews D.A., Bolin J.T., Burridge J.M., Filman D.J., Volz K.W., Kaufman B.T., Beddell C.R., Champness J.N., Stammers D.K. & Kraut J.** (1985) Refined Crystal Structures of *Escherichia coli* and Chicken Liver Dihydrofolate Reductase Containing Bound Trimethoprim. *Journal of Biological Chemistry*. 260: 381 - 390
- McLeod G.X. & Hammer S.M.** (1992) Zidovudine: Five Years Later. *Annals of Internal Medicine*. 117: 487 - 501
- Medina I., Mills J., Leong G., Hopwell P.C., Lee B., Modin G., Benowitz N. & Wofsy C.B.** (1990) Oral Therapy for *Pneumocystis carinii* Pneumonia in the Acquired Immunodeficiency Syndrome. A Controlled Trial of Trimethoprim-Sulphamethoxazole Versus Trimethoprim-Dapsone. *New England Journal of Medicine*. 323: 776 - 782
- Medina R. & Muller K.** (1990) Molecular Modeling Studies on the Urease Active Site and the Enzyme-Catalysed Urea Hydrolysis. *Journal of Computer-Aided Molecular Design*. 4: 355 - 367
- Mellera P.W., Davide J.P., Hession C.A. & Scotto K.W.** (1984) Phenotypic Expression in *Escherichia coli* and Nucleotide Sequence of Two Chinese Hamster Lung Cell cDNAs Encoding Different Dihydrofolate Reductases. *Molecular and Cellular Biology*. 4: 38 - 48
- Meyer A.M. & Richards W.G.** (1991) Similarity of Molecular Shape. *Journal of Computer-Aided Drug Design*. 5: 426 - 439
- Millard P.S. & van der Horst C.** (1991) Reversible Diabetes Mellitus after Intravenous Pentamidine. *American Journal of Medicine*. 91: 442
- Montgomery A.B., Debs R.J., Luce J.M., Corkery K.J., Turner J., Brunnette E.N., Lin E.T. & Hopewell P.C.** (1987) Aerosolized Pentamidine as a Sole Therapy for *Pneumocystis carinii* Pneumonia in Patients with Acquired Immunodeficiency Syndrome. *Lancet*. ii: 480 - 482
- Moon J.B. & Howe W.J.** (1991) Computer Design of Bioactive Molecules - A Method for Receptor-Based *De Novo* Ligand Design. *Protein: Structure, Function and Genetics*. 11: 314 - 328

Morley S.B., Abraham R.J., Haworth I.S., Jackson D.E., Saunders M.R. & Vinter J.G. (1991) COSMIC(90): An Improved Molecular Mechanics Treatment of Hydrocarbons and Conjugated Systems. *Journal of Computer-Aided Molecular Design*. 5: 475 - 504.

Morrison J. & Boyd J. (1992) Organic Chemistry. Published by Prentice Hall, Englewood, Cliffs, New Jersey.

Motherwell W.D.S. & Clegg W. (1978) PLUTO78. Program for Plotting Molecular and Crystal Structures, University of Cambridge.

Murray J.F., Felton C.P., Garay J.M., Gottlieb M.S., Hopewell P.C., Stover D.E. & Terstein A.S. (1984) Pulmonary Complications of Acquired Immunodeficiency Syndrome. A Report of the National Heart Lung and Blood Institute Workshop. *New England Journal of Medicine*. 310: 1682 - 1688

Needleman S.B. & Wunsch C.D. (1970) A General Method Applicable to the Search for Similarities in the Amino Acid Sequence of Two Proteins. *Journal of Molecular Biology*. 48: 443 - 453

Oefner C., D'Arcy A., & Winkler F.K. (1988) The Crystal Structure of Human DHFR Complexed with Folate. *European Journal of Biochemistry*. 174: 377 - 385

Orozco M. & Luque F.J. (1990) On the Use of AM1 and MNDO Wave Functions to Compute Accurate Electrostatic Charges. *Journal of Computational Chemistry*. 11: 909 - 923

Ozment J.L., Schmiedekamp A.M., Schultz-Merkel L.A., Smith R.H., jnr, Michejda C.J. (1991) Theoretical Analysis of Acetyltriene and the Mechanistic Implications of Its Reaction with Acid. *Journal of the American Chemical Society*. 113: 397 - 405.

Peterson D.L., Gleisner J.M. & Blakely R.L. (1975) The Structure of the Mutant Dihydrofolate Reductase from *Streptococcus faecium*. *Journal of Biological Chemistry*. 250: 4945 - 4954

Petukhov M.G. & Mazur A.K. (1992) Modeling of Possible Methods of Substrate Binding in the Active-site of α -Amylase A from *Aspergillus oryzae*. *Molecular Biology*. 26: 217 - 223

Piela L., Nemethy G. & Scheraga H.A. (1987) Conformational Constraints of Amino Acid Side Chains in α -Helices. *Biopolymers*. 26: 1273 - 1286

Ponder J.W. & Richards F.M. (1987) Tertiary Templates for Proteins. Use of Packing Criteria in the Enumeration of Allowed Sequences for Different Structural Classes. *Journal of Molecular Biology*. 193: 775 - 791

Ramos M.N. & Pereira S.R. (1986) Conformation, Electronic Structure and Biological Activity of Antitumour Triazines. *Journal of the Chemical Society, Perkin Transactions 2*. : 131 - 133

Randall A.J. & Schwalbe C.H. (1984) The Crystal Structure of 3-Methyl-1-*para*-Tolyltriene. *Journal of the Chemical Society, Perkin Transactions 2*. : 251 - 251

Richards W.G. (1992) Computer-Aided Drug Discovery. *Proceedings of the Royal Society of Edinburgh*. 99b(1/2): 105 - 111

Richards W.G. (1990) Quantum Mechanics in Drug Design. *Computer-Aided Molecular Design*. 43 - 51. Editor Richards W.G. Publishers IBC Technical Services Ltd., London.

Rosowsky A., Freisheim J.H., Hynes J.B., Queener S.F., Bartlett M., Smith J.W., Lazarus H. & Modest E.J. (1989) Tricyclic 2,4-Diaminopyrimidines with Broad Antifolate Activity and the Ability to Inhibit *Pneumocystis carinii* Growth in Cultured Lung Fibroblasts in the Presence of Leucovorin. *Biochemical Pharmacology*. 38: 2677 - 2684

Roth B., Bliss E. & Beddell C.R. (1983) Inhibitors of Dihydrofolate Reductase. *Molecular Aspects of Anti-Cancer Drug Action*. Editors Neidle S. and Waring M.J. Published by Jon Wiley and Sons Ltd.

Roth B. & Stammers D.K. (1992) Drug Interactions with Target Enzymes. *The Design of Drugs to Macromolecular Targets*. Chapter 4: 85 - 118. Edited by Beddell C.R., Published by John Wiley & Sons Ltd.

Rouch D.A., Messroiti L.J., Loo L.S.L., Jackson C.A. & Skurray R.A. (1989) Trimethoprim Resistance Transposon TN4003 from *Staphylococcus aureus* Encodes Genes for a Dihydrofolate Reductase and Thymidylate Synthase Flanked by 3 Copies of 1s257. *Molecular Microbiology*. 3: 161 - 175

Rouzer C.A., Thompson E.J., Skinner T.L., Heaver P.A., Bartolini W.P., Mitchell K., Kurz E., Smith R.H., jr & Michejda C.J. (1993) An Unexpected Pathway for the Metabolic Degradation of 1,3-Dialkyl-3-acyltriazines. *Biochemical Pharmacology*. 46: 165 - 173

Ruf B., Rohde I & Pohle H.D. (1991) Efficacy of Clindamycin/Primaquine Versus Trimethoprim/Sulphamethoxazole in the Primary Treatment of *Pneumocystis carinii* Pneumonia. *European Journal of Clinical Microbiology and Infectious Diseases*. 10: 207 - 210

Ruf B., Schurmann D., Bergmann F., Schulermane W., Grunewald T., Gottschalk H.J., Witt H. & Pohle H.D. (1993) Efficacy of Pyrimethamine and Sulfadoxine on the Prevention of Toxoplasmic Encephalitis and *Pneumocystis carinii* Pneumonia in HIV Infected Patients. *European Journal of Clinical Microbiology and Infectious Diseases*. 5: 325 - 329

Ruiz J., Lopez M., Mila J., Lozoya E., Lozano J.J. & Pouplana R. (1993) QSAR and Conformational Analysis of the Antiinflammatory Agent Amfenac and Analogues. *Journal of Computer-Aided Molecular Design*. 7: 183 - 198

Ruskin J. & La Riviere M. (1991) Low-Dose Cotrimoxazole for the Prevention of *Pneumocystis carinii* Pneumonia in Human Immunodeficiency Virus Diseases. *Lancet*. 337: 468 - 471

Safrin S., Sattler F.R., Lee B.L., Young T., Bill R., Boylan C.T. & Mills J. (1991) Dapsone as a Single Agent is Suboptimal Therapy for *Pneumocystis carinii* Pneumonia. *Journal of Acquired Immunodeficiency Syndrome*. 4: 244 - 249

Salim M., Zhang Y., Sloan D. & Meshnick S. (1989) Inhibition of *Pneumocystis carinii* Dihydropteroate Synthetase by Sulfa Drugs. *Antimicrobial Agents and Chemotherapy*. 34: 1075 - 1078

Sandorfy C. (1970) *The Chemistry of the Carbon - Nitrogen Double Bond*. Chapter 1, page 3, Edited by S. Patai. Published by Interscience, John Wiley and Sons Ltd.

Santiago-Delpin E.A., Mora E., Gonzalez Z.A., Morales-Otero L.A. & Bermudez R. (1988) Factors in an Outbreak of *Pneumocystis carinii* in a Transplant Unit. *Transplant Proceedings*. 20: 462 - 465

Sali A. & Blundell T.L. (1990) Definition of Topological Equivalence in Protein Structures - Procedure Involving Comparison of Properties and Relationships. *Journal of Molecular Biology*. 212: 403 - 428

Sali A., Overington J.P., Johnson M.S. & Blundell T.L. (1990) From Comparison of Protein Sequences and Structures to Protein Modeling and Design. *Trends in Biochemical Sciences*. 15: 235 - 241

Sansom C.E., Schwalbe C.H., Lambert P.A., Griffin R.J., Stevens M.F.G. (1989) Structural Studies on Bio-active Compounds. Part 15. Structure-activity Relationships for Pyrimethamine and a Series of Diaminopyrimidine Analogues Versus Bacterial Dihydrofolate Reductase. *Biochimica et Biophysica Acta*. 995: 21 - 27

Saqi M.A.S. & Sternberg M.J.E. (1991) A Simple Method to Generate Non-trivial Alternate Alignments of Protein Sequences. *Journal of Molecular Biology*. 219: 727 - 732

Sattler F.R., Allegra C.J., Verdegem T.D., Bisher A., Tuazon C.U., Hughlett C., Ogata-Arakaki D., Feinberg J., Shelhamer J., Lane H.C., Davis R., Boylen C.T., Leedom J.M. & Masur H. (1990) Trimetrexate-Leucovorin Dosage Evaluation Study for the Treatment of *Pneumocystis carinii* Pneumonia. *Journal of Infectious Diseases*. 161: 91 - 96

Sawyer T.K., Staples D.J., Liu L., Tomasselli A.G., Hui O.G., O'Connell K., Schostarez H., Hester J.B., Moon J., Howe W.J., Smith C.W., Decamp D.L., Craik C.S., Dunn B.M., Lowther W.T., Harris J., Poorman R.A., Wlodawer A., Jaskolski M. & Heinrikson R.L. (1992) HIV Protease (HIV PR) Inhibitor Structure-Activity-Selectivity, and Active Site Molecular Modeling of High Affinity Leu[CH(OH)CH₂]Val Modified Viral and Nonviral Substrate Analogs. *International Journal of Peptide and Protein Research*. 40: 274 - 281

Schmatz D.M., Abruzzo G., Powles M.A., McFadden D.C., Balkovec J.M., Black R.M., Nollstadt K. & Bartizal K. (1992) Pneumocandins from *Zalerion arboricola* IV. Biological Evaluation of Natural and Semisynthetic Pneumocandins for Activity Against *Pneumocystis carinii* and *Candida* Species. *Journal of Antibiotics*. 45: 1886 - 1891

Schmiedekamp A., Smith R.H. & Michejda C.J. (1988) *Ab initio* Studies of Triazenes in Relation to Experimental Findings. *Journal of Organic Chemistry*. 53: 3433 - 3436

Schusterman A.J., Johnson A.S. & Hansch C. (1989) Correlation of Mutagenicity of 1,1-Dimethyl-3-(X-Phenyl)-triazenes with Molecular Orbital Energies and Hydrophobicity. *International Journal of Quantum Chemistry*. 36: 19 - 33

Schwalbe C.H. & Cody V. (1983) *Chemistry and Biology of Pteridines*. 511 - 515. Editor Blair J.A.. Published by Walter de Gruyter, Berlin.

Schwalbe C.H. & Williams G.J.B. (1987) The Structure of 2,4,6-Triaminopyrimidine. *Acta Crystallographica*. C43: 1097 - 1100

Schwalbe C.H. & Williams G.J.B. (1982) Triamterene. *Acta Crystallographica*. B38: 1840 - 1843

Sheldrick G. (1976) SHELX. A Program for Crystal Structure Determination. University of Cambridge.

Shotkoski F.A. & Fallon A.M. (1991) An Amplified Insect Dihydrofolate Reductase Gene Contains a Single Intron. *European Journal of Biological Chemistry*. 201: 157 - 160

Sibanda B.L., Blundell T.L. & Thornton J.M. (1989) Conformation of β -Hairpins in Protein Structures. A Systematic Classification of Modeling by Homology, Electron Density Fitting and Protein Engineering. *Journal of Molecular Biology*. 206: 759 - 777

Simmonds R.J., Katenga R.B. & Schwalbe C.H. (1988) The Geometry of N-Hydroxymethyl Compounds. 2. Crystal Structures of 1-(4-Carboxyphenyl)-3-Hydroxymethyl-3-Methyltriazene, N-Hydroxymethylpentamethylmelamine and N-Hydroxymethylbenzamide. *Anti-Cancer Drug Design*. 3: 91 - 101

Simonsen C.C. & Levinson A.D. (1983) Isolation and Expression of an Altered Mouse Dihydrofolate Reductase cDNA. *Proceedings of the National Academy of Sciences of the U.S.A.* 80: 2495 - 2499

Singh J. & Thornton J.M. (1990) An Automated Method for the Analysis of the Preferred Packing Arrangements Between Protein Groups. *Journal of Molecular Biology*. 211: 595 - 615

Slattery J.T. & Unadkat J.D. (1991) Adverse Reactions to Cotrimoxazole in HIV Patients. *Lancet*. 338: 1216

Smith D.A., Ulmer C.W. II, & Gilbert M.J. (1992) Structural Studies of Aromatic Amines and the DNA Intercalating Compounds *m*-AMSA and *o*-AMSA: Comparison of MNDO, AM1 and PM3 to Experimental and *Ab Initio* Results. *Journal of Computational Chemistry*. 13: 640 - 650

Smith D. & Gazzard B. (1991) Treatment and Prophylaxis of *Pneumocystis carinii* Pneumonia in AIDS Patients. *Drugs*. 42: 628 - 639

Smith S.L., Patrick D., Stone D., Phillips A.W. & Burchall J.J. R Plasmid Dihydrofolate Reductase with Subunit Structure (1979) *Journal of Biological Chemistry*. 254: 6222 - 6225

So S.S. & Richards W.G. (1992) Applications of Neural Networks - Quantitative Structure-Activity Relationships of the Derivatives of 2,4-Diamino-5-(Substituted-benzyl)pyrimidines as DHFR Inhibitors. *Journal of Medicinal Chemistry*. 35: 3201 - 3207

Spassova M.K. & Golovinsky E.V. (1985) Pharmacobiochemistry of Arylalkyltriazenes and their Application in Cancer Chemotherapy. *Pharmacology and Therapeutics*. 27: 333 - 352

Stammers D.K., Champness J.N., Beddell C.R., Dann J.G., Eliopoulos E.E., Geddes A.J. & North A.C.T. (1987) The Structure of Mouse L1210 Dihydrofolate Reductase-Drug Complexes and the Construction of a Model of Human Enzyme. *FEBS Letters*. 218: 178 -184

Stammers D.K., Delves C., Ballantine S., Jones E.V., Stuart D.I., Achari A., Bryant P.K. & Champness J.M. (1993) Preliminary Crystallographic Data for *Pneumocystis carinii* Dihydrofolate Reductase. *Journal of Molecular Biology*. 230: 679 - 680

Stewart J.J.P. (1992) Semi-empirical Molecular Orbital Methods. *Reviews in Computational Chemistry*. 45 - 81 Editors Lipkowitz K.B. & Boyd D.B. Publishers VCH Publishers (UK) Ltd, Cambridge.

Stewart J.J.P. (1990 a) MOPAC: A Semi-empirical Molecular Orbital Program. *Journal of Computer-Aided Molecular Design*. 4: 1 - 105.

Stewart J.J.P. (1990 b) Reply to "Comments on a Comparison of AM1 with the Recently Developed PM3 Method." *Journal of Computational Chemistry*. 11: 543 - 544

Stewart J.J.P. (1989) Optimisation of Parameters for Semi-empirical Methods. I. Method. *Journal of Computational Chemistry*. 10: 209 - 220

Sutton P.A. & Cody V. (1986) Crystal Structures of Methotrexate and its Conformational Comparison with Related Structures. *Proceedings of the American Association of Cancer Research*. 27

Sutcliffe M.J., Hayes F.R.F & Blundell T.L. (1987) Knowledge Based Modeling of Homologous Proteins. *Protein Engineering*. 1: 385 - 392

Taylor R. & Kennard O. (1984) Hydrogen Bond Geometry in Organic Crystals. *Accounts in Chemical Research*. 17: 320 - 326

Taylor J.E. & Metzler D.E. (1990) Modeling Inhibitors in the Active Site of Aspartate Aminotransferase. *Annals of the New York Academy of Sciences*. 585: 58 - 67

Taylor J.W. & Bowman B.H. (1993) *Pneumocystis carinii* and the Ustomycetous Red Yeast Fungi. *Molecular Microbiology*. 8: 425 - 426.

Theil W. (1988) Semiempirical Methods: Current Status and Perspectives. *Tetrahedron*. 44: 7393 - 7408

Tidwell R.R., Jones S.K., Dykstra C.C., Gorton L. & Hall J.E. (1991) Treatment of Experimental *Pneumocystis carinii* Pneumonia with 1,3-di-(4-Imidazolino-2-Methoxyphenoxy)propane Lactate. *Journal of Protozoology*. 38: 148S - 150S

Tidwell R.R., Jones S.K., Geratz J.D., Ohemeng K.A., Cory M. & Hall J.E. (1990) Analogues of 1,5-bis-(4-Amidinophenoxy)pentane (Pentamidine) in the Treatment of Experimental *Pneumocystis carinii* Pneumonia. *Journal of Medicinal Chemistry*. 33: 1252 - 1257

Toma E. (1991) Clindamycin/Primaquine for the Treatment of *Pneumocystis carinii* Pneumonia in AIDS. *European Journal of Clinical Microbiology and Infectious Diseases*. 10: 210 - 213

Trimble J.J., Murthy S.C.S., Bakker A., Grassmann R. & Desrosiers R.C. (1988) A Gene for Dihydrofolate Reductase in a Herpesvirus. *Science*. 239: 1145 - 1147

van Alsenoy C. (1988) *Ab initio* Calculations on Large Molecules: the Multiplicative Integral Approximation. *Journal of Computational Chemistry*. 9: 620 - 626

van der Ven A.J., Koopmans P.P., Vree T.B. & van der Meer J.W. (1991) Adverse Reactions to Cotrimoxazole in HIV Infections. *Lancet*. 338: 431 - 433

Vanek J. & Jirovec O. (1952) Parasitic Pneumonia; Interstitial Plasma Cell Pneumonia of Premature Infants Caused by *Pneumocystis carinii*. *Zentralbl. Bakteriologie*. 158: 120

Varva J. & Kucera K. (1970) *Pneumocystis carinii* Delanoe, its Ultrastructure and Ultrastructure Affinities. *Journal of Protozoology*. 17: 463 - 469

Vingron M. & Sibbald P.R. (1993) Weighting in Sequence Space: A Comparison of Methods in Terms of Generalised Sequences. *Proceedings of the National Academy of Sciences of the U.S.A.* 90: 8777 - 8781

Volberding P.A. (1991) Recent Advances in the Early Management of HIV Disease. *Journal of General Internal Medicine*. 6: S7 - S12

Wakefield A.E., Hopkin J.M., Bridge P.D. & Hawksworth D.L. (1993) *Pneumocystis carinii* and the Ustomycetous Red Yeast Fungi. *Molecular Microbiology*. 8: 426 - 427.

Wakefield A.E., Peters S.E., Banerji S., Bridge P.D., Hall G.S., Hawksworth D.L., Guiver L.A., Allen G.A. & Hopkin J.M. (1992) *Pneumocystis carinii* Shows DNA Homology with Ustomycetous Red Yeast Fungi. *Molecular Microbiology*. 6: 1903 - 1911

Wakefield A.E., Pixeley F.J., Banerji S., Sinclair K., Miller R.F., Moxon E.R. & Hopkin J.M. (1990) Detection of *Pneumocystis carinii* with DNA Amplification. *Lancet*. 336: 451 - 453

Walzer P.D., Foy J., Steele P. & White M. (1992) Treatment of Experimental Pneumocystis: Review of 7 Years of Experience and Development of a New System for Classifying Antimicrobial Drugs. *Antimicrobial Agents and Chemotherapy*. 36: 1943 - 1950

Walzer P.D., Schelle V., Armstrong D. & Rosen P.P. (1977) Nude Mouse: A New Experimental Model for *Pneumocystis carinii* Infection. *Science*. 197: 177 - 179

Watanabe J.I., Hori H., Tanabe K., Nakamura Y. (1989) 5s Ribosomal RNA Sequence of *Pneumocystis carinii* and its Phylogenetic Association with Rhizopoda/Myxomycota/Zygomycota Group. *Molecular and Biochemical Parasitology*. 32: 163 - 173

Weber I.T. (1990) Evaluation of Homology Modeling of HIV Protease. *Proteins: Structure, Functions and Genetics*. 7: 172 - 184

- Welsh W.J.** (1990) AM1 Molecular Orbital Studies on the Structures, Conformations, Protonation Energies and Electronic Properties of Triazine Dihydrofolate Reductase Inhibitors. *Journal of Computational Chemistry*. 11: 644 - 653
- White D.N.J. & Kitson D.H.** (1986) Computational Conformational Analysis of Cyclohexaglycyl. *Journal of Molecular Graphics*. 4: 112- 116
- Worth D.F. & Werbel L.M.** (1984) Repository Preparations. *Anti-malarial Drugs II*. Chapter 7: 237 - 252. Editors Peters W. & Richards W.H.G. Published by Springer-Verlag, Berlin, Germany
- Wrigglesworth J.M.** (1991) The Active Site Structure of Cytochrome Oxidase: Modeling by "Catalytic Homology." *Biochemical Society Transactions*. 19: 258S
- Yeh J.C., Borchardt R.T. & Vedani A.** (1991) A Molecular Model for the Active Site of *S*-Adenosyl-L-Homocysteine Hydrolase. *Journal of Computer-Aided Molecular Design*. 5: 213 - 234
- Yomawong M.F., Fonzi W.A. & Sypherd P.S.** (1992) Fungus-specific Translation Elongation Factor-III Present in *Pneumocystis carinii*. *Infection and Immunity*. 60: 4140 - 4145
- Yoneda K. & Walzer P.D.** (1983) Attachment of *Pneumocystis carinii* to Type I Alveolar Cells Studied by Freeze-Fracture Electron Microscopy. *Infection and Immunity*. 40: 812 - 815
- Yoneda K. & Walzer P.D.** (1984) The Effect of Corticosteroid Treatment on the Cell Surface of the Rat Pulmonary Alveolus - Relevance to the Host-Parasite Relationship in *Pneumocystis carinii* Infection. *British Journal of Experimental Pathology*. 65: 347 - 354
- Yoshida Y.** (1989) Ultrastructural Studies of *Pneumocystis carinii*. *Journal of Protozoology*. 36: 53 - 60
- Zackrison L.H. & Tsou E.** (1991) *Pneumocystis carinii*: A Deadly Opportunist. *American Family Physician*. 44: 528 - 541
- Zheung Q.A., Rosefield P., Vadja S. & Delisi C.** (1993) Determination of Protein Loop Conformation Using Scaling-Relaxation Techniques. *Protein Science*. 2: 1242 - 1248
- Zolg J.W., Plitt J.R., Chen G.-X.X. & Palmer S.** (1989) Point Mutations in the Dihydrofolate Reductase - Thymidylate Synthase Gene as the Molecular Basis for Pyrimethamine Resistance in *Plasmodium falciparum*. *Molecular and Biochemical Parasitology*. 36: 253 - 262
- Zuker M.** (1991) Suboptimal Sequence Alignments in Molecular Biology. Alignment with Error Analysis. *Journal of Molecular Biology*. 221: 403 - 420
- Zusman T., Rosenshine J.S., Boehm G., Jaenicke R., Leskiw B. & Mevavech M.** (1989) Dihydrofolate Reductase of the Extremely Halophilic archaeobacterium *Halobacterium volcanii*. The Enzyme and its Coding Gene. *Journal of Biological Chemistry*. 264: 18878 - 18883

Appendix Structure Factor Data for AM90021

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR AM90021

H	K	L	10FO	10FC	H	K	L	10FO	10FC	H	K	L	10FO	10FC	H	K	L	10FO	10FC	H	K	L	10FO	10FC
1	0	0	785	845	-5	4	0	44	49	-4	8	0	89	-89	3	-8	1	190	-188	-3	-4	1	74	73
2	0	0	325	-335	-4	4	0	280	274	-2	8	0	80	-73	4	-8	1	122	-121	-2	-4	1	276	259
4	0	0	111	78	-3	4	0	247	243	-1	8	0	83	-81	5	-8	1	77	75	-1	-4	1	100	-111
5	0	0	173	172	-2	4	0	422	382	0	8	0	19	26	6	-8	1	52	49	0	-4	1	419	-418
6	0	0	30	25	-1	4	0	145	128	1	8	0	36	-34	8	-8	1	52	-53	1	-4	1	43	-56
7	0	0	70	-59	0	4	0	211	-200	2	8	0	123	-117	-4	-7	1	86	89	2	-4	1	40	-33
-7	1	0	60	-64	1	4	0	160	146	3	8	0	130	-132	-3	-7	1	105	109	3	-4	1	500	455
-6	1	0	65	-62	2	4	0	36	33	-7	9	0	34	39	-2	-7	1	34	34	4	-4	1	103	100
-4	1	0	77	-75	3	4	0	172	-174	-6	9	0	41	31	-1	-7	1	26	-23	5	-4	1	178	-186
-3	1	0	592	-579	4	4	0	177	-185	-5	9	0	107	-107	0	-7	1	49	45	6	-4	1	113	-96
-2	1	0	673	-642	7	4	0	53	-51	-4	9	0	85	-95	1	-7	1	28	18	7	-4	1	101	93
1	1	0	408	358	-9	5	0	28	20	-3	9	0	42	47	2	-7	1	97	100	-7	-3	1	30	20
2	1	0	295	-299	-8	5	0	44	45	-2	9	0	58	61	3	-7	1	59	61	-6	-3	1	106	104
3	1	0	188	197	-6	5	0	106	-103	-1	9	0	41	46	4	-7	1	209	-211	-5	-3	1	22	10
4	1	0	145	137	-5	5	0	119	-103	0	9	0	34	41	6	-7	1	31	20	-4	-3	1	144	-146
5	1	0	62	68	-4	5	0	177	-168	1	9	0	39	-31	-5	-6	1	37	-42	-3	-3	1	106	96
6	1	0	35	-22	-3	5	0	93	80	-6	10	0	51	48	-4	-6	1	165	-159	-2	-3	1	109	107
7	1	0	71	-65	-1	5	0	264	267	-5	10	0	52	-53	-3	-6	1	131	-121	-1	-3	1	186	220
-8	2	0	43	50	0	5	0	107	103	-4	10	0	105	-106	-2	-6	1	46	46	0	-3	1	156	164
-6	2	0	81	-79	1	5	0	78	76	-3	10	0	23	21	-1	-6	1	71	72	1	-3	1	110	-100
-5	2	0	87	-82	2	5	0	29	23	-2	10	0	124	119	0	-6	1	166	-163	2	-3	1	42	37
-4	2	0	365	-321	3	5	0	192	-182	-1	10	0	64	53	1	-6	1	154	-146	3	-3	1	87	-75
-3	2	0	77	67	5	5	0	23	33	-6	11	0	41	35	2	-6	1	131	-118	4	-3	1	47	-53
-2	2	0	339	-316	-8	6	0	67	69	-4	11	0	58	-58	3	-6	1	49	52	5	-3	1	93	-70
-1	2	0	429	-441	-7	6	0	24	28	1	-11	1	51	53	4	-6	1	39	-30	6	-3	1	86	85
0	2	0	76	-65	-6	6	0	120	-112	3	-11	1	63	-58	5	-6	1	125	-117	8	-3	1	46	56
1	2	0	193	-186	-5	6	0	25	-28	4	-11	1	24	34	6	-6	1	85	84	-8	-2	1	61	-58
2	2	0	226	-199	-4	6	0	187	178	5	-11	1	40	42	7	-6	1	37	-59	-7	-2	1	41	47
3	2	0	265	-261	-2	6	0	112	126	-1	-10	1	71	-75	8	-6	1	69	-72	-6	-2	1	30	16
4	2	0	204	198	-1	6	0	117	-107	0	-10	1	36	38	-5	-5	1	87	-91	-5	-2	1	97	-88
5	2	0	139	143	0	6	0	48	53	1	-10	1	83	81	-4	-5	1	125	-110	-4	-2	1	180	-186
6	2	0	43	56	1	6	0	176	175	2	-10	1	34	-28	-3	-5	1	90	91	-2	-2	1	168	181
7	2	0	51	-44	2	6	0	158	154	3	-10	1	72	-70	-2	-5	1	142	128	-1	-2	1	670	-626
-6	3	0	150	-140	3	6	0	126	129	5	-10	1	58	56	-1	-5	1	143	-142	0	-2	1	86	91
-5	3	0	84	86	4	6	0	23	24	6	-10	1	54	46	0	-5	1	26	18	1	-2	1	93	-68
-4	3	0	324	328	-8	7	0	52	48	7	-10	1	40	44	1	-5	1	87	86	2	-2	1	314	290
-3	3	0	354	-345	-7	7	0	63	68	-2	-9	1	41	-42	2	-5	1	449	427	3	-2	1	103	92
-2	3	0	864	-820	-5	7	0	179	-172	0	-9	1	30	-31	3	-5	1	318	287	4	-2	1	150	145
-1	3	0	521	-530	-4	7	0	49	-44	1	-9	1	48	50	4	-5	1	84	-88	6	-2	1	90	-91
0	3	0	28	-17	-3	7	0	108	116	2	-9	1	46	46	5	-5	1	172	-163	7	-2	1	107	98
1	3	0	186	195	-1	7	0	21	-15	3	-9	1	65	-64	6	-5	1	30	28	9	-2	1	52	-53
3	3	0	329	-333	0	7	0	149	-151	5	-9	1	91	88	7	-5	1	79	81	-7	-1	1	60	57
4	3	0	124	-126	1	7	0	32	30	7	-9	1	58	-62	8	-5	1	61	-74	-6	-1	1	126	118
5	3	0	131	124	2	7	0	121	121	-3	-8	1	88	95	9	-5	1	65	-59	-5	-1	1	105	-108
7	3	0	35	33	3	7	0	48	52	-2	-8	1	36	39	-7	-4	1	78	74	-4	-1	1	255	-235
-8	4	0	60	68	-7	8	0	63	63	-1	-8	1	75	-79	-6	-4	1	32	-29	-3	-1	1	93	-85
-7	4	0	51	-43	-6	8	0	48	52	1	-8	1	124	120	-5	-4	1	68	-71	-2	-1	1	628	592
-6	4	0	24	-36	-5	8	0	41	-40	2	-8	1	87	91	-4	-4	1	37	50	-1	-1	1	816	807
1	-1	1	54	-93	5	2	1	47	36	-1	6	1	187	-200	7	-9	2	52	-52	-3	-4	2	119	122
2	-1	1	292	-273	6	2	1	30	-4	-1	6	1	73	-80	8	-9	2	34	-40	-2	-4	2	112	100
3	-1	1	495	479	8	2	1	34	44	5	6	1	45	55	-3	-8	2	46	-46	-1	-4	2	154	-168
4	-1	1	117	93	-8	3	1	29	22	-9	7	1	31	24	-2	-8	2	98	-101	0	-4	2	100	102
5	-1	1	173	-153	-7	3	1	98	-101	-8	7	1	24	18	-1	-8	2	38	36	1	-4	2	259	-256
6	-1	1	54	-63	-6	3	1	76	74	-6	7	1	57	-46	0	-8	2	86	86	2	-4	2	18	-26
-8	0	1	51	-52	-5	3	1	310	298	-4	7	1	89	78	1	-8	2	40	42	3	-4	2	28	17
-7	0	1	50	58	-4	3	1	310	-287	-1	7	1	32	-37	2	-8	2	100	-103	4	-4	2	223	-214
-6	0	1	80	83	-3	3	1	82	77	3	-8	2	39	-33	5	-4	2	39	-33	7	-4	2	119	-131
-5	0	1	40	-43	-2	3	1	329	297	2	7	1	213	-207	4	-8	2	48	44	8	-4	2	52	-51
-4	0	1	446	-430	-1	3	1	173	166	3	7	1	86	-93	6	-8	2	30	45	9	-4	2	47	46
-3	0	1	700	-641	2	3	1	154	149	-4	7	1	27	35	-4	-7	2	45	-43	-6	-3	2	118	-114
-2	0	1	845	-815	3	3	1	172	189	-7	8	1	52	-45	-3	-7	2	69	-75	-5	-3	2	165	-160
-1	0	1	137	135	4	3	1	72	72	-4	8	1	85	86	0	-7	2	52	-55	-4	-3	2	76	74
1	0	1	258	-201	5	3	1	150	-133	-2	8	1	35	-35	1	-7	2	66	-64	-3	-3	2	30	35
2	0	1	959	934	6	3	1	53	-57	-1	8	1	83	88	1	-7	2	17	-9	-2	-3	2	163	-163
3	0	1	408	-409	9	4	1	36	47	0	8	1	154	161	2	-7	2	17	-9	-1	-3	2	139	-126
4	0	1	247	-229	-7	4	1	76	-68	1	8	1	147	146	3	-7	2	50	-57	0	-3	2	205	166
5	0	1	154	-155	-6	4	1	204	-208	2	8	1	90	89	4	-7	2	190	191	1	-3	2	749	785
6	0	1	82	-81	-5	4	1	284	-261	-6	9	1	100	-91	6	-7	2	117	-131	1	-3	2	260	235
7	0	1	47	32	-4	4	1	152	-158	-5	9	1	55	-54	9	-7	2	46	37	2	-3	2	285	-297
8	0	1	41	37	-3	4	1	215	-195	-4	9	1	33	26	-4	-6	2	136	138	3	-3	2	220	-235
-8	1	1	64	-68	-2	4	1	202	176	-2	9	1	62	-69	-3	-6	2	100	93	5	-3	2	93	107
-6	1	1	77	83	-1	4	1	22																

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR AM90021

H	K	L	10FO	10FC	H	K	L	10FO	10FC	H	K	L	10FO	10FC	H	K	L	10FO	10FC	H	K	L	10FO	10FC
-1	2	1	1986	2041	5	5	1	28	-11	0	-9	2	44	43	6	-5	2	44	41	6	-2	2	175	174
0	2	1	504	526	-9	6	1	38	32	2	-9	2	68	-68	7	-5	2	108	-109	8	-2	2	59	-56
1	2	1	487	-484	-5	6	1	120	103	3	-9	2	69	67	9	-5	2	67	66	-7	-1	2	68	57
2	2	1	39	-39	-4	6	1	154	141	4	-9	2	85	85	-6	-4	2	28	18	-6	-1	2	25	-18
3	2	1	297	279	-3	6	1	181	167	5	-9	2	74	-75	-5	-4	2	47	-43	-5	-1	2	60	-67
4	2	1	108	121	-2	6	1	62	-71	6	-9	2	138	-128	-4	-4	2	103	-101	-4	-1	2	40	-18
-3	-1	2	301	-284	-3	2	2	254	-219	-7	6	2	92	-99	2	-10	3	49	46	1	-5	3	247	-224
-2	-1	2	402	-358	-2	2	2	172	165	-6	6	2	26	-34	3	-10	3	58	62	2	-5	3	163	-152
-1	-1	2	540	-511	-1	2	2	532	-544	-4	6	2	54	-60	4	-10	3	42	-35	3	-5	3	130	130
0	-1	2	138	165	0	2	2	609	-616	-2	6	2	36	-39	5	-10	3	126	-126	4	-5	3	176	168
1	-1	2	717	675	1	2	2	337	374	-1	6	2	89	-90	8	-10	3	27	-18	6	-5	3	55	-53
2	-1	2	569	556	2	2	2	153	-150	0	6	2	180	-191	-2	-9	3	33	38	-5	-4	3	129	131
3	-1	2	244	-213	3	2	2	319	-292	1	6	2	121	-109	1	-9	3	84	-88	-4	-4	3	149	147
4	-1	2	160	-149	4	2	2	198	-190	3	6	2	96	89	2	-9	3	36	31	-3	-4	3	113	-110
5	-1	2	153	148	5	2	2	110	-108	4	6	2	24	43	3	-9	3	83	84	-2	-4	3	80	-94
6	-1	2	80	64	6	2	2	29	-34	5	6	2	43	-40	4	-9	3	62	68	-1	-4	3	667	684
7	-1	2	64	78	7	2	2	56	50	-8	7	2	54	-55	5	-9	3	73	65	0	-4	3	136	136
-8	0	2	69	70	-8	3	2	50	-48	-6	7	2	101	103	6	-9	3	38	35	1	-4	3	269	-260
-7	0	2	38	34	-5	3	2	213	-210	-5	7	2	143	143	-2	-8	3	69	73	2	-4	3	96	-96
-6	0	2	53	58	-4	3	2	249	256	-4	7	2	133	-134	-1	-8	3	22	30	3	-4	3	225	-213
-5	0	2	105	103	-3	3	2	98	104	-3	7	2	156	-150	1	-8	3	63	62	4	-4	3	462	450
-4	0	2	243	235	-2	3	2	697	-686	-2	7	2	71	67	2	-8	3	45	36	5	-4	3	89	81
-3	0	2	115	-118	-1	3	2	87	-71	-1	7	2	134	133	3	-8	3	100	107	6	-4	3	162	-166
-2	0	2	1181	-1132	0	3	2	143	-142	0	7	2	106	102	4	-8	3	59	-68	7	-4	3	68	-72
-1	0	2	571	-621	1	3	2	134	123	1	7	2	36	-42	5	-8	3	98	-102	-7	-3	3	50	-41
1	0	2	939	944	2	3	2	238	238	-8	8	2	58	-53	7	-8	3	58	65	-6	-3	3	94	97
2	0	2	23	-28	3	3	2	46	50	-6	8	2	52	44	8	-8	3	64	69	-5	-3	3	273	266
3	0	2	111	91	4	3	2	82	-87	-5	8	2	90	93	-4	-7	3	62	-58	-4	-3	3	117	112
4	0	2	120	-118	5	3	2	66	-64	-4	8	2	26	26	-1	-7	3	111	112	-3	-3	3	332	-331
5	0	2	119	-128	-7	4	2	105	107	-3	8	2	237	-242	0	-7	3	70	74	-2	-3	3	47	-67
6	0	2	147	136	-6	4	2	180	181	-2	8	2	171	-177	1	-7	3	157	-160	-1	-3	3	362	354
7	0	2	52	52	-5	4	2	68	-71	0	8	2	88	90	2	-7	3	37	-36	0	-3	3	430	454
8	0	2	38	-36	-4	4	2	453	-440	1	8	2	43	44	3	-7	3	208	205	2	-3	3	179	-175
-9	1	2	28	-29	-3	4	2	95	-95	2	8	2	31	-32	4	-7	3	74	85	3	-3	3	263	-235
-8	1	2	54	52	-2	4	2	382	367	-7	9	2	29	-24	5	-7	3	176	-183	4	-3	3	46	50
-7	1	2	47	55	-1	4	2	307	294	-6	9	2	23	25	6	-7	3	46	-32	5	-3	3	174	180
-6	1	2	59	-66	0	4	2	134	-149	-5	9	2	85	83	7	-7	3	55	-59	6	-3	3	77	-79
-5	1	2	48	41	1	4	2	202	-194	-3	9	2	106	-102	-5	-6	3	56	53	8	-3	3	75	76
-4	1	2	196	194	2	4	2	162	156	-2	9	2	95	-96	-4	-6	3	110	-118	-8	-2	3	41	-34
-3	1	2	631	591	3	4	2	151	145	-1	9	2	22	27	-3	-6	3	239	-234	-6	-2	3	37	33
-2	1	2	862	787	4	4	2	69	64	1	9	2	92	-98	-2	-6	3	19	17	-5	-2	3	79	-85
-1	1	2	270	347	-9	5	2	31	24	2	9	2	73	-78	-1	-6	3	132	141	-4	-2	3	60	68
0	1	2	113	-176	-8	5	2	58	-56	-5	10	2	82	80	1	-6	3	241	-237	-3	-2	3	102	97
1	1	2	192	-162	-7	5	2	89	86	-3	10	2	71	-55	2	-6	3	126	-121	-2	-2	3	184	-198
2	1	2	566	545	-6	5	2	346	339	-1	10	2	28	37	3	-6	3	25	25	-1	-2	3	367	368
3	1	2	167	158	-4	5	2	131	-118	0	10	2	46	45	4	-6	3	170	159	0	-2	3	106	-110
5	1	2	128	-113	-3	5	2	170	143	-5	11	2	65	60	5	-6	3	130	129	1	-2	3	140	147
6	1	2	31	20	-2	5	2	228	210	-4	11	2	85	87	6	-6	3	47	-41	2	-2	3	177	-152
7	1	2	52	46	0	5	2	99	119	-2	11	2	25	-24	8	-6	3	49	-51	3	-2	3	186	177
-9	2	2	45	-52	1	5	2	45	44	-4	-5	3	101	-101	-4	-5	3	101	-101	4	-2	3	161	165
-7	2	2	81	87	2	5	2	88	88	4	-11	3	77	-76	-3	-5	3	136	-139	5	-2	3	19	17
-6	2	2	90	-103	3	5	2	139	127	-2	-5	3	39	-39	-3	-5	3	39	-39	6	-2	3	100	-94
-5	2	2	210	-185	4	5	2	38	-23	-1	-10	3	40	-35	-1	-5	3	78	-78	7	-2	3	38	-37
-4	2	2	105	-116	0	-10	3	54	-58	0	-5	3	380	-375	8	-2	3	103	104	8	-2	3	103	104
-8	-1	3	38	39	-8	6	2	38	-40	0	-10	3	54	-58	0	-5	3	380	-375	5	-5	4	44	-37
-7	-1	3	85	-65	-4	2	3	76	-67	-5	6	3	212	-206	8	-10	4	26	27	6	-5	4	37	-34
-6	-1	3	105	-114	-3	2	3	107	-113	-4	6	3	49	-44	0	-9	4	43	-46	7	-5	4	71	74
-5	-1	3	78	79	-2	2	3	134	-153	2	-9	4	33	40	2	-9	4	33	40	-6	-4	4	53	52
-4	-1	3	131	-117	-1	2	3	371	378	-3	6	3	242	232	3	-9	4	87	-89	-5	-4	4	115	-120
-3	-1	3	263	-269	0	2	3	532	559	0	6	3	129	-132	4	-9	4	40	-38	-4	-4	4	316	-309
-2	-1	3	626	-610	1	2	3	146	-146	1	6	3	50	-53	6	-9	4	74	66	-3	-4	4	46	-50
-1	-1	3	771	-796	2	2	3	80	-103	4	6	3	29	-24	8	-9	4	62	-71	-3	-4	4	267	269
0	-1	3	621	609	3	2	3	39	-38	-9	7	3	44	-43	-3	-8	4	32	26	-1	-4	4	86	-86
1	-1	3	225	241	4	2	3	29	-35	-8	7	3	42	39	-2	-8	4	24	-19	-1	-4	4	382	-393
2	-1	3	355	-356	5	2	3	90	92	-7	7	3	119	124	-1	-8	4	28	-31	0	-4	4	122	116
4	-1	3	360	359	6	2	3	58	63	-6	7	3	60	66	0	-8	4	28	-31	2	-4	4	185	-173
6	-1	3	72	-59	7	2	3	56	-47	-5	7	3	154	-141	2	-8	4	149	159	3	-4	4	191	185
7	-1	3	134	-139	-8	3	3	65	67	-3	7	3	146	136	3	-8	4	28	-34	4	-4	4	168	161
8	-1	3	86	-117	-6	3	3	39	-47	-2	7	3	38	37	4	-8	4	231	-235	5	-4	4	93	-83
-9	0	3	58	51	-4	3	3	132	131	5	-8	4	38	-29	5	-8	4	38	-29	6	-4	4	83	82
-7	0	3	75	-69	-3	3	3	114	117	6	-8	4	61	54	8	-8	4	66	-75	9	-4	4	34	-35
-6	0	3	41	48	-2	3	3	632	618	-3	-7													

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR AM90021

H	K	L	10FO	10FC	H	K	L	10FO	10FC	H	K	L	10FO	10FC	H	K	L	10FO	10FC	H	K	L	10FO	10FC
3	1	3	183	-177	-4	5	3	115	119	3-11	4	72	-78	-4	-5	4	52	59	4	-2	4	21	-11	
4	1	3	137	129	-3	5	3	422	-422	-1-10	4	44	-40	-3	-5	4	96	108	5	-2	4	213	-192	
5	1	3	69	77	-2	5	3	190	-179	0-10	4	34	-39	-2	-5	4	215	214	6	-2	4	198	-211	
6	1	3	22	35	-1	5	3	30	-27	1-10	4	26	26	-1	-5	4	94	92	-8	-1	4	66	-65	
7	1	3	36	-9	0	5	3	55	-57	2-10	4	65	57	0	-5	4	104	-116	-7	-1	4	63	68	
-9	2	3	48	48	2	5	3	61	57	3-10	4	60	53	1	-5	4	293	292	-6	-1	4	353	348	
-7	2	3	26	12	3	5	3	77	-76	5-10	4	70	62	2	-5	4	247	249	-5	-1	4	221	223	
-6	2	3	75	74	-7	6	3	115	129	6-10	4	28	-24	3	-5	4	122	-120	-4	-1	4	161	143	
-5	2	3	185	184	-6	6	3	34	30	7-10	4	50	-49	4	-5	4	82	-85	-3	-1	4	239	249	
-2	-1	4	230	212	1	2	4	178	151	3	6	4	47	-46	3	-8	5	163	164	4	-4	5	143	-134
-1	-1	4	297	347	3	2	4	42	40	-7	7	4	93	-101	6	-8	5	73	-79	5	-4	5	305	306
0	-1	4	294	-376	4	2	4	161	159	-6	7	4	62	-63	7	-8	5	71	75	6	-4	5	207	202
1	-1	4	416	-424	6	2	4	46	-56	-5	7	4	49	53	8	-8	5	118	124	-5	-3	5	93	96
2	-1	4	89	106	-9	3	4	29	17	-4	7	4	43	-44	-4	-7	5	39	43	-4	-3	5	171	159
3	-1	4	203	185	-6	3	4	66	-70	-3	7	4	270	-269	-2	-7	5	39	-39	-3	-3	5	185	183
4	-1	4	25	29	-5	3	4	68	81	-2	7	4	197	-192	-1	-7	5	21	-37	-2	-3	5	20	-16
5	-1	4	36	-39	-4	3	4	100	86	-1	7	4	175	-177	0	-7	5	102	103	-1	-3	5	127	-144
7	-1	4	27	32	-3	3	4	161	-159	0	7	4	33	-35	1	-7	5	18	28	0	-3	5	34	-37
8	-1	4	69	80	-2	3	4	455	-435	-4	8	4	220	210	2	-7	5	211	-222	1	-3	5	494	463
-9	0	4	52	-43	-1	3	4	344	-323	-3	8	4	187	180	3	-7	5	183	-189	2	-3	5	141	156
-8	0	4	50	-61	0	3	4	186	-189	-2	8	4	102	-109	4	-7	5	40	-34	3	-3	5	100	-89
-7	0	4	36	-46	1	3	4	610	-603	-1	8	4	104	-113	5	-7	5	130	131	4	-3	5	281	-277
-6	0	4	114	-118	2	3	4	319	-308	0	8	4	62	-67	6	-7	5	38	-55	5	-3	5	133	-143
-5	0	4	58	-55	4	3	4	106	120	2	8	4	27	27	7	-7	5	110	-121	6	-3	5	101	106
-4	0	4	137	137	-8	4	4	42	-42	-8	9	4	40	39	8	-7	5	124	-140	-7	-2	5	70	71
-3	0	4	281	264	-6	4	4	97	97	-6	9	4	78	-79	-4	-6	5	81	81	-5	-2	5	150	-144
-2	0	4	250	257	-5	4	4	134	134	-4	9	4	105	105	-2	-6	5	93	-88	-2	-2	5	466	462
-1	0	4	40	48	-3	4	4	189	-165	-3	9	4	62	63	0	-6	5	91	89	-1	-2	5	563	-582
0	0	4	50	48	-2	4	4	150	131	-1	9	4	38	-38	1	-6	5	226	237	0	-2	5	77	68
1	0	4	244	-255	-1	4	4	304	304	-4	10	4	65	55	2	-6	5	149	-149	1	-2	5	372	-372
2	0	4	496	472	1	4	4	165	-168	-3	10	4	63	51	3	-6	5	464	-438	2	-2	5	75	79
3	0	4	390	389	2	4	4	172	-174	-2	10	4	37	-34	4	-6	5	195	-171	3	-2	5	201	212
5	0	4	157	-155	3	4	4	64	-68	-1	10	4	38	-42	5	-6	5	186	196	4	-2	5	49	39
7	0	4	80	85	4	4	4	63	62	1	11	5	20	-2	6	-6	5	170	167	5	-2	5	30	30
-9	1	4	55	58	5	4	4	96	94	3	-11	5	40	44	7	-6	5	110	-118	6	-2	5	54	48
-7	1	4	111	-104	6	4	4	38	-33	5	-11	5	39	-40	8	-6	5	58	-63	8	-2	5	37	-42
-6	1	4	52	-51	-7	5	4	95	-93	-1	-10	5	25	15	-6	-5	5	58	-59	-7	-1	5	28	-55
-4	1	4	175	-164	-5	5	4	150	145	1	-10	5	60	-63	-5	-5	5	23	-16	-6	-1	5	248	-243
-3	1	4	136	-143	-4	5	4	141	134	1	-10	5	54	-58	-5	-1	5	52	52	-5	-1	5	119	-115
-2	1	4	205	-197	-3	5	4	355	344	2	-10	5	53	58	-3	-5	5	30	-26	-4	-1	5	238	224
-1	1	4	64	-53	-2	5	4	117	119	3	-10	5	53	58	-1	-5	5	179	-174	-3	-1	5	75	86
0	1	4	506	494	-1	5	4	209	-215	4	-10	5	135	136	1	-5	5	181	174	-2	-1	5	167	-166
1	1	4	406	421	0	5	4	71	72	5	-10	5	42	48	2	-5	5	26	10	-1	-1	5	638	-663
1	1	4	487	-481	1	5	4	208	222	6	-10	5	36	-39	3	-5	5	52	-57	0	-1	5	89	-102
2	1	4	70	60	3	5	4	48	-48	4	-9	5	46	51	4	-5	5	97	91	1	-1	5	179	196
3	1	4	528	494	-9	6	4	52	52	1	-9	5	29	29	5	-5	5	58	67	2	-1	5	356	-362
6	1	4	68	-67	-8	6	4	46	56	2	-9	5	174	-169	9	-5	5	48	58	4	-1	5	93	-89
7	1	4	68	-75	-6	6	4	207	-202	3	-9	5	75	-68	-6	-4	5	35	-40	4	-1	5	74	68
-8	2	4	30	-28	-5	6	4	80	77	4	-9	5	74	72	-4	-4	5	128	129	5	-1	5	114	115
-7	2	4	39	-36	-4	6	4	197	186	5	-9	5	93	87	-3	-4	5	61	-60	6	-1	5	32	47
-6	2	4	23	35	-3	6	4	92	86	7	-9	5	51	-50	-2	-4	5	262	-245	7	-1	5	54	-55
-5	2	4	143	129	-2	6	4	31	34	-2	-8	5	70	-72	-1	-4	5	126	-117	8	-1	5	43	-48
-4	2	4	112	107	-1	6	4	42	39	0	-8	5	25	22	0	-4	5	329	326	-7	0	5	81	-78
-3	2	4	285	-300	0	6	4	131	131	1	-8	5	72	75	1	-4	5	92	91	-6	0	5	96	-98
-2	2	4	373	-371	1	6	4	107	106	2	-8	5	42	39	2	-4	5	44	41	-5	0	5	84	-78
-1	2	4	91	-82	1	6	4	81	-90	3	-8	5	159	152	0	-7	6	70	74	-2	-3	6	226	220
0	2	4	154	188	-8	4	5	78	78	1	-7	6	33	33	1	-7	6	194	-195	3	-3	6	275	252
-4	0	5	138	-114	-6	4	5	96	97	2	-7	6	79	75	3	-7	6	47	42	5	-3	6	50	-53
-3	0	5	16	17	-4	4	5	263	-265	-1	8	5	71	67	4	-7	6	270	246	6	-3	6	166	-167
-2	0	5	25	37	-3	4	5	237	-236	0	8	5	40	44	5	-7	6	101	93	7	-3	6	93	-99
-1	0	5	608	-624	-1	4	5	109	123	2	8	5	39	-37	8	-7	6	29	-5	-7	-2	6	68	-71
0	0	5	343	-395	0	4	5	268	260	-8	9	5	36	-50	5	-6	6	57	55	-6	-2	6	95	-97
1	0	5	37	-49	1	4	5	19	-17	-7	9	5	31	-34	-5	-6	6	105	-110	-4	-2	6	29	26
2	0	5	134	135	2	4	5	22	28	-5	9	5	25	-20	-3	-6	6	78	73	-3	-2	6	150	-142
4	0	5	50	-41	3	4	5	35	29	-4	9	5	121	-118	-1	-6	6	89	97	-1	-2	6	453	435
5	0	5	158	-161	4	4	5	23	-32	0	-6	6	85	-90	0	-6	6	29	-40	0	-2	6	87	-97
6	0	5	27	-47	5	4	5	51	-45	1	-6	6	160	-162	1	-6	6	112	110	1	-2	6	189	-195
-9	1	5	48	-46	-8	5	5	55	-64	2	-6	6	59	-70	3	-6	6	59	-70	4	-2	6	483	466
-7	1	5	318	318	-7	5	5	78	78	5	-6	6	79	87	5	-6	6	79	87	5	-2	6	50	-50
-6	1	5	167	171	-6	5	5	70	68	6	-6	6	258	272	6	-6	6	132	152	7	-2	6	29	12
-5	1	5	43	46	-4	5	5	128	-122	7	-6	6	104	109	8	-6	6	104	109	-4	-1	6	94	-92
-4	1	5	16	19	-3	5	5	321	-313	8	-6	6	27	31	-6	-5	6	27	31					

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR AM90021

H	K	L	10FO	10FC	H	K	L	10FO	10FC	H	K	L	10FO	10FC	H	K	L	10FO	10FC	H	K	L	10FO	10FC
6	2	5	79	84	-7	7	5	32	35	1	-8	6	53	-49	1	-4	6	180	-177	1	0	6	245	-270
-8	3	5	70	-78	-6	7	5	124	111	2	-8	6	51	-51	2	-4	6	357	-342	3	0	6	139	132
-7	3	5	35	-36	-5	7	5	206	203	3	-8	6	200	205	3	-4	6	502	-458	5	0	6	76	77
-6	3	5	89	89	-4	7	5	194	195	4	-8	6	43	34	4	-4	6	41	42	6	0	6	52	59
-4	3	5	84	-99	-3	7	5	158	160	5	-8	6	170	-173	5	-4	6	120	119	-8	1	6	37	43
-3	3	5	253	237	-2	7	5	59	-51	6	-8	6	42	-44	6	-4	6	176	-168	-7	1	6	180	-185
-2	3	5	520	532	0	7	5	76	79	7	-8	6	84	-93	7	-4	6	83	-89	-6	1	6	115	-114
-1	3	5	476	497	1	7	5	53	-54	8	-8	6	44	-56	-7	-3	6	53	-56	-5	1	6	127	137
0	3	5	219	-234	2	7	5	44	-42	-4	-7	6	49	-53	-6	-3	6	83	-81	-4	1	6	120	-114
1	3	5	131	-114	-8	8	5	34	-36	-3	-7	6	82	-81	-5	-3	6	64	-66	-3	1	6	31	-19
2	3	5	145	132	-7	8	5	107	-125	-2	-7	6	36	26	-4	-3	6	78	74	-2	1	6	142	-137
-9	4	5	57	-53	-5	8	5	136	138	-1	-7	6	125	133	-3	-3	6	42	40	-1	1	6	30	-30
0	1	6	138	143	-3	5	6	361	362	2	-9	7	116	118	0	-4	7	183	-190	-4	0	7	186	-192
1	1	6	31	-19	-2	5	6	146	141	3	-9	7	49	-48	1	-4	7	78	90	-3	0	7	120	-104
2	1	6	292	-300	-1	5	6	146	-162	4	-9	7	91	-90	2	-4	7	273	251	-2	0	7	35	29
3	1	6	130	-124	0	5	6	76	-73	5	-9	7	34	-25	4	-4	7	120	-121	-1	0	7	80	86
6	1	6	55	52	1	5	6	113	114	6	-9	7	67	60	5	-4	7	96	-97	0	0	7	80	60
7	1	6	94	-87	2	5	6	112	103	7	-9	7	46	47	6	-4	7	183	182	1	0	7	172	-174
-8	2	6	38	-40	4	5	6	29	-27	-3	-8	7	76	70	7	-4	7	144	150	2	0	7	35	-46
-7	2	6	65	-65	-7	6	6	74	79	-1	-8	7	73	-77	-5	-3	7	48	42	3	0	7	267	257
-6	2	6	115	-116	-6	6	6	75	78	0	-8	7	63	-65	-4	-3	7	63	-54	4	0	7	209	208
-5	2	6	31	29	-5	6	6	138	-132	1	-8	7	26	-35	-2	-3	7	331	322	5	0	7	66	-56
-4	2	6	44	42	-4	6	6	55	59	2	-8	7	107	103	-1	-3	7	61	66	6	0	7	127	-128
-3	2	6	122	125	-3	6	6	95	89	3	-8	7	113	108	0	-3	7	367	-370	-8	1	7	59	-70
-2	2	6	61	49	-2	6	6	89	97	4	-8	7	103	-101	1	-3	7	69	-77	-7	1	7	33	29
-1	2	6	72	-80	-1	6	6	26	22	-3	-7	7	64	63	2	-3	7	359	363	-6	1	7	106	101
0	2	6	46	50	4	6	6	52	-47	-2	-7	7	28	24	3	-3	7	96	96	-5	1	7	83	-82
2	2	6	29	-19	-9	7	6	49	-54	0	-7	7	28	33	4	-3	7	127	116	-4	1	7	73	75
3	2	6	190	-187	-8	7	6	62	-58	1	-7	7	57	57	6	-3	7	45	-59	-3	1	7	132	-136
4	2	6	96	-100	-5	7	6	25	-24	2	-7	7	55	59	7	-3	7	85	83	-2	1	7	257	-266
5	2	6	49	-49	-4	7	6	78	-74	3	-7	7	92	-87	8	-3	7	39	40	-1	1	7	162	193
6	2	6	33	-35	-3	7	6	98	104	4	-7	7	166	-156	-6	-2	7	59	59	0	1	7	332	326
-8	3	6	111	116	-2	7	6	128	121	5	-7	7	44	45	-4	-2	7	107	-103	1	1	7	36	41
-7	3	6	86	87	0	7	6	129	-128	6	-7	7	266	258	-3	-2	7	85	69	-3	2	7	123	-128
-6	3	6	134	-120	2	7	6	36	33	7	-7	7	109	122	-2	-2	7	371	376	3	1	7	27	-21
-5	3	6	156	-156	-8	8	6	32	-37	-5	-6	7	65	-70	-1	-2	7	279	290	4	1	7	61	51
-3	3	6	76	85	-6	8	6	85	-83	-4	-6	7	65	-65	0	-2	7	58	-63	5	1	7	90	86
-2	3	6	328	319	-5	8	6	246	-246	-3	-6	7	75	73	1	-2	7	464	-475	-9	2	7	76	85
-1	3	6	166	167	-4	8	6	140	-136	-1	-6	7	135	-133	2	-2	7	696	-668	-8	2	7	76	81
1	3	6	78	74	-3	8	6	54	57	2	-6	7	228	235	3	-2	7	126	-108	-7	2	7	42	-42
2	3	6	121	-135	-2	8	6	91	89	3	-6	7	67	-64	4	-2	7	30	36	-5	2	7	92	94
3	3	6	47	-40	-1	8	6	74	-90	4	-6	7	404	-395	5	-2	7	126	-120	-4	2	7	222	218
5	3	6	27	-7	0	8	6	118	-116	5	-6	7	178	-174	6	-2	7	99	-91	-3	2	7	211	198
6	3	6	55	44	1	8	6	80	-82	6	-6	7	66	-62	-7	-1	7	60	-57	-2	2	7	312	-322
-8	4	6	68	68	2	8	6	28	24	7	-6	7	46	-58	-6	-1	7	50	-51	-1	2	7	129	113
7	4	6	74	81	-7	9	6	39	42	-4	-5	7	63	-59	-5	-1	7	41	28	0	2	7	154	174
-5	4	6	130	-129	-6	9	6	54	51	-2	-5	7	166	168	-4	-1	7	65	62	1	2	7	105	118
-4	4	6	31	38	-5	9	6	45	40	-1	-5	7	59	-50	-3	-1	7	55	-61	3	2	7	89	-90
-2	4	6	150	-133	-4	9	6	29	38	0	-5	7	233	-236	-2	-1	7	111	-103	5	2	7	85	74
-1	4	6	27	32	-1	9	6	39	40	-1	-1	7	129	-126	-1	-1	7	351	354	6	2	7	89	86
0	4	6	102	-94	-6	10	6	51	57	2	-5	7	20	7	0	-1	7	35	40	-9	3	7	87	87
1	4	6	18	12	-5	10	6	59	62	3	-5	7	257	248	1	-1	7	49	-46	-8	3	7	78	-80
2	4	6	73	79	1	-11	7	30	31	4	-5	7	133	121	2	-1	7	104	93	-7	3	7	133	-135
3	4	6	91	-89	3	-11	7	45	-47	5	-5	7	81	-83	4	-1	7	31	-28	-6	3	7	28	39
5	4	6	97	95	4	-11	7	25	-29	6	-5	7	72	-72	5	-1	7	173	-173	-5	3	7	56	51
6	4	6	54	48	0	-10	7	19	21	7	-5	7	26	25	6	-1	7	49	-62	-4	3	7	107	-92
-7	5	6	84	90	1	-10	7	39	35	-6	-4	7	41	41	8	-1	7	32	41	-3	3	7	103	-91
-6	5	6	61	-53	6	-10	7	39	33	-4	-4	7	72	-73	-8	0	7	64	-76	-2	3	7	170	-164
-5	5	6	134	-133	-1	-9	7	70	-64	-2	-4	7	82	74	-7	0	7	105	-105	0	3	7	245	267
-4	5	6	22	16	1	-9	7	81	77	-1	-4	7	42	41	-5	0	7	76	80	1	3	7	76	-74
2	3	7	131	-131	-3	8	7	75	70	6	-5	8	165	169	1	-1	8	374	394	-4	3	8	115	-103
3	3	7	94	-86	0	8	7	38	40	7	-5	8	42	-32	2	-1	8	66	60	-3	3	8	229	-226
5	3	7	82	77	-7	9	7	26	28	8	-5	8	54	-65	3	-1	8	30	9	-1	3	8	268	276
-8	4	7	30	-38	-1	9	7	50	-48	8	-5	8	33	-36	5	-1	8	31	-28	0	3	8	26	-9
-7	4	7	94	-97	-5	10	7	89	-95	-7	-4	8	59	-62	6	-1	8	37	26	1	3	8	94	-112
-5	4	7	163	161	-3	10	7	39	32	-6	-4	8	74	-73	7	-1	8	26	34	5	3	8	55	-50
-4	4	7	161	164	-3	11	8	39	38	-7	0	8	146	-146	-7	0	8	87	85	-9	4	8	65	64
-3	4	7	137	-143	4	-11	8	30	30	-6	0	8	91	-95	-6	0	8	84	81	-6	4	8	81	74
-2	4	7	71	-60	0	-10	8	67	68	-5	0	8	62	68	-5	0	8	141	-136	-5	4	8	137	133
-1	4	7	65	-75	1	-10	8	28	33	-4	0	8	27	-40	-4	0	8	27	-40	-4	4	8	59	54
0	4	7	120	117	2	-10	8	120	-115	-3	0	8	274	281	-3	0	8	227	212	-3	4	8	70	-65
1	4	7	189	195	2	-10	8	88	-81	-2	-4	8	53	-45	-2	0	8	83	87	-2	4	8	27	-2

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR AM90021

H	K	L	10FO	10FC	H	K	L	10FO	10FC	H	K	L	10FO	10FC	H	K	L	10FO	10FC	H	K	L	10FO	10FC
4	6	7	70	61	-4	-6	8	87	86	1	-2	8	226	218	-7	2	8	25	-26	2	6	8	83	-83
-7	7	7	70	-85	-2	-6	8	94	-95	2	-2	8	101	-116	-6	2	8	55	55	-4	7	8	60	-61
-6	7	7	148	-142	1	-6	8	240	243	3	-2	8	240	-237	-5	2	8	179	-180	-3	7	8	95	-99
-5	7	7	100	-106	2	-6	8	315	310	4	-2	8	111	-103	-4	2	8	225	-214	-2	7	8	101	106
-4	7	7	150	157	3	-6	8	43	-51	5	-2	8	95	89	-3	2	8	280	-263	-1	7	8	40	55
-3	7	7	93	110	4	-6	8	60	-58	6	-2	8	99	95	-1	2	8	191	174	0	7	8	29	21
-2	7	7	77	-82	5	-6	8	82	-76	7	-2	8	44	45	0	2	8	101	-109	2	7	8	34	-25
-1	7	7	54	-52	6	-6	8	26	-14	8	-2	8	31	-20	1	2	8	32	37	-4	8	8	81	-75
0	7	7	41	35	7	-6	8	50	51	-7	-1	8	71	77	2	2	8	131	140	-1	8	8	24	23
1	7	7	48	47	-6	-5	8	64	-58	-6	-1	8	97	103	3	2	8	128	122	-6	9	8	37	-39
-8	8	7	87	85	-3	-5	8	59	52	-5	-1	8	75	-72	4	2	8	112	-111	-5	9	8	39	41
-7	8	7	35	48	-2	-5	8	53	-51	-4	-1	8	68	-62	5	2	8	129	-127	-4	9	8	32	43
-6	8	7	42	-30	0	-5	8	135	139	-3	-1	8	115	108	-9	3	8	60	-66	-3	9	8	51	-50
-5	8	7	46	55	3	-5	8	145	-144	-2	-1	8	274	275	-7	3	8	119	133	-2	9	8	66	-65
-4	8	7	81	84	5	-5	8	261	259	-1	-1	8	141	-138	-6	3	8	49	44	3	-11	9	33	29
4	-11	9	27	25	1	-4	9	213	214	3	0	9	20	-26	2	4	9	183	182	-1	-7	10	105	-109
5	-11	9	30	-25	2	-4	9	135	130	4	0	9	92	93	4	4	9	51	-42	0	-7	10	29	-29
1	-10	9	90	-90	4	-4	9	32	-23	5	0	9	119	115	-8	5	9	47	42	1	-7	10	227	222
2	-10	9	25	29	6	-4	9	74	76	6	0	9	86	90	-7	5	9	38	42	2	-7	10	306	302
3	-10	9	77	77	7	-4	9	65	64	-8	1	9	60	61	-6	5	9	63	-57	3	-7	10	204	200
6	-10	9	35	21	-5	-3	9	58	-43	-7	1	9	50	51	-5	5	9	145	-133	4	-7	10	51	-47
-2	-9	9	65	67	-4	-3	9	87	78	-6	1	9	84	-90	-4	5	9	69	-59	5	-7	10	62	-53
-1	-9	9	36	42	-1	-3	9	234	242	-5	1	9	61	-64	-2	5	9	156	154	6	-7	10	107	108
0	-9	9	55	-54	0	-3	9	59	55	-4	1	9	62	69	-1	5	9	90	-101	7	-7	10	40	43
2	-9	9	102	104	-3	-3	9	65	-67	-3	1	9	166	163	0	5	9	106	-116	-1	-6	10	68	-68
3	-9	9	137	137	3	-3	9	75	-60	-2	1	9	223	-207	4	5	9	37	-32	0	-6	10	169	-169
5	-9	9	65	-63	4	-3	9	152	149	-1	1	9	142	-161	-7	6	9	49	43	1	-6	10	219	-219
7	-9	9	63	67	5	-3	9	266	255	0	1	9	213	228	-6	6	9	57	58	2	-6	10	20	14
-3	-8	9	42	-37	6	-3	9	47	-39	1	1	9	252	239	-5	6	9	75	-75	3	-6	10	128	129
-1	-8	9	42	40	7	-3	9	105	-115	2	1	9	56	66	-4	6	9	80	-74	5	-6	10	63	-65
1	-8	9	126	-127	-7	-2	9	35	-33	3	1	9	93	-94	-1	6	9	44	45	7	-6	10	42	48
2	-8	9	154	-150	-5	-2	9	25	-8	4	1	9	88	-83	0	6	9	60	-60	-6	-5	10	45	35
4	-8	9	93	87	-3	-2	9	91	-80	6	1	9	52	43	-7	7	9	61	77	-4	-5	10	46	-58
5	-8	9	63	-59	-2	-2	9	193	-206	7	1	9	44	46	-6	7	9	48	41	-3	-5	10	88	-83
8	-8	9	44	52	-1	-2	9	188	202	-8	2	9	102	114	-5	7	9	31	-24	-2	-5	10	75	69
-3	-7	9	88	-95	0	-2	9	390	399	-7	2	9	109	114	-4	7	9	47	-46	-1	-5	10	63	73
-2	-7	9	51	-49	1	-2	9	80	79	-6	2	9	131	-131	-3	7	9	58	57	0	-5	10	22	-36
-1	-7	9	61	59	2	-2	9	202	-194	-4	2	9	210	193	-2	7	9	39	49	1	-5	10	71	-74
0	-7	9	69	69	3	-2	9	59	-53	3	2	9	72	82	2	7	9	52	50	2	-5	10	118	-122
2	-7	9	53	-49	4	-2	9	75	73	-2	2	9	147	134	-5	8	9	26	-14	4	-5	10	108	-104
4	-7	9	29	-30	7	-2	9	43	-35	-1	2	9	235	-219	-2	8	9	29	35	6	-5	10	80	85
5	-7	9	145	-146	-6	-1	9	87	-84	0	2	9	113	-114	-1	8	9	59	61	-6	-4	10	44	44
6	-7	9	100	-96	-5	-1	9	36	-47	1	2	9	147	150	-6	9	9	34	36	0	-4	10	83	-92
-2	-6	9	56	56	2	-1	9	87	88	2	2	9	215	227	-3	9	9	29	22	1	-4	10	86	-79
0	-6	9	44	-45	4	2	9	91	-83	-2	9	9	91	-83	-2	9	9	40	42	2	-4	10	69	72
3	-6	9	225	222	-2	-1	9	207	-200	5	2	9	34	-32	1	-10	10	97	91	3	-4	10	60	59
4	-6	9	185	173	-1	-1	9	325	-340	6	2	9	47	-54	3	-10	10	151	-151	5	-4	10	168	-167
6	-6	9	103	-114	0	-1	9	121	-112	6	2	9	65	62	4	-10	10	74	-72	6	-4	10	107	-113
7	-6	9	61	-65	1	-1	9	281	-259	-6	3	9	65	62	6	-10	10	39	50	8	-4	10	65	69
8	-6	9	56	52	2	-1	9	59	60	-3	3	9	120	114	-1	-9	10	59	-62	-5	-3	10	45	38
-1	-5	9	179	182	3	-1	9	31	34	-2	3	9	39	30	0	-9	10	36	-33	-4	-3	10	52	-51
0	-5	9	81	76	-1	3	9	228	-211	-1	3	9	75	-75	2	-9	10	31	25	-3	-3	10	140	-139
1	-5	9	36	-40	5	-1	9	122	-124	1	3	9	163	160	4	-9	10	52	-55	-2	-3	10	194	208
2	-5	9	216	-219	6	-1	9	39	-45	2	3	9	41	47	5	-9	10	81	-69	-1	-3	10	216	213
3	-5	9	162	-153	7	-1	9	50	-49	3	3	9	37	-35	6	-9	10	44	-39	0	-3	10	126	122
4	-5	9	116	111	-8	0	9	56	58	4	3	9	62	-49	-2	-8	10	60	66	1	-3	10	34	-31
6	-5	9	99	-106	-6	0	9	49	-49	-9	4	9	89	-92	0	-8	10	67	-70	2	-3	10	82	-80
7	-5	9	49	-57	-5	0	9	60	67	-6	4	9	124	-116	2	-8	10	109	105	3	-3	10	112	111
-5	-4	9	40	42	-4	0	9	25	-33	-5	4	9	105	-106	3	-8	10	62	53	4	-3	10	96	93
-3	-4	9	50	-52	-3	0	9	78	76	-4	4	9	111	121	4	-8	10	81	-82	5	-3	10	146	-136
-2	-4	9	48	-52	-2	0	9	72	-79	-3	4	9	111	121	5	-8	10	57	-46	6	-3	10	136	-136
-1	-4	9	40	-44	-1	0	9	176	159	-1	4	9	138	-137	-4	-7	10	32	-30	8	-3	10	40	-43
0	-4	9	89	92	0	0	9	175	-187	0	4	9	90	-95	-3	-7	10	86	84	-5	-2	10	38	35
-3	-2	10	297	-294	1	4	9	134	139	-1	4	9	134	139	0	-3	11	59	-65	-1	2	11	145	130
-2	-2	10	164	-165	2	-10	11	41	-47	2	-10	11	22	-20	3	-3	11	37	-45	0	2	11	290	-296
-1	-2	10	35	46	4	-10	11	62	65	1	-9	11	62	65	1	-3	11	72	-81	1	2	11	189	-206
0	-2	10	164	184	1	2	10	146	-139	2	-9	11	41	33	3	-3	11	99	111	2	2	11	190	111
1	-2	10	263	267	0	2	10	149	133	2	-8	11	63	-57	4	-3	11	157	157	3	2	11	163	157
3	-2	10	70	-56	2	2	10	26	23	-2	-8	11	63	-57	5	-3	11	157	157	4	2	11	74	75
4	-2	10	38	-46	0	-8	11	124	132	6	-3	11	77	81	6	-3	11	77	81	7	3	11	36	34
5	-2	10	66	74	3	2	10	69	66	7	-2	11	47	47	-7	-2	11	47	47	-6	3	11	81	79
6	-2	10	49	42	4	2	10	60	61															

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR AM90021

H	K	L	10FO	10FC	H	K	L	10FO	10FC	H	K	L	10FO	10FC	H	K	L	10FO	10FC	H	K	L	10FO	10FC
-7	1	10	78	-85	1	5	10	44	-42	0	-5	11	196	194	-2	0	11	94	87	-6	6	11	47	37
-6	1	10	55	50	2	5	10	55	-53	1	-5	11	200	196	-1	0	11	74	74	-5	6	11	59	53
-5	1	10	128	132	3	5	10	34	-37	2	-5	11	129	132	0	0	11	68	66	-3	6	11	93	-94
-4	1	10	93	87	-8	6	10	35	37	3	-5	11	105	-101	3	0	11	162	-170	-2	6	11	45	-44
-3	1	10	102	-99	-4	6	10	54	51	4	-5	11	106	-111	4	0	11	54	-51	-1	6	11	28	-28
-2	1	10	315	-311	-3	6	10	58	-59	5	-5	11	105	111	5	0	11	74	-69	0	6	11	45	50
-1	1	10	148	-158	-2	6	10	49	-45	6	-5	11	51	54	6	0	11	52	-49	-5	7	11	29	27
0	1	10	116	-110	0	6	10	98	107	-2	-4	11	41	39	-6	1	11	39	40	-3	7	11	56	-55
1	1	10	109	110	1	6	10	103	110	-1	-4	11	106	-110	-3	1	11	42	-31	-1	7	11	28	20
2	1	10	39	34	-6	7	10	65	-53	0	-4	11	105	-113	-2	1	11	54	51	-5	8	11	40	33
3	1	10	106	-112	-5	7	10	39	-41	1	-4	11	98	89	1	1	11	20	35	-2	8	11	21	-18
4	1	10	229	-210	-2	7	10	39	-39	2	-4	11	93	98	2	1	11	249	231	-1	8	11	30	25
5	1	10	115	-123	1	7	10	22	20	3	-4	11	61	70	3	1	11	400	394	2	-10	12	36	38
6	1	10	64	-64	-7	8	10	37	39	4	-4	11	63	-63	4	1	11	155	161	3	-10	12	47	44
-8	2	10	37	-31	-5	8	10	23	-27	-7	-3	11	46	41	-7	2	11	39	39	-1	-9	12	32	31
-7	2	10	90	-92	-4	9	10	24	18	-5	-3	11	52	-42	-6	2	11	46	57	0	-9	12	26	-23
-5	2	10	128	129	0	-10	11	65	61	-4	-3	11	167	-158	-5	2	11	73	-73	1	-9	12	116	-112
-4	2	10	67	-74	1	-10	11	38	40	-1	-3	11	67	-75	-2	2	11	26	34	2	-9	12	86	-83
3	-9	12	115	115	-5	-2	12	77	75	-1	2	12	107	-105	0	-7	13	61	-58	0	-1	13	115	116
4	-9	12	82	79	-4	-2	12	95	96	0	2	12	87	-93	1	-7	13	143	-146	1	-1	13	153	-157
6	-9	12	76	-78	-3	-2	12	162	159	2	2	12	25	-39	2	-7	13	78	80	2	-1	13	47	-52
0	-8	12	41	-43	-2	-2	12	134	-134	2	2	12	155	-151	3	-7	13	121	116	4	-1	13	38	24
1	-8	12	106	-112	-1	-2	12	119	-122	3	2	12	57	-59	5	-7	13	98	-101	-7	0	13	26	-32
2	-8	12	46	-40	0	-2	12	142	136	-5	3	12	22	-17	-4	-6	13	73	-70	-6	0	13	53	50
3	-8	12	28	20	1	-2	12	117	128	-3	3	12	38	43	-3	-6	13	42	-44	-5	0	13	67	67
4	-8	12	42	38	2	-2	12	82	83	-2	3	12	41	35	-1	-6	13	37	-45	-4	0	13	81	84
-2	-7	12	65	67	3	-2	12	123	-117	-1	3	12	53	46	0	-6	13	68	-73	-3	0	13	112	-100
2	-7	12	22	-23	4	-2	12	57	-49	0	3	12	91	-88	1	-6	13	98	-100	-2	0	13	98	-87
4	-7	12	36	31	-6	-1	12	71	-68	1	3	12	67	55	2	-6	13	39	-31	-1	0	13	33	27
6	-7	12	61	-59	-5	-1	12	118	-113	2	3	12	146	152	-5	-5	13	40	40	0	0	13	24	41
-3	-6	12	47	-44	-4	-1	12	93	82	3	3	12	50	64	-3	-5	13	46	49	2	0	13	113	-117
-1	-6	12	170	176	-3	-1	12	73	75	-7	4	12	31	29	-2	-5	13	45	-42	4	0	13	30	19
0	-6	12	166	171	-2	-1	12	38	-34	-6	4	12	43	-50	0	-5	13	83	89	5	0	13	76	73
2	-6	12	103	-103	-1	-1	12	82	-86	-4	4	12	32	45	1	-5	13	78	83	-7	1	13	48	-45
3	-6	12	68	67	0	-1	12	37	42	-3	4	12	70	64	2	-5	13	31	-31	-6	1	13	41	-45
4	-6	12	163	157	1	-1	12	36	-28	-2	4	12	168	166	3	-5	13	47	37	-5	1	13	68	65
5	-6	12	77	79	2	-1	12	126	-132	-1	4	12	30	-31	5	-5	13	58	-58	-4	1	13	87	72
6	-6	12	47	-38	4	-1	12	61	64	0	4	12	88	-85	6	-5	13	48	42	-3	1	13	25	26
7	-6	12	67	-72	5	-1	12	56	56	1	4	12	51	46	-5	-4	13	66	65	-1	1	13	57	71
-5	-5	12	53	44	6	-1	12	69	-67	2	4	12	119	121	-3	-4	13	19	-28	0	1	13	117	-107
-4	-5	12	30	-37	-4	0	12	65	66	-7	5	12	28	29	-2	-4	13	77	83	1	1	13	213	-221
-3	-5	12	122	-123	-3	0	12	140	137	-6	5	12	82	75	-1	-4	13	187	200	2	1	13	143	-135
-2	-5	12	80	-73	-2	0	12	104	-108	-5	5	12	31	-23	1	-4	13	122	-129	3	1	13	57	44
0	-5	12	67	-63	-1	0	12	39	-40	-3	5	12	43	-41	3	-4	13	70	65	4	1	13	105	98
1	-5	12	27	27	0	0	12	38	43	-2	5	12	81	-90	4	-4	13	146	145	5	1	13	41	-40
2	-5	12	32	33	1	0	12	174	183	-1	5	12	21	-19	6	-4	13	49	-50	-5	2	13	37	34
3	-5	12	142	-140	2	0	12	103	93	-7	6	12	26	23	-6	-3	13	55	50	-4	2	13	39	37
-3	-4	12	21	26	3	0	12	88	91	-5	6	12	31	25	-4	-3	13	41	-46	-2	2	13	43	44
-2	-4	12	87	-80	5	0	12	28	-35	-2	6	12	57	-56	-3	-3	13	72	-70	-1	2	13	133	143
-1	-4	12	177	-187	6	0	12	37	32	-1	6	12	54	-55	-2	-3	13	56	47	0	2	13	114	123
0	-4	12	33	-30	-6	1	12	55	-55	0	6	12	52	-62	0	-3	13	52	51	1	2	13	24	-21
1	-4	12	35	-45	-5	1	12	39	-48	1	6	12	26	-43	2	-3	13	89	-102	2	2	13	63	73
2	-4	12	55	58	-4	1	12	89	89	-2	7	12	30	46	3	-3	13	36	-40	3	2	13	59	51
3	-4	12	49	47	-3	1	12	212	201	0	7	12	25	28	5	-3	13	36	35	-4	3	13	96	82
5	-4	12	94	90	-2	1	12	46	-52	2	-10	13	71	70	-4	-2	13	22	19	-2	3	13	87	-95
-6	-3	12	24	-22	-1	1	12	337	-334	4	-10	13	41	-38	-3	-2	13	72	-70	-1	3	13	114	-107
-5	-3	12	56	60	0	1	12	268	-272	0	-9	13	48	-46	-2	-2	13	195	-191	1	3	13	93	104
-4	-3	12	115	119	1	1	12	63	-71	2	-9	13	63	66	0	-2	13	112	-123	2	3	13	52	-39
-2	-3	12	194	-195	2	1	12	182	183	3	-9	13	46	29	1	-2	13	46	47	3	3	13	65	-70
0	-3	12	111	112	3	1	12	26	-24	5	-9	13	46	46	2	-2	13	40	40	-4	4	13	29	-35
1	-3	12	85	78	4	1	12	60	-68	-2	-8	13	37	33	4	-2	13	56	46	-3	4	13	24	17
2	-3	12	175	-173	5	1	12	46	39	0	-8	13	56	-53	5	-2	13	32	-25	-2	4	13	64	-64
3	-3	12	133	-133	-6	2	12	31	-33	2	-8	13	34	-32	-7	-1	13	32	-26	0	4	13	33	26
5	-3	12	65	61	-5	2	12	73	-65	6	-8	13	58	56	-5	-1	13	90	93	1	4	13	60	-59
6	-3	12	59	58	-3	2	12	70	55	-2	-7	13	67	64	-3	-1	13	115	-119	2	4	13	92	-79
-7	-2	12	25	-32	-2	2	12	93	88	-1	-7	13	42	50	-1	-1	13	136	137	-6	5	13	30	-29
-3	5	13	35	-30	0	-3	14	93	101	3	3	14	76	68	1	-3	15	143	154	-2	-7	16	69	67
-1	5	13	76	-75	2	-3	14	74	72	3	3	14	55	50	2	-3	15	53	59	-1	-7	16	79	-77
1	5	13	29	32	-1	4	14	41	-40	-2	4	14	57	57	3	-3	15	111	-106	3	-7	16	38	38
-5	6	13	34	27	3	-3	14	86	-84	0	4	14	33	-40	-6	-2	15	51	-54	-3	-6	16	28	-24
-4	6	13	79	73	4	-3	14	41	28	1														

OBSERVED AND CALCULATED STRUCTURE FACTORS FOR AM90021

H	K	L	10FO	10FC	H	K	L	10FO	10FC	H	K	L	10FO	10FC	H	K	L	10FO	10FC	H	K	L	10FO	10FC
-2	-6	14	39	-35	-6	1	14	51	54	5	-6	15	54	62	-1	1	15	43	46	-1	-3	16	34	33
-1	-6	14	22	-15	-5	1	14	48	-50	6	-6	15	36	44	0	1	15	32	35	1	-3	16	30	-39
0	-6	14	36	40	-4	1	14	31	-37	-4	-5	15	52	51	3	1	15	33	37	2	-3	16	37	-44
3	-6	14	50	-43	-3	1	14	61	65	-2	-5	15	45	-46	-4	2	15	46	-36	-5	-2	16	46	51
5	-6	14	71	72	-2	1	14	192	189	-1	-5	15	49	-47	-3	2	15	82	-73	-4	-2	16	33	-32
6	-6	14	53	57	-1	1	14	142	137	0	-5	15	70	72	-2	2	15	43	-55	-2	-2	16	95	98
-3	-5	14	60	58	0	1	14	65	-64	1	-5	15	35	37	0	2	15	51	-53	-1	-2	16	101	108
-2	-5	14	45	50	1	1	14	96	-102	5	-5	15	32	-27	2	2	15	41	52	1	-2	16	85	-83
-1	-5	14	63	-62	2	1	14	53	-51	-4	-4	15	60	60	-3	3	15	35	37	3	-2	16	39	43
0	-5	14	163	-172	-4	2	14	90	-86	-3	-4	15	33	-30	-1	3	15	31	-27	4	-2	16	68	70
2	-5	14	140	141	-3	2	14	51	-34	-1	-4	15	88	96	0	3	15	41	-41	5	-2	16	32	-21
4	-5	14	146	-141	-1	2	14	73	74	0	-4	15	56	61	2	3	15	43	38	-4	-1	16	38	41
6	-5	14	26	36	1	2	14	78	-77	1	-4	15	34	41	-2	4	15	30	21	-1	-1	16	39	40
-5	-4	14	73	-75	2	2	14	85	-89	2	-4	15	39	-35	1	4	15	73	-77	0	-1	16	138	142
-3	-4	14	98	103	3	2	14	55	-59	4	-4	15	45	37	-3	5	15	35	33	1	-1	16	32	39
-1	-4	14	86	-86	4	2	14	29	27	5	-4	15	42	51	-2	5	15	36	36	2	-1	16	77	-84
0	-4	14	163	-170	-7	3	14	30	-21	-4	-3	15	35	36	0	5	15	44	-40	3	-1	16	42	-37
3	-4	14	78	-85	-2	3	14	70	76	-3	-3	15	44	49	-3	6	15	46	35	-6	0	16	53	-58
4	-4	14	69	-71	0	3	14	40	-39	-2	-3	15	67	-55	2	-8	16	25	27	-4	0	16	37	30
5	-4	14	35	-45	1	3	14	74	-67	-1	-3	15	124	-129	3	-8	16	32	26	-3	0	16	101	-98
-1	-3	14	25	30	2	3	14	48	39	0	-3	15	40	-35	4	-8	16	52	-47	-2	0	16	122	-115
-1	0	16	97	94	2	-6	17	71	66	3	-2	17	34	35	-1	-6	18	29	-24	-2	2	18	26	-11
0	0	16	81	86	4	-6	17	97	-99	4	-2	17	44	37	0	-6	18	22	27	-1	2	18	38	-38
-4	1	16	103	-102	-3	-5	17	52	54	-4	-1	17	83	-69	1	-6	18	48	54	-2	3	18	42	34
-3	1	16	79	-86	-2	-5	17	54	50	-3	-1	17	90	-82	-2	-5	18	50	-51	-1	-5	19	49	46
-2	1	16	78	-78	0	-5	17	64	-62	-1	-1	17	46	44	-1	-5	18	31	-26	0	-5	19	23	-21
0	1	16	48	49	2	-5	17	55	53	-4	0	17	77	-71	-2	-4	18	44	-39	1	-5	19	55	-57
1	1	16	83	-70	3	-5	17	71	70	-2	0	17	56	55	1	-4	18	52	46	0	-4	19	37	34
2	1	16	75	-74	-4	-4	17	33	-22	0	0	17	33	-38	3	-4	18	79	-77	-3	-3	19	51	-46
-4	2	16	85	81	-3	-4	17	52	51	2	0	17	23	21	-3	-3	18	46	42	0	-3	19	28	-20
-3	2	16	55	54	1	-4	17	37	-43	3	0	17	27	31	-1	-3	18	41	-40	-3	-2	19	55	-38
0	2	16	52	49	4	-4	17	24	-20	-5	1	17	54	55	2	-3	18	32	35	-2	-2	19	35	-33
2	2	16	25	-25	-3	-3	17	49	48	-4	1	17	56	63	0	-2	18	19	-7	0	-2	19	34	38
3	2	16	34	-33	-2	-3	17	75	72	-1	1	17	89	79	-2	-1	18	34	-35	2	-2	19	77	-64
-4	3	16	35	24	-1	-3	17	33	-33	0	1	17	41	45	-1	-1	18	55	-62	-2	-1	19	77	-67
-3	3	16	63	51	0	-3	17	41	-48	-5	2	17	62	51	0	-1	18	55	55	2	-1	19	28	23
0	3	16	24	-21	1	-3	17	55	64	-3	2	17	72	-68	1	-1	18	105	111	-3	0	19	33	36
-5	4	16	25	-24	2	-3	17	84	87	-1	2	17	38	37	3	-1	18	34	-20	-1	0	19	27	-32
-2	4	16	30	32	4	-3	17	36	-25	-5	3	17	52	-48	-3	0	18	34	-31	1	0	19	23	28
0	4	16	31	22	-5	-2	17	42	-45	-4	3	17	42	-45	-2	0	18	40	-47	-2	1	19	42	-43
-1	5	16	30	-20	-4	-2	17	40	-35	-3	3	17	30	-35	-4	1	18	63	-54	0	1	19	66	60
-1	-7	17	35	-33	1	-2	17	37	-35	-1	3	17	28	31	-3	1	18	32	-38	0	-3	20	36	-35
3	-7	17	32	-39	2	-2	17	35	-38	-3	4	17	34	-26	-3	2	18	41	44	-1	-1	20	25	-21
-1	-6	17	44	-40																				

DEVIATIONS GREATER THAN 2 SIGMA

H	K	L	FO	FC	D/SIGMA
4	0	0	11.14	7.80	3.88
1	1	0	40.76	35.80	2.01
0	2	0	7.61	6.47	2.18
0	3	0	2.75	1.69	2.66
1	-4	1	4.31	5.58	2.74
-1	-3	1	18.56	21.97	2.93
5	-3	1	9.28	6.97	2.86
1	-2	1	9.32	6.76	3.92
1	-1	1	5.42	9.26	8.00
4	-1	1	11.71	9.28	2.81
1	0	1	25.84	20.05	3.68
-5	1	1	12.89	10.39	2.64
-3	1	1	17.67	14.06	3.22
-2	1	1	10.45	12.42	2.74
-4	2	1	9.63	7.75	2.35
6	2	1	3.03	0.45	2.10
-2	4	1	20.18	17.61	2.05
0	-3	2	20.53	16.62	3.25
-4	-1	2	4.04	1.83	3.47
0	-1	2	13.82	16.54	3.33
3	-1	2	24.41	21.26	2.08
6	-1	2	8.02	6.36	2.30
3	0	2	11.11	9.10	2.48
-1	1	2	26.96	34.74	4.75
0	1	2	11.29	17.59	9.23
1	1	2	19.19	16.22	2.51
-3	2	2	25.36	21.85	2.25
-1	3	2	8.70	7.13	2.32
-3	5	2	16.97	14.31	2.39
0	5	2	9.88	11.91	3.03
-3	10	2	7.10	5.50	2.23
-2	-3	3	4.70	6.71	4.06
2	-2	3	17.67	15.21	2.22
-7	-1	3	8.49	6.52	2.66
8	-1	3	8.57	11.67	3.67
-3	0	3	16.97	14.50	2.27
0	0	3	7.28	3.35	7.99
-1	1	3	31.97	27.16	2.48
0	1	3	18.30	13.06	4.88
1	1	3	3.86	5.88	4.46
7	1	3	3.57	0.94	2.22
-2	2	3	13.44	15.26	2.06
2	2	3	8.03	10.30	3.22
1	3	3	13.68	11.30	2.59
-6	4	3	5.74	3.64	3.00
0	4	3	11.90	13.50	2.13
-1	-6	4	6.31	4.59	3.02
-3	-3	4	9.94	7.57	2.85
3	-3	4	12.63	10.27	2.71
2	-2	4	14.84	12.59	2.34
-1	-1	4	29.74	34.73	2.75
0	-1	4	29.40	37.59	4.78
2	-1	4	8.90	10.59	2.40
0	2	4	15.44	18.84	3.69
1	2	4	17.79	15.07	2.41
-5	3	4	6.83	8.13	2.16
3	-4	5	3.14	0.94	2.96
-1	-3	5	12.69	14.42	2.01
0	-1	5	8.94	10.18	2.12
-4	0	5	13.85	11.42	2.49
0	0	5	34.28	39.48	2.61
1	0	5	3.70	4.94	2.64
-2	1	5	22.73	29.16	4.54
5	1	5	4.54	2.54	2.54
0	2	5	4.29	2.19	5.82
-1	-1	6	4.83	5.92	2.38
0	-1	6	12.15	14.54	3.19
2	-1	6	10.32	8.62	2.24
-2	2	6	6.09	4.93	2.30
-1	8	6	7.37	8.97	2.22
-3	-2	7	8.51	6.92	2.20
3	-2	7	12.61	10.79	2.01
0	0	7	8.05	6.00	3.67
-1	1	7	16.25	19.31	2.94
0	2	7	15.44	17.45	2.16
-2	5	7	7.19	8.86	2.71
3	-1	8	3.01	0.93	2.72
-1	1	8	6.88	5.42	2.41
0	3	8	2.61	0.87	2.53
1	3	8	9.38	11.18	2.28
-5	-3	9	5.77	4.26	2.22
3	-3	9	7.53	5.96	2.50
3	-2	10	6.99	5.64	2.21
-1	3	10	6.99	8.38	2.20
-2	-2	11	7.16	8.98	3.05
-3	2	12	6.97	5.53	2.29
-4	1	13	8.73	7.19	2.28
-1	1	13	5.70	7.13	2.11
-3	0	14	7.32	5.88	2.18
-3	2	14	5.13	3.38	2.48
5	1	15	4.83	2.96	2.11
-2	-4	16	3.52	1.56	2.14
-3	-2	19	5.54	3.83	2.24

1ANALYSIS OF VARIANCE FOR KEITH1

		GGG	UGG	GUG	UUG	GGU	UGU	GUU	UUU	ALL					
	N	368	387	373	365	384	363	378	373	2991					
	V	70	69	86	79	87	73	83	87	80					
SIN THETA		0.00	- 0.18	- 0.23	- 0.26	- 0.29	- 0.31	- 0.33	- 0.36	- 0.38	- 0.40	- 0.43			
	N	316	343	278	335	245	280	415	275	254	250				
	V	145	93	76	67	62	67	63	62	55	50				
SQRT(F/FMAX)		0.00	- 0.13	- 0.15	- 0.16	- 0.18	- 0.20	- 0.22	- 0.24	- 0.27	- 0.33	- 1.00			
	N	366	395	189	368	326	282	216	252	311	286				
	V	61	70	77	81	85	85	99	80	82	83				
ABS (H)	0	1	2	3	4	5	6	7	8	9	10	11	12	13	REST
N	256	485	482	442	390	339	269	183	106	39	0	0	0	0	0
V	117	90	78	77	73	65	63	66	70	47	0	0	0	0	0
ABS (K)	0	1	2	3	4	5	6	7	8	9	10	11	12	13	REST
N	203	412	398	350	340	313	273	230	191	150	97	34	0	0	0
V	96	112	86	86	71	69	64	61	54	48	54	34	0	0	0
ABS (L)	0	1	2	3	4	5	6	7	8	9	10	11	12	13	REST
N	120	250	249	235	228	230	216	206	185	175	160	140	138	117	342
V	79	98	99	99	85	86	73	80	68	67	65	69	60	69	62