1 of/de 1

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS
STANDARD REFERENCE MATERIAL 1010a
(ANSI and ISO TEST CHART No. 2)

National Library
of Canada

Bibliothèque nationale
du Canada

Acquisitions and
Bibliographic Services Branch

Direction des acquisitions et
des services bibliographiques

395 Wellington Street
Ottawa, Ontario
K1A 0N4

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file - Votre référence

Our file - Notre référence

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

**A MODEL OF THE
INTERACTIONS OF AN ANTI-RETROVIRAL NUCLEOTIDE
WITH
HUMAN IMMUNODEFICIENCY VIRUS REVERSE TRANSCRIPTASE**

by

ANNA PURDY

B.Sc. (Hons.) Carleton University, 1991

***A Thesis Submitted to the Faculty of
Graduate Studies in partial fulfilment
of the requirements for the degree of
Master of Science***

***Department of Chemistry
Carleton University
Ottawa, Ontario
June, 1993.***

National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file   Votre référence

Our file   Notre référence

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-89869-0

Canada

Name _____

Dissertation Abstracts International is arranged by broad, general subject categories. Please select the one subject which most nearly describes the content of your dissertation. Enter the corresponding four-digit code in the spaces provided.

0 1 3 7

U·M·I

SUBJECT TERM

SUBJECT CODE

Subject Categories

THE HUMANITIES AND SOCIAL SCIENCES

COMMUNICATIONS AND THE ARTS

Architecture 0729
Art History 0377
Cinema 0900
Dance 0378
Fine Arts 0357
Information Science 0723
Journalism 0391
Library Science 0399
Mass Communications 0708
Music 0413
Speech Communication 0459
Theater 0465

EDUCATION

General 0515
Administration 0514
Adult and Continuing 0516
Agricultural 0517
Art 0273
Bilingual and Multicultural 0282
Business 0688
Community College 0275
Curriculum and Instruction 0727
Early Childhood 0518
Elementary 0524
Finance 0277
Guidance and Counseling 0519
Health 0680
Higher 0745
History of 0520
Home Economics 0528
Industrial 0521
Language and Literature 0279
Mathematics 0280
Music 0522
Philosophy of 0998
Physical 0523

Psychology 0525
Reading 0535
Religious 0527
Sciences 0714
Secondary 0533
Social Sciences 0534
Sociology of 0340
Special 0529
Teacher Training 0530
Technology 0710
Tests and Measurements 0288
Vocational 0747

LANGUAGE, LITERATURE AND LINGUISTICS

Language 0679
General 0289
Ancient 0290
Linguistics 0291
Modern 0291
Literature 0401
General 0294
Classical 0295
Comparative 0297
Medieval 0298
Modern 0316
African 0591
American 0305
Asian 0352
Canadian (English) 0355
Canadian (French) 0593
English 0311
Germanic 0312
Latin American 0315
Middle Eastern 0313
Romance 0314
Slavic and East European 0314

PHILOSOPHY, RELIGION AND THEOLOGY

Philosophy 0422
Religion 0318
General 0321
Biblical Studies 0319
Clergy 0320
History of 0322
Philosophy of 0469
Theology 0323

SOCIAL SCIENCES

American Studies 0323
Anthropology 0324
Archaeology 0326
Cultural 0327
Physical 0310
Business Administration 0272
General 0770
Accounting 0454
Banking 0338
Management 0385
Marketing 0501
Canadian Studies 0503
Economics 0505
General 0508
Agricultural 0509
Commerce-Business 0510
Finance 0511
History 0358
Labor 0366
Theory 0351
Folklore 0351
Geography 0351
Gerontology 0351
History 0578
General 0578

Ancient 0579
Medieval 0581
Modern 0582
Black 0328
African 0331
Asia, Australia and Oceania 0332
Canadian 0334
European 0335
Latin American 0336
Middle Eastern 0333
United States 0337
History of Science 0585
Law 0398
Political Science 0615
General 0616
International Law and 0617
Relations 0617
Public Administration 0814
Recreation 0452
Social Work 0626
Sociology 0627
General 0938
Criminology and Penology 0631
Demography 0628
Ethnic and Racial Studies 0629
Individual and Family 0630
Studies 0700
Industrial and Labor 0344
Relations 0709
Public and Social Welfare 0999
Social Structure and 0453
Development 0344
Theory and Methods 0709
Transportation 0999
Urban and Regional Planning 0453
Women's Studies 0453

THE SCIENCES AND ENGINEERING

BIOLOGICAL SCIENCES

Agriculture 0473
General 0285
Agronomy 0475
Animal Culture and 0476
Nutrition 0359
Animal Pathology 0478
Food Science and 0479
Technology 0480
Forestry and Wildlife 0817
Plant Culture 0777
Plant Pathology 0306
Plant Physiology 0287
Range Management 0746
Wood Technology 0306
Biology 0287
General 0308
Anatomy 0309
Biostatistics 0379
Botany 0329
Cell 0353
Ecology 0369
Entomology 0793
Genetics 0410
Limnology 0307
Microbiology 0317
Molecular 0416
Neuroscience 0433
Oceanography 0821
Physiology 0778
Radiation 0474
Veterinary Science 0786
Zoology 0760
Biophysics 0760
General 0425
Medical 0996

EARTH SCIENCES

Biogeochemistry 0425
Geochemistry 0996

Geodesy 0370
Geology 0372
Geophysics 0373
Hydrology 0388
Mineralogy 0411
Paleobotany 0345
Paleoecology 0426
Paleontology 0418
Paleozoology 0985
Palynology 0427
Physical Geography 0368
Physical Oceanography 0415

HEALTH AND ENVIRONMENTAL SCIENCES

Environmental Sciences 0768
Health Sciences 0566
General 0300
Audiology 0992
Chemotherapy 0567
Dentistry 0350
Education 0769
Hospital Management 0758
Human Development 0982
Immunology 0564
Medicine and Surgery 0347
Mental Health 0569
Nursing 0570
Nutrition 0380
Obstetrics and Gynecology 0354
Occupational Health and 0381
Therapy 0571
Ophthalmology 0419
Pathology 0572
Pharmacology 0382
Pharmacy 0573
Physical Therapy 0574
Public Health 0575
Radiology 0575
Recreation 0575

Speech Pathology 0460
Toxicology 0383
Home Economics 0386

PHYSICAL SCIENCES

Pure Sciences 0485
Chemistry 0749
General 0486
Agricultural 0487
Analytical 0488
Biochemistry 0490
Inorganic 0491
Nuclear 0494
Organic 0495
Pharmaceutical 0754
Physical 0405
Polymer 0605
Radiation 0986
Mathematics 0606
Physics 0608
General 0748
Acoustics 0607
Astronomy and 0798
Astrophysics 0759
Atmospheric Science 0609
Atomic 0610
Electronics and Electricity 0752
Elementary Particles and 0756
High Energy 0611
Fluid and Plasma 0463
Molecular 0346
Nuclear 0984
Optics 0346
Radiation 0984
Solid State 0346
Statistics 0984
Applied Sciences 0346
Applied Mechanics 0984
Computer Science 0984

Engineering 0537
General 0538
Aerospace 0539
Agricultural 0540
Automotive 0541
Biomedical 0542
Chemical 0543
Civil 0544
Electronics and Electrical 0348
Heat and Thermodynamics 0545
Hydraulic 0546
Industrial 0547
Marine 0794
Materials Science 0548
Mechanical 0743
Metallurgy 0551
Mining 0552
Nuclear 0549
Packaging 0765
Petroleum 0554
Sanitary and Municipal 0790
System Science 0428
Geotechnology 0796
Operations Research 0795
Plastics Technology 0994
Textile Technology 0994

PSYCHOLOGY

General 0621
Behavioral 0384
Clinical 0622
Developmental 0620
Experimental 0623
Industrial 0624
Personality 0625
Physiological 0989
Psychobiology 0349
Psychometrics 0632
Social 0451

**The undersigned recommend to the
Faculty of Graduate studies acceptance of the Thesis
"A Model of the Interactions of an Anti-Retroviral Nucleotide with
Human Immunodeficiency Virus Reverse Transcriptase."
Submitted by Anna Purdy
in partial fulfilment of the requirements for the
degree of Master of Science**

A handwritten signature in black ink, appearing to read "G.W. Buchanan", written over a horizontal line.

**Dr. G.W. Buchanan
Supervisor**

A handwritten signature in black ink, appearing to read "K. Houghton for RAB", written over a horizontal line.

**Chairman
Department of Chemistry**

ABSTRACT

The tetrapeptide isoleucine-lysine-lysine-lysine (IKKK) was synthesized using the fluorenylmethoxycarbonyl (Fmoc) polyacrilamide method of solid phase peptide synthesis. This sequence is highly conserved in retroviral reverse transcriptases and represents a possible nucleotide triphosphate binding sequence on human immunodeficiency virus reverse transcriptase. The interactions of the metabolically active nucleotide analogue, 2',3'-dideoxyadenosine 5'-triphosphate (ddATP), with this sequence as well as with di-L-lysine was determined via ^{13}C nuclear magnetic resonance (NMR) relaxation times (T_1) and 2 dimensional nuclear Overhauser NMR spectroscopy (NOESY). As well, molecular modelling was used to determine those interactions of the lowest energy. The solution conformational equilibrium of ddATP was determined by analysis of $^3J_{\text{H-H}}$ coupling constants.

Experimental T_1 data suggest there is both an electrostatic and hydrophobic interaction of ddATP with di-L-lysine and IKKK. Decreases in spin-lattice relaxation time upon mixing of the two peptides with ddATP were seen for carbons closest to the positively charged amino groups of the lysine side chains. This indicates the negatively charged phosphates of ddATP are interacting with these positively charged amino groups, as predicted by molecular modelling. The changes in spin-lattice relaxation times for two of the carbons on the side chain of isoleucine were extremely large, indicating a hydrophobic association with ddATP in an aqueous environment. A NOESY spectrum of di-L-lysine with ddATP reveals a connectivity between the ribose of ddATP and the side chains of the lysine residues.

ACKNOWLEDGEMENTS

There are so many wonderful people I have had the opportunity to work with, I only hope I am so lucky in the future.

I would like to thank Dr. G.W. Buchanan for his thoughtful guidance and ideas during this project.

I would like to thank Dr. C.S. Tsai who performed all molecular modelling calculations.

Thanks to Keith Bourque for running all 400MHz spectra, and for making me laugh.

Thanks to both Tony O'Neil and Fred Cassalman, who put up with my ongoing HPLC nightmares.

Special thanks to Karl Diedrich and Rob Kehle in chemistry stores without whom I would never have been able to complete this thesis.

Thanks to Marilyn, Barb and Kim for helping me out of all those administrative jams.

Thanks to Marcin Kolbuszewski, without his help, I probably would still be unsure how to turn on a computer.

Thank you to all those people with whom I worked in the lab for their helpful hints and insights, Abolghassem Mohigmi, Mark Reynolds, Marielle Gerzain, Kane Denike, and Alex Dreiga.

To Beth "Ruskonovitch" Russell, Aleixo "All I heard was Dead tunes" Muise, and Donna "Slide me some southern" Douglas, thanks for being the remarkable friends you are.

Finally, I would like to thank my parents for their encouragement.

TABLE OF CONTENTS

TITLE	PAGE
ABSTRACT	i
ACKNOWLEDGEMENTS	ii
TABLE OF CONTENTS	iii
LIST OF TABLES	v
LIST OF FIGURES	vi
LIST OF ABBREVIATIONS	viii
 INTRODUCTION	 1
 I. Human Immunodeficiency Virus Reverse Transcriptase	 1
II. Conformational Analysis of Nucleotides	14
III. Nuclear Magnetic Resonance Spectroscopy	19
A.Chemical Shift	20
B.Spin-Spin Coupling	21
C.Fourier Transform NMR	23
D.Spin-Lattice Relaxation Time	25
E.Nuclear Overhauser Enhancement Spectroscopy	30
IV. Peptide Synthesis	34
V. Molecular Modelling	38
VI. Experimental Design	39
 RESULTS AND DISCUSSION	 40
 I. Peptide Synthesis	 40
II. Spectral Analysis	43
III. Molecular Modelling	59
IV. Nucleotide Conformation	62

CONCLUSIONS	68
FUTURE DIRECTIONS	71
EXPERIMENTAL	72
I. Peptide Synthesis	72
II. NMR	75
REFERENCES	77

LIST OF TABLES

TABLE	TITLE	PAGE
1	α -H Chemical Shifts and ^3J ^1H - ^1H Coupling Constants of IKKK	47
2	^{13}C T_1 Values for Di-L-Lysine in D_2O	52
3	^{13}C T_1 Values for IKKK in D_2O	53
4	$^2J_{HH}$, $^3J_{HH}$ and $^3J_{HP}$ Coupling Constants and Chemical Shifts for ddATP.	63
5	Comparison of the conformation of ddATP with related compounds	64
6	Population of γ rotamers of ddATP	64

LIST OF FIGURES

FIGURE	TITLE	PAGE
1	Structure of Human Immunodeficiency Virus	2
2	Mechanism of DNA Chain Termination by 2'3'-dideoxynucleotides	5
3	Metabolic Pathway for the Phosphorylation of ddI and ddA to ddATP <i>in vivo</i>	8
4	The Polymerase Domain of p66	12
5	Conformational parameters of Nucleosides	15
6	Definition of Positive and Negative Substituents	16
7	Mechanism of Transferring Spin Information	22
8	Variance of the Coupling Constant with the Dihedral Angle α	22
9	Pulsed Fourier Transform Techniques	24
10	Effect of Correlation time on Spin-lattice relaxation Time (T_1)	25
11	Orientation of the Dipolar Interaction Relative to the Magnetic Field	26
12	Relaxation of the Magnetization Vector M_z via Spin-Lattice Relaxation	28
13	Inversion Recovery Pulse Sequence	29
14	Spin States and Transition Probabilities (W_i) for a coupled Two Spin System.	30
15	Protecting Group Combinations for the Merrifield Synthesis	35
16	Protecting Group Combinations for the Fmoc Method of Peptide Synthesis	36
17	Base Catalyzed Elimination of the Fmoc group	36
18	Synthetic Sequence	41
19	HPLC trace of Crude Tetrapeptide	40
20	Numbering Scheme for Di-L-lysine, ddATP, and IKKK	43
21	200 MHz ^1H NMR of Di-L-lysine in D_2O	44
22	^{13}C NMR of di-L-lysine in D_2O	45
23	FABMS of di-L-lysine	45
24	C_α region of the ^1H 400 MHz NMR spectrum of IKKK in D_2O	46

LIST OF FIGURES CON'T

FIGURE	TITLE	PAGE
25	¹ H NMR spectrum of IKKK in D ₂ O	46
26	COSY spectrum of IKKK in D ₂ O	48
27	HETCOR of IKKK in D ₂ O	50
28	¹³ C Spectrum of IKKK in D ₂ O	50
28a	Amide Region of ¹³ C NMR spectrum of IKKK in D ₂ O	51
29	FABMS spectrum of IKKK	51
30	T ₁ Spectrum of di-L-lysine in D ₂ O	54
31	T ₁ Spectrum of di-L-lysine with ddATP in D ₂ O	54
32	T ₁ Spectrum of IKKK in D ₂ O	55
33	T ₁ Spectrum of IKKK with ddATP in D ₂ O	55
34	Changes in ¹³ C Spin-lattice Relaxation Times for di-L-lysine and IKKK in Solution with ddATP.	56
35	NOESY spectrum of di-L-lysine with ddATP	58
36	¹ H NMR of Decomposed and Unaltered ddATP	58
37	Lowest Energy Conformation of ddATP	59
38	Lowest Energy Conformation of IKKK	60
39	Brick Map depicting interaction of ddATP and IKKK with interaction energies lower than -26 kcal mol ⁻¹	61
40	Brick Map depicting interaction of ddATP and IKKK with interaction energies lower than -30 kcal mol ⁻¹	61
41	Brick Map depicting interaction of ddATP and IKKK with interaction energies lower than -35 kcal mol ⁻¹	61
42	¹ H NMR of ddATP in D ₂ O	65
43a-c	Experimental and Simulated ¹ H NMR of ddATP at 278K, 283K, 288K, 293K, and 299K	66 67
44	Principal Solution Conformation of ddATP	68
45	Amino Acid Sequence Comparison between HIV and 10 other Retroviral RTs	69
46	Bubbler Apparatus used in Peptide Synthesis	72

LIST OF ABBREVIATIONS

IKKK	Isoleucine-Lysine-Lysine-Lysine
FMOC	Fluoromethoxycarbonyl
ddATP	2',3'-dideoxyadenosine 5'-triphosphate
NMR	Nuclear Magnetic Resonance
T₁	Spin-lattice Relaxation Time
NOESY	Nuclear Overhauser Enhancement Spectroscopy
HIV	Human Immunodeficiency Virus
AIDS	Acquired Immunodeficiency Syndrome
RT	Reverse Transcriptase
AZT	3'-Azido-2',3'-dideoxythymidine
ddI	2',3'-dideoxyinosine
LTR	Long Terminal Repeat
dNTP	2'-deoxynucleotide 5'-triphosphate
ddAMP	2',3'-dideoxyadenosine 5'-monophosphate
ddC	2',3'-dideoxycytidine
ddT	2',3'-dideoxythymidine
MuLV	Moloney Murine Leukemia Virus
PLP	Pyridoxyl 5-Phosphate
Pol	Polymerase
Hg-UTP	5-mercuriuridine 5'-triphosphate
P	Pseudorotational Phase Angle
FT-NMR	Fourier Transform Nuclear Magnetic Resonance
FID	Free Induction Decay
NOE	Nuclear Overhauser Enhancement
T_c	Molecular Correlation Time
HPLC	High Pressure Liquid Chromatography
FAB-MS	Fast Atom Bombardment Mass Spectrometry
COSY	Correlation Spectroscopy
HETCOR	Heteronuclear Correlation Spectroscopy
DMF	Dimethylformamide

INTRODUCTION

Human Immunodeficiency Virus Reverse Transcriptase

Human immunodeficiency virus (HIV) is a pathogenic retrovirus, and the causative agent of acquired immunodeficiency syndrome (AIDS) ¹⁻³. Retroviruses are RNA containing viruses that replicate through a DNA intermediate by virtue of a viral coded RNA-dependant DNA polymerase, called reverse transcriptase^{4,5}. The World Health Organization estimates between 10 and 12 million adults, and 1 million children have been infected with the HIV virus worldwide⁶. It is estimated that by the year 2000 at least 38 million adults and 10 million children will be infected with HIV. AIDS has become a global medical emergency, and the scientific community has responded to this challenge by making progress at a remarkable rate in the characterization of HIV. The life cycle and complex mechanisms of gene regulation are now more fully understood for HIV than that for any other retrovirus.

Variability is the hallmark of HIV, and thus HIV cannot be treated in the traditional manner of elimination through vaccinations⁷. Studies indicate that the gene encoding the envelope glycoprotein (env) varies at the rate of 1% per year⁶. Instead, those infected with the virus are being treated through anti-retroviral drug therapy directed against an essential enzyme in the viral life cycle. The viral enzyme reverse transcriptase is the target of the most commonly used drugs which are nucleoside analogues such as, 3'-azido-2',3'-dideoxythymidine (AZT) and

dideoxyinosine (ddI). These drugs have furnished dramatic results in clinical trials, but are invariably toxic when used on a continuing basis⁸.

Electron microscopy has revealed a cylindrical core in HIV particles, characteristic of animal lentiviruses (Fig.1)⁷.

Fig.1 Structure of Human Immunodeficiency Virus⁷

Like other retroviruses, HIV has two copies of 35S single stranded genomic RNA with the polarity of mRNA. The genomic RNA is thought to exist as a ribonucleoprotein complex containing reverse transcriptase (p61/p55), endonuclease (p32) and retroviral RNA binding Gag protein (p9). The viral core is composed mainly of p24, while the myristylated Gag protein (p17) forms the outer shell of

the capsid. The amino terminal of this protein is inserted into the lipid membrane envelope which is derived from the plasma membrane of the infected cell. The envelope glycoprotein (gp120), and transmembrane glycoprotein (gp41) exist as a noncovalent complex on the surface of the envelope⁶. HIV has at least 9 known genes and is the most complex retrovirus studied thus far⁷. These nine genes are flanked at each end by sequences called long terminal repeats (LTR). The LTRs contain several characteristic regulatory elements, a TATA box, which is essential for core promoter function, upstream enhancer regions to which cellular transcription factors bind to activate HIV transcription, a negative regulatory element, a trans-acting responsive sequence, and a polyadenylation signal sequence^{8,9,10}. The LTRs are necessary to initiate the expression of viral genes, via a complex process requiring cellular polymerases and transcription factors.

The first step in cellular HIV infection is the binding of gp120, the envelope glycoprotein to its cellular receptor, which in the case of HIV is CD4, a protein molecule found on the surface of T4 cells¹¹. This binding is thought to produce a conformational change in the HIV envelope unmasking a fusion domain in the transmembrane protein gp41 which is responsible for cell-virion fusion followed by HIV entry into the cell¹². After cell entry, the virus is thought to lose its envelope and release its RNA genome into the cytoplasm as a ribonucleoprotein complex composed of the genomic RNA, gag, RT, and endonuclease⁷. After uncoating, HIV RT uses the viral RNA as a template for proviral DNA synthesis.

A biproduct-bireactant mechanism, analogous to the mechanism described for other well characterized DNA polymerases has been suggested for DNA synthesis by HIV RT¹³. In this mechanism, binding of the reactants is ordered with the template-primer binding first followed by the binding of the 2'-deoxynucleoside 5'-triphosphate having Watson-Crick complementarity to the next appropriate template residue. The resulting ternary complex is proposed to undergo a conformational change to form the catalytically competent complex. Incorporation of the bound dNTP into the nascent chain then occurs via an S_N2 displacement of pyrophosphate by the 3' hydroxyl at the primer terminus¹⁴. The presence of homopolymeric template-primers does not result in selective binding of the complementary 2'dNTP indicating that Watson-Crick base pairing is not involved in the initial binding reaction¹⁵. This is in accord with a general observation for DNA polymerases in which the low energy step of Watson-Crick base pairing does not contribute to the assembly of a catalytically competent complex¹⁶. HIV RT first generates a DNA copy of the viral RNA using cellular tRNA as a primer⁷. During the synthesis of the first strand of DNA, the RNA in the RNA-DNA duplex is degraded by the ribonuclease activity of RT which is necessary for the production of double strand DNA. One segment of the RNA is resistant to this hydrolysis and serves as the primer for the synthesis of the second strand of DNA by HIV RT¹⁷. The newly synthesized proviral double strand DNA, by poorly understood mechanisms, migrates to the nucleus of the host cell, where it integrates

into the host chromosome⁷. HIV RT is extremely error prone (5-10 per round of replication) and thus is probably responsible for the hypermutability of HIV.

Reverse transcriptase represents an excellent target for antiviral design as there are few examples of cellular eukaryotic reverse transcription, and RT is absolutely necessary for viral replication¹⁷. Additionally RT is the only nonorganelar polymerase which operates in the cytoplasm of HIV infected cells, and thus specific targeting of cytoplasmic polymerization has been undertaken. The most common RT inhibitors are nucleoside analogues, which lack a 3' hydroxyl group. It is reported that 2',3'-dideoxynucleosides as their triphosphates inhibit HIV RT¹⁸⁻²¹. Incorporation of an analogue into a growing DNA chain would render the next 3'-5' phosphodiester linkage impossible, and thus result in DNA chain termination. (Fig.2).

Fig.2 Mechanism of DNA Chain Termination by 2',3'-dideoxynucleotides.

HIV RT is composed of two subunits, a 66 kDa polypeptide and a 51kDa polypeptide²². The amino termini of the two subunits are identical, and the smaller 51kDa subunit is derived from proteolytic processing of the p66 subunit by HIV protease. The polymerization domain maps to the amino terminus of the p66 chain. The ribonuclease domain, which is lacking from the p51 chain maps to the carboxyl terminus of the p66 chain²³⁻²⁵. Partial proteolysis of the p66/p51 dimer produces an active p66/p30/p29 "complex", indicating the p66 chain is providing the polymerization active site²⁶. Furthermore, cross linking studies of p66/p51 with substrate analogues have shown only the p66 subunit is labelled²⁷.

A broad family of nucleosides with a 2',3'-dideoxyribose moiety can restrict the infectivity and replication of divergent HIV strains *in vitro*. These nucleosides were originally synthesized in the 1960s²⁸. Among those which have been administered to HIV⁺ patients are AZT, ddC, ddA, and ddI. These nucleoside analogues are capable of suppression of HIV replication in monocytes, macrophages and T-cells^{14,29,30}. AZT was first administered to patients with severe HIV infection in July of 1985⁸. These individuals underwent immunological, viral and clinical improvement during the AZT therapy. After a placebo-controlled trial of AZT therapy confirmed these results as well as showing a reduced mortality, AZT was approved as a prescription drug for patients who have had episodes of *Pneumocystis carinii* pneumonia or whose CD4⁺ cell count was below 200/mL. AZT is not without side effects, which include bone marrow suppression, myalgias,

seizures, and liver function abnormalities^{8,31,32}. ddA and its deamination product ddI were identified in 1985 as potent HIV inhibitors in vitro¹⁸. In cell culture, ddI appears to be less toxic than AZT¹⁸, which is probably a result of greater specificity for HIV RT. ddI was first administered to individuals with AIDS in July of 1988. Improvement in immunological function and evidence of a decrease in viral load was seen in these patients⁸. Doses of up to 9.6mg/Kg of body weight have been tolerated in some patients for 2 years without toxicity⁸.

The unphosphorylated precursors to metabolically active nucleotides, dda, ddI, and AZT enter both infected and noninfected cells through passive diffusion²⁹. The phosphorylation of AZT is thought to be catalyzed by cytosolic thymidine kinase, which is one of the key enzymes in the salvage pathway for pyrimidine deoxyribonucleotide synthesis³⁴. ddA can be directly phosphorylated to 2'3'-dideoxyadenosine 5'-monophosphate (ddAMP) in the cytoplasm by deoxycytidine kinase or adenosine kinase³⁵. However most ddA is deaminated by adenosine deaminase³⁵, and converted to ddI which is anabolized in human lymphoid cells by an initial step of phosphorylation and subsequent amination to dideoxy-AMP via adenylosuccinate synthetase/lyase³⁵. Additionally, in the presence of MgCl₂, KCl, and inosine-5'-monophosphate as a phosphate donor, cytosolic 5'-nucleosidase catalyzes the phosphorylation of ddI³⁶. Following the initial phosphorylation, nucleoside analogues are further anabolized by cellular nucleoside monophosphate kinase and nucleoside diphosphate kinase to the metabolically active

Activity profiles *in vivo* and *in vitro* of seemingly similar dideoxynucleosides are very different, showing differences in metabolic activity, solubility and reversal of antiviral activity by normal nucleoside counterparts³⁷. Reasons for the different *in vivo* and *in vitro* activities of dideoxynucleosides could either be a result of efficiency of cellular anabolism of the analogue to its metabolically active triphosphorylated form, or a result of a higher or lower affinity for HIV RT. Closely related 2'3'-dideoxynucleosides can differ greatly in anti-HIV activity

levels. AZT (EC_{50} =.002-.009 μ M), 2'3'-dideoxycytidine (ddC) (EC_{50} =.011 μ M), ddA (EC_{50} =.011 μ M) and 2'3'-dideoxythymidine (ddT) (EC_{50} =.17 μ M) are quite active but 2'3'-dideoxyuridine (ddU) (EC_{50} =96.8 μ M) and 2'3'-dideoxy-5-ethyluridine (EC_{50} =4.9 μ M) are much less active³⁸. Trends indicate the addition of a 3'-azido group increases the activity of thymidine and uridine related compounds, for example AZT (EC_{50} =.002-.009), and 3'-azido-2'3'-dideoxyuridine (EC_{50} =.018-.46 μ M). The nucleoside analogue 3'-(propyl-2-ene)-2'3'-dideoxyuridine has a very low activity (EC_{50} =100 μ M). This suggests the size and polarity of the 3' substituent are important³⁸. Correlation of activity with molecular conformations in solution can be used to determine any trends relating solution conformation with activity.

Several dideoxynucleoside triphosphates have a higher affinity for HIV RT than for the cellular DNA polymerase α . However, these dideoxynucleosides show a similar affinity for the cellular DNA polymerases β and γ (mitochondrial polymerase)³⁹. This may explain the development of mitochondrial myopathy in individuals receiving long term AZT therapy⁴⁰. AZT-TP inhibits DNA polymerase γ by competition with dTTP, and not by incorporation into the DNA chain, causing chain termination as it does for HIV RT⁴¹. Unlike AZT, the antiretroviral activity of ddI and ddA is not reversed by any of the normal 2'deoxy nucleosides²⁹. This is a startling result, and data are yet inconclusive as to the reason for this discrepancy. This could be a result of more efficient cellular phosphorylation of

ddA and ddI or a greater affinity of ddA and ddI for HIV RT.

In some patients receiving AZT therapy for more than 6 months, strains of HIV which are resistant to AZT have been isolated⁴². Nucleotide sequence analysis of the region coding for reverse transcriptase has revealed common amino acid substitutions for the AZT resistant variants, (Asp⁶⁷ -> Asn, Lys⁷⁰ ->Arg, Thr²¹⁵ ->Phe or Tyr) and in some isolates (Lys²¹⁹ ->Gln)⁴⁴. Partially resistant RT isolates had a combination of these four changes⁴². Interestingly, these AZT insensitive isolates show no difference in sensitivity to AZT-TP, or other 2',3'-dideoxynucleosides including ddI and ddC⁴³. Site directed mutagenesis of reverse transcriptase can render the enzyme insensitive to AZT-TP, but paradoxically recombinant infectious virus carrying the gene for the insensitive mutant showed hypersensitivity to AZT in culture^{44,45}.

The design of drugs specific for HIV RT has been impeded by the lack of a high resolution crystal structure for HIV RT. Nucleoside analogues are invariably toxic and HIV RT, over a continuing time frame develops resistance to them. Experimental data suggest that in some cases ddI and AZT show a greater affinity for HIV RT than normal 2'-deoxynucleosides²⁹. Detailed information about enzyme substrate interactions at the active site of HIV RT, as well as the nucleoside triphosphate binding site on HIV RT is yet unknown. Information about enzyme substrate interactions would certainly lead to the design and use of specific RT inhibitors with limited cellular toxicity.

It is likely all retroviral RTs have diverged from a common ancestor⁴⁶. The high similarity of RT amino acid sequences across a wide range of RT's suggests a considerable degree of structural relatedness. The amino acid sequences of ten retroviral reverse transcriptases were aligned in order to determine those residues which are highly conserved, and thus important for enzymatic function⁴⁷. Sequence comparisons in the polymerase domains reveal significant sequence conservation from residue Q23 through W266. It is thought the nucleotide binding site of RT lies within this region. Studies on Moloney murine leukaemia virus (MuLV) RT have shown that a lysine conserved in all ten retroviral RTs is specifically protected from chemical modification by its substrate⁴⁸. For HIV as well as MuLV the essential part of RT for polymerization is located in the N terminus. MuLV RT was labelled with pyridoxal 5'-phosphate, (PLP) which is a substrate binding site directed reagent for DNA polymerases⁴⁸. Treatment of MuLV RT with PLP results in a loss of RNA directed DNA polymerase activity, but has no effect on ribonuclease activity. When in solution with template-primer and complementary substrate, inhibition of RNA directed DNA polymerase activity by PLP of RT is repressed indicating PLP does in fact block the active site of the enzyme. Comparative tryptic peptide mapping of RT modified in the presence and absence of substrate by PLP indicated two peptides which were protected from pyridoxylation in the presence of natural substrate. The targets of pyridoxylation were found to be two distinct lysine residues, as determined by amino acid analysis

and sequencing. These results indicate two lysine residues are involved in binding dNTP to MuLV RT⁴⁸. One of these lysine residues corresponds to lysine in position 65 in HIV RT when the amino acid sequences are aligned⁴⁷. AZT resistant isolates of HIV RT show a point mutation at this lysine residue, again implicating a role in nucleoside binding¹⁷.

Recently, the 3 dimensional crystal structure of HIV RT at 3.5^oA resolution has been reported⁴⁹. The p66-p51 heterodimer is asymmetric in nature. Surprisingly, the two polymerase (pol) domains have very different conformations, despite having the same amino acid sequence⁴⁹. There is a cleft between the RNAase Hydrolysis and polymerase domains of p66 sufficient to accommodate an A-form RNA-DNA hybrid⁴⁹. The polymerase domain is folded into 4 different subdomains. The 3D model of its conformation resembles a right hand, and thus its subdomains have been named "fingers", "palm" and "thumb"(Fig.4)⁴⁹.

Fig.4 The Polymerase Domain of p66.

Residues implicated in rendering RT resistant to AZT and ddI (Met⁴¹, Asp⁶⁷, Asp⁶⁹, Lys⁷⁰, Leu⁷⁴, Thr²¹⁵, and Lys²¹⁹)^{17,50} are in the polymerase active site, and cluster in two regions, the former on the "fingers" subdomain, and the latter on the "palm". Another group has reported the crystallization of RT complex with DNA at a resolution of 7⁰A⁵¹. In this study 5-mercuriuridine 5'-triphosphate (Hg-UTP) was used for phasing purposes. It was found that Hg-UTP binds to one site in the DNA containing crystals that is different from the sites to which nonnucleoside mecurial derivatives bind. This site is located to the end of the DNA opposite to that which is in contact with the RNase domain⁵¹. This is separated from the RNase domain by a distance of about 15 nucleotides consistent with biochemical measurement of the separation between RNase and polymerization activities⁵². Thus it seems likely that the Hg-UTP is interacting with the triphosphate binding site of HIV RT. The position at which it is interacting corresponds to the "fingers" domain in the structure previously presented, which contains the polymerase active site and within this active site, the sequence Ile⁶³-lys⁶⁴-lys⁶⁵-lys⁶⁶. Mutation of ile⁶³ to serine and lys⁶⁴ to arginine decreases HIV RT activity to 50% of wildtype RT⁶⁰. Mutation of lys⁶⁶ to arginine results in no difference in polymerase activity⁵². However, mutation of lys⁶⁵ to arginine, an evolutionary conservative mutation eradicates the polymerase function of HIV RT but does not effect the RNase capability, suggesting an important role for this residue in the polymerization function⁵⁰.

Conformational Analysis of Nucleotides via ^1H - ^1H coupling Constants.

The conformations of nucleosides are described by three essential parameters, the geometry of the glycosyl link, the rotation about the exocyclic C4'-C5' bond, and the puckering of the sugar ring⁵³(Fig.5). The torsion angles $\chi(\text{C2-N1-C1'-O4'})$ and $\gamma(\text{C3'-C4'-C5'-O5'})$ relate the orientations of the base and the 5'-hydroxyl group relative to the sugar ring. The puckering of the five-membered sugar ring is a continuous displacement of one (envelope) or two (twist) atoms out of the plane of the others. This displacement can be described by the pseudorotational phase angle P , which depends on all intra-ring torsion angles, and on the maximum torsion angle, ν_{max} ⁵³ (Fig.5).

Solid state studies indicate the preferred conformations of active nucleoside analogue sugar rings is the 3'exo conformation (Fig.5)³⁸. However, solid state studies show only one conformation of the nucleoside.

In solution, the furanose ring can be described as an equilibrium between N ($P=0^\circ\pm90^\circ$) and S ($P=180^\circ\pm90^\circ$) type conformers, each occurring in a confined pseudorotational range⁵⁴. It is possible to determine the nature of this equilibrium in solution through the use of proton-proton spin-spin coupling constants which give a good indication of proton-proton torsion angles⁵⁵. These torsion angles for a five membered ring are related to the endocyclic torsion angles which are correlated via the laws of pseudorotation. The vicinal couplings can be expressed as a function of the pseudorotation parameters, P and ν_{max} ⁵⁵. The coupling

Fig.5 Conformational Parameters of Nucleosides

Sugar Ring Pucker

C_2' endo- C_3' -exo Twist
 $P=180^\circ$
 S Type Conformer

C_3' endo- C_2' -exo Twist
 $P=0^\circ$
 N Type Conformer

C_3' -exo Envelope
 $P=198^\circ$
 S Type Conformer

Torsion Angle γ Describing the Orientation of the 5'OR Group

Trans

Gauche γ^+ Gauche γ^-

The Glycosyl Link

Syn

Anti

constants are time averaged values of the couplings of the individual conformers of the two state equilibrium. Thus, the coupling constants can be expressed as a function of the five conformational parameters involved, i.e. the mole fraction of one conformer, the phase angle for each conformer, and the puckering amplitude for each conformer⁵⁵.

A generalization of the classical Karplus equation which takes into account the influence of an electronegative substituent on $^3J_{HH}$ was introduced by Altona and coworkers.

$$^3J_{HH} = P_1 \cos^2 \theta_{HH} + P_2 \cos \theta_{HH} + P_3 + \Sigma \{ P_4 + P_5 \cos^2 (\zeta_i \theta_{HH} + P_6 \Delta X_i) \} \quad (\text{Eq. 1.1})$$

where θ_{HH} is the proton-proton torsion angle, ΔX_i is the difference in electronegativity between the substituent and hydrogen on the scale of Huggins⁵⁶ and ζ_i is a substituent orientation parameter (Fig.6).

Fig.6 Definition of Positive and Negative Substituents

With the concept of pseudorotation the endocyclic torsion angles in a five-membered ring can be approximated by

$$\theta_j = \theta_m \cos (P + 4\pi j/5) \quad j=1,2,3,4 \quad (\text{Eq.1.2})$$

assuming a linear relationship between proton-proton torsion angles and endocyclic torsion angles, θ_{HH} can be expressed as

$$\theta_{HH} = a\theta_m \cos(P+4\pi j/5) + b \quad (\text{Eq.1.3})$$

The parameters a and b for 2',3'-dideoxyribonucleosides have been developed using crystal structures of ddU, ddT, ddA and ddC⁵⁷.

In the case of a two state equilibrium, $I \rightleftharpoons II$

$$J_{\text{exp}} = (1-X_{II})J_I + X_{II}J_{II} \quad (\text{Eq.1.4})$$

where X_{II} is the mole fraction of conformer II, and J_I and J_{II} are the coupling constants of the pure conformers. Combination of equations 1.1-1.4 gives ${}^3J_{HH}$ as a function of the conformational parameters: ${}^3J_{HH}=f(P_I, \theta_{mI}, P_{II}, \theta_{mII}, X_{II})$ (Eq.1.5)

When coupling constants are measured under n different conditions (i.e. temperature, pH etc), equation 1.5 can be extended to a set of n equations:

$${}^3J_{HH} = f(P_i, \theta_{mI}, P_{II}, \theta_{mII}, X_{II}) \quad i=1 \dots n$$

The computer program PSEUROT⁵⁸ calculates the best fit of the conformational parameters to n sets of experimental couplings.

The conformation for ddG, ddI, ddA, AZT and ddC in solution have been determined by NMR using this method⁵⁷, however, there has been no conformational analysis of the metabolically active triphosphorylated forms of these

drugs. In the case of the dideoxynucleosides for all but AZT, the N conformer predominated in solution. In the case of AZT, the N and S conformers were equally populated in solution. The purine dideoxynucleosides, ddA, ddI and ddG adopt a twist C_2' -*exo*- C_3' -*endo* form, while ddC, a pyrimidine nucleoside adopts a C_3' -*endo* envelope form³⁹. The unphosphorylated natural substrate, 2'-deoxyadenosine shows a preference for the S-type conformation ($P_3=162$, $X_N=.26$)³⁹ and adopts a C2'-*endo*/C3'-*exo* twist type conformation in solution.

The X-ray structure of ddA depicts the dideoxynucleoside in a S type C_3' -*exo* conformation³⁸ which is the conformation active anti-HIV nucleosides show a preference for in the solid state. This conformation would place C5' in an axial position, which in turn affects the position of the 5' triphosphate group in relation to the base. The 5' position is the position at which the analogue becomes phosphorylated and incorporated into the growing DNA chain³⁸. The torsion angle, γ (C_3' - C_4' - C_5' - O_5') describes the orientation of the 5'-phospho group relative to the sugar ring (Fig.5). The solution conformation of the 5'-CH₂OR-moiety relative to the furanose ring may compromise three types of rotamers denoted γ' , γ , and γ'' (Fig.5)⁵³. Rapid interconversion between the three conformers yields weighted time-averaged NMR couplings that are related to the of the individual conformers and their relative populations by:

$$^3J_{4,5'} = X_{\gamma'}J_{\gamma'}' + X_{\gamma}J_{\gamma}' + X_{\gamma''}J_{\gamma}'' \quad (\text{Eq.1.6})^{53}$$

$$^3J_{4,5''} = X_{\gamma'}J_{\gamma'}'' + X_{\gamma}J_{\gamma}'' + X_{\gamma''}J_{\gamma}'' \quad (\text{Eq.1.7})^{53}$$

Where X is the relative population of each rotamer and $\sum X_i = 1$. The values J' and J'' for each rotamer were determined by analysis of x-ray crystallographic structures, which yielded the proton-proton torsion angle, in combination with equation 1.1⁵³. Thus the population of each rotamer and the position of the 5' substituent can be described.

Nuclear Magnetic Resonance Spectroscopy

Nuclear spin (I) is a fundamental property of the atomic nucleus. The nuclear spin has values of $0, 1/2, 1, 1 1/2, \dots$ depending on the mass number and atomic number of the nucleus. The nuclear magnetic moment (μ) is directly proportional to the spin,

$$\mu = \gamma \hbar / 2\pi \quad (\text{Eq. 2.1})^{60}$$

where γ , the proportionality constant, is called the magnetogyric ratio and is a constant for each particular nucleus. When a magnetic field is applied, the nuclear moments orient themselves with only certain orientations as we are considering a quantum mechanical system. A nucleus of spin I has $2I+1$ possible orientations given by the magnetic quantum number m_I , which has values of $-I, -I+1, I-1, I$. The interaction energy is proportional to the nuclear moment and the applied field

$$E = -\gamma \hbar m_I B / 2\pi \quad (\text{Eq. 2.2})^{60}$$

where B is the applied field.

The selection rule for NMR transitions is that m_I can only change by one unit, thus the transition energy is given by

$$\Delta E = \gamma \hbar B / 2\pi \quad (\text{Eq.2.3})^{60}$$

For detection of this transition energy, radiation given by $E = h\nu$ must be applied. Combining this with equation 2.3 gives the fundamental resonance condition for all NMR experiments

$$\nu = \gamma B / 2\pi \quad (\text{Eq.2.4})^{60}$$

i.e. when a nucleus of magnetogyric ratio γ is placed in a magnetic field B , the resonant condition is satisfied when the frequency of the applied radiation, ν is given by equation 2.4.

In the case of a nucleus with spin $= 1/2$, there are two possible orientations of the nuclear spin. $m_I = +1/2$ is the lower energy more stable state, denoted by the symbol α , the upper state, $m_I = -1/2$ is given by the symbol β . There are two allowed transitions, $\alpha \rightarrow \beta$ corresponding to an adsorption of energy, and $\beta \rightarrow \alpha$ corresponding to an emission of energy. The Boltzman distribution of energies is maintained between these two states

$$N_\beta / N_\alpha = \exp(-h\nu / kT) = 1 - h\nu / kT \quad (\text{Eq.2.5})$$

and therefore $N_\beta < N_\alpha$ giving a net adsorption of energy when the resonance condition is satisfied, and an NMR signal.

Chemical Shift

The electrons within a molecule precess in the applied magnetic field

producing an electric current. This current produces a magnetic field at the nucleus which opposes the external field, shielding the nuclei from an applied magnetic field. Therefore the field at the nucleus is not equal to the applied external field, and it is necessary to increase the applied field over that of the isolated nucleus in order to obtain the resonant condition. The chemical shift is defined as the nuclear shielding divided by the external field, and is a function only of the nucleus and its environment.

$$\delta = [(B_{\text{reference}} - B_{\text{sample}})/B_{\text{reference}}] \times 10^6 \text{ ppm (Eq.3.1)}$$

Spin-Spin Coupling

The magnetic field at each proton is made up of contributions due to circulating electrons plus the effects of other nuclear magnets in the molecule, for example another proton, which can have two possible orientations in the magnetic field, and will produce either a small field, or a small opposing field at nucleus A. The energy of interaction between two nuclei is given by the coupling constant J, which is independent of the applied field. In the case of two coupled protons, A and B, the chemical shift of proton A will be affected by the direction of the spin of proton B. If the field of proton B is aligned with the applied field, A will be slightly deshielded because it will feel the effect of proton B plus the applied field. The converse is true if the field of proton B opposes the applied field, i.e. proton A will be shielded (Fig.7). Internuclear couplings are transmitted via the bonding electrons. For couplings involving hydrogen, the dominant mechanism for

internuclear couplings is the interactions between the nuclear moments and the electronic spins in the s-orbitals (Fig.7).

Fig.7 Mechanism of Transferring Spin Information⁶¹.

There is a dependence of the ^3J proton-proton coupling constant on the CH-CH dihedral angle, originally theoretically predicted by Karplus⁶² to be

$$J_{(\text{CHCH})} = 10 \cos^2 \phi$$

If the dihedral angle is 180° , the coupling is maximal (10Hz), and if the dihedral angle is 90° there is little or no orbital overlap and the coupling constant is minimal (Fig.8).

Fig.8 Variance of the Coupling Constant with the Dihedral Angle α .

Fourier Transform NMR

The basic Fourier Transform experiment (FT) and the behaviour of the magnetization in the rotating frame can be described as follows: If a sample containing nuclei of nuclear spin $1/2$ is placed in a strong magnetic field, the nuclei will precess around the direction of the field with a frequency ω_0 , the Larmour frequency. The nuclei will either be aligned or opposed to the applied field. The slight Boltzman excess of nuclei aligned with the applied field will give rise to a resultant magnetization vector, M_0 , which lies in the same direction of the applied field. If this system can be considered with reference to a rotating frame, which rotates at the Larmour frequency, the magnetic behaviour now is completely described by a magnetization vector M_0 acting along the applied field, B_0 (Fig.9). Application of a radiofrequency (RF) pulse at the resonant frequency, ω_0 , along the x' -axis results in the bulk magnetization flipping into the $x'y'$ plane, the angle of which is determined by the power and duration of the RF pulse. A 90° pulse will cause the bulk magnetization to become aligned with the y' -axis. It is possible to excite all the nuclei of a given species by using a very short pulse of frequency F for t seconds, which will give a range of frequencies centred about F with a bandwidth of $1/t$. Following the RF pulse, the magnetization begins to fan out due to magnetic field inhomogeneity and to spin-spin relaxation. Using a rotating frame terminology, this results in the transverse magnetization splitting into its different component parts corresponding to the different nuclear magnetic moments

which rotate at different speeds in the $x'y'$ plane. In the continuous wave experiment previously described absorption of energy by the spins is measured. In the pulsed experiment, all the nuclei in the sample are excited by a short, high energy pulse. The pulse is then turned off, and the energy emitted by the spin system as it returns to equilibrium is recorded. As the components of the transverse magnetization span out, the intensity of the signal reaching the receiver gradually decreases and this exponential decay of signal intensity with time is known as the free induction decay (FID). Fourier transformation gives a frequency dependant function, which is the normal NMR spectrum⁶⁰.

Fig.9 Pulsed Fourier Transform Techniques

9.1 Motion of spin 1/2 nuclei in a magnetic field rotating at ω_0

9.2 The effect of an RF pulse with frequency ω_0 for a time t on the bulk magnetization vector M_0

9.3 Relaxation mechanisms

Spin-Lattice Relaxation Time

Following the perturbation of the magnetization vector due to an assembly of ^{13}C nuclei by the application of a pulse, the magnetization vector starts to relax back to its equilibrium value by two processes known as spin-lattice and spin-spin relaxation. ^{13}C relaxation times cover a considerable range from milliseconds in molecules of high molecular weight, to several hundred seconds for unprotonated carbons⁶⁰. This discussion will focus on spin-lattice relaxation mechanisms, and most specifically dipole-dipole relaxation as frequently the only significant relaxation mechanism for carbon atoms which bear directly bonded protons is dipolar relaxation⁶³.

In solution the freedom of motion of molecules interacting with other molecular species is restricted resulting in a longer correlation time. Excluding molecules with very large molecular weight, this increase in the correlation time normally results in a decrease in the NMR spin-lattice relaxation time (T_1) for molecules undergoing primarily dipolar relaxation (Fig.10)⁶⁰. Measurements of T_1 have been reported in studies of the binding of sulfonamides and penicillin to bovine serum albumin⁶⁴, lipid-protein interactions⁶⁵, cellulose-water interactions⁶⁶, calcium-parvabumin binding⁶⁷, and bilirubin-dilysine interactions⁶⁸.

Fig.10 Effect of Correlation time on Spin-lattice relaxation Time (T_1)

There are several mechanisms which can contribute to spin-lattice relaxation in a molecule. The most common of these mechanisms are; dipole-dipole (T_{1DD}), spin-rotation (T_{1SR}), quadrupolar (T_{1Q}), scalar (T_{1SC}), and chemical shift anisotropy (T_{1CSA}). These combine to generate an overall spin-relaxation time (T_1) given by the equation:

$$1/T_1 = 1/T_{1DD} + 1/T_{1SR} + 1/T_{1Q} + 1/T_{1SC} + 1/T_{1CSA} \text{ (Eq. 4.1)}^{60}$$

There is the potential for an efficient relaxation mechanism when the nucleus undergoing relaxation is directly bonded to a second nucleus possessing a magnetic spin. In the case of a ^{13}C nucleus bonded to a proton, the two spins can be considered as small dipoles located at the centre of the ^{13}C and ^1H atoms. Hence the ^{13}C nucleus will experience a small field due to the dipolar interaction with the proton. This will depend on the magnetic field which has a component at or close to the Larmour frequency (the frequency with which the nucleus will precess around the direction of the magnetic field) of their line of interaction (i.e. the orientation along the C-H bond) relative to the magnetic field in the spectrometer. (Fig.11)

Fig.11 Orientation of the Dipolar Interaction Relative to the Magnetic Field.

The magnetic field H_{DD} created at the ^{13}C nucleus is given by

$$H_{DD} = (3 \cos^2\theta - 1) \gamma_H h / 4\pi r^3 \quad (\text{Eq.4.2})^{60}$$

As the molecule tumbles in solution, the changing θ will result in fluctuations in H_{DD} . Relaxation can be induced by any oscillating electric or magnetic field which has a component at or close to the Larmour frequency of the nucleus concerned. Consequently, the oscillations in H_{DD} comprise a possible relaxation mechanism. For most solutions the contribution by dipole-dipole relaxation can be expressed as

$$1/T_{1DD} = [\gamma^2 H \gamma_C^2 (h/2\pi)]^2 T_C \Sigma (1/r_{C-H}^6) \quad (\text{Eq.4.3})^{60}$$

T_C is the correlation time, and is a measure of how quickly the molecule undergoes reorientation in solution; the summation is carried out over all atoms in the molecule. The effect is inversely proportional to r^6 and thus drops off very rapidly with distance. For this reason, dipole-dipole relaxation is normally only seen for carbons bearing protons. This equation is only valid if $\omega_C T_C \ll 1$, where ω_C is the ^{13}C Larmour frequency. This condition is normally met for samples with $MW < 1000$ in a nonviscous solution⁶⁰.

There are two customary techniques for determining spin-lattice relaxation times. In the inversion recovery method, a 180° pulse is applied to a system at equilibrium which inverts the magnetization to the $-z$ -axis. Immediately following the pulse, the magnetization vector M_z equals $-M_z^0$. M_z will now start to relax back along the z -axis toward its equilibrium value M_z^0 via the spin-relaxation process.

Fig.12 Relaxation of the Magnetization Vector M_z via Spin-Lattice Relaxation

It is possible to express this mathematically as

$$M_z = M_z^0(1 - 2\exp(-t/T_1)) \quad (\text{Eq.4.4})^{60}$$

where M_z is the component of M along the z -axis t s following the application of the 180° pulse, M_z^0 is its equilibrium value and T_1 is the spin-lattice relaxation time (considering the simplest case of a spectrum consisting of one line). At one point M_z will pass through zero, i.e.

$$0 = M_z^0(1 - 2\exp(-t/T_1)) \quad (\text{Eq.4.5})$$

$$T_1 = t_0 / 2.303 \log 2 = t_0 / 0.693$$

where t_0 is the time at which $M_z=0$. In practice, after a delay of t s a second 90° pulse is applied which tips the magnetization onto the $-y'$ -axis where it can be detected by the spectrometer. As a single pulse is not adequate to detect ^{13}C signals at natural abundance the cycle must be repeated. At the end of the cycle, the magnetization lies along the $-y'$ -axis, therefore before a second 180° pulse can be applied. It is necessary to wait a period of $5T_1$ s to allow the magnetization vector to relax back to its equilibrium value (M_z^0). The structure of the cycle is

then $(180^\circ \text{ pulse} - \text{delay } t_s - 90^\circ \text{ pulse} - \text{Pulse delay } (5T_1))n$, where n is the number of times the cycle is repeated.

Fig.13 Inversion Recovery Pulse Sequence

At the end of the experiment, the accumulated FID is transformed and the intensity of the signal determined. Rearranging Eq.4.4 gives

$$M_z - M_z^0 = -2M_z^0 \exp(-t/T_1) \quad (\text{Eq.4.6})^{60}$$

taking logs gives

$$\ln(M_z - M_z^0) = -\ln(2M_z^0) - t/T_1$$

A plot of $\ln(M_z - M_z^0)$ against t will give a straight line with a gradient of $-1/T_1$. T_1 values are calculated automatically by the computer which performs a least squares analysis on the data.

Nuclear Overhauser Enhancement Spectroscopy (NOESY)

NOESY has been used in the analysis of large molecules such as the determination of protein secondary and tertiary structure⁶³.

Enhancement of the carbon signal under conditions of proton noise decoupling is much greater than expected due to the collapse of the multiplet structure. This enhancement is known as the nuclear Overhauser effect (nOe). The spin states and transition probabilities of a coupled two spin system, in which the spins A and B (¹³C-H) are coupled by means of a relaxation interaction are shown in Fig.14.

Fig.14 Spin States and Transition Probabilities (W_i) for a coupled Two Spin System.

If the total number of spins is N , then it is assumed the intermediate energy levels each have a population of $N/4$. The $\alpha\alpha$ state, being lower energy will be slightly more populated than the higher energy $\beta\beta$ state, and thus the population of the $\alpha\alpha$ state is $N/4+\delta$, and the population of the $\beta\beta$ state is $N/4-\delta$. The intensity of the two allowed transitions (W_{1A}) is proportional to the difference in population between the two spin states involved, ($P_{\alpha\alpha}-P_{\beta\alpha}$ and $P_{\alpha\beta}-P_{\beta\beta}$) giving δ .

Following perturbation, the spin states will begin to relax back to their equilibrium populations via spin-lattice and spin-spin relaxation mechanisms. The relaxation pathways are the reverse of the labelled transitions (W_{1A}) W_{1B} , and most effective, the mechanism by which the $\beta\beta$ state relaxes to the $\alpha\alpha$ state W_2 . This transition corresponds to a change in total spin of two units, and would normally be forbidden by the selection rules, however the transition occurs without involving the direct emission of radiation and hence without violating the selection rules⁶⁰.

When nucleus B is irradiated with a strong decoupling field inducing rapid transitions between spin states, the population of the spin states become equal. The efficient relaxation via the W_2 mechanism maintains the Boltzman distribution between the $\alpha\alpha$ and $\beta\beta$ states. Since the populations linked by the W_{1B} transitions are equal, the population of the $\alpha\beta$ state must be equal to that of the $\alpha\alpha$ state ($N/4+\delta$), and the population of the $\beta\alpha$ state equal to that of the $\beta\beta$ state ($N/4-\delta$). The difference in population in the A transition states is now equal to 2δ ($P_{\alpha\alpha}-P_{\beta\alpha}$ and $P_{\alpha\beta}-P_{\beta\beta}$), compared to δ in the absence of decoupling of B. This model would predict ^{13}C intensities would increase by a factor of 2 in the presence of proton decoupling.

Quantitatively, the derivation of nOe considers the total magnetization (M_z) for each spin given as the difference in population of the individual spin states α and β .

$$M_z = n_\alpha - n_\beta \quad (\text{Eq.5.1})^{60}$$

At equilibrium.

$$M_z^0 = n_{\alpha 0} - n_{\beta 0} \quad (\text{Eq.5.2})^{60}$$

where the Boltzman distribution of the spins is achieved,

$$n_{\alpha 0} - n_{\beta 0} = \exp(-E/kT) = 1 - \gamma h B_0 / 2\pi kT \quad (\text{From Eq.1.4}) \quad (\text{Eq.5.3})^{60}$$

Thus,

$$n_{\alpha 0} - n_{\beta 0} = N/2(\gamma h B_0 / 2\pi kT) \quad (\text{Eq.5.4})^{60}$$

For two different spins, the different magnetizations (M_z^0) at equilibrium for the same field and numbers of spins is given by

$$M_{A0}/M_{B0} = \gamma_A/\gamma_B \quad (\text{Eq.5.5})^{60}$$

For the coupled spin system (Fig.14) how the magnetization of the spins relaxes back to the equilibrium value after a perturbation must be considered. The relaxation will be proportional to the perturbation ($M_z - M_z^0$). Considering the A spin, the number of ways of relaxing the A spin from β to α are, from Fig.14 equal to $2W_{1A} + W_2 + W_o$. The W_2 and W_o transitions affect both the A and B spins, which must be accounted for, giving an overall rate equation of

$$dM_A/dt = -(2W_{1A} + W_2 + W_o)(M_A - M_{A0}) - (W_2 - W_o)(M_B - M_{B0}) \quad (\text{Eq.5.6})^{60}$$

When the B-spin is totally decoupled the populations in the α and β state are equal, and $M_B = 0$, therefore

$$dM_A/dt = 0 = -(2W_{1A} + W_2 + W_o)(M_A - M_{A0}) + (W_2 - W_o)M_{B0} \quad (\text{Eq.5.7})^{60}$$

And the new equilibrium value of the A spin under conditions of total decoupling of the B-spin is related to the unperturbed value (M_{A0}) by the equation

$$M_A/M_{A0} = 1 + [(W_2 - W_0)/(2W_{1A} + W_2 + W_0)](\gamma_A/\gamma_B) = 1 + n \text{ (Eq.5.8)}^{60}$$

where n is the nOe. For most systems

$$(W_2 - W_0)/(2W_{1A} + W_2 + W_0) = 0.5$$

$$\text{and } n = 0.5(\gamma_A/\gamma_B)$$

For a homonuclear experiment, $\gamma_B = \gamma_A$, and the maximum enhancement is .5, or 50%. This is extensively used to determine protons which are in close spacial proximity.

NOESY

The NOESY experiment is the 2D experiment of a transient 1D experiment described above. In the case that A has no near neighbours, and B has several, when a strong RF pulse is applied at the resonance frequency of nucleus B, and then switched off, the spin populations of B will return to the equilibrium value via dipolar interactions with A and other surrounding nuclei (B', B"...). If the resonance frequency of A is irradiated in a similar manner, A will relax through B, which will in turn relax through B', B"... Obviously, in this case the relaxation of the spin which is irradiated is as important to that which it is dipolar coupled as far as the resultant nuclear Overhauser enhancement is concerned⁶⁰. Thus the size of the nOe observed in the transient experiment depends on the time at which the FID is collected. The basic NOESY experiment centres around a three pulse sequence (delay time-90° pulse-t₁-90° pulse-T_m-90° pulse-t₂)⁷⁰. The first pulse flips the bulk magnetization to the x'y' plane. During the first delay time, t₁, free

precession occurs, which has the effect of frequency labelling the spins. The experiment is designed such that the magnetization which evolved with some frequency during t_1 (i.e. labelled with this frequency) evolves with a different frequency during t_2 , giving rise to off diagonal, or cross peaks. The second 90° pulse rotates the x' and y' components through 90° and the y' component is transferred to the $-z'$ axis. This is followed by the mixing time, in which the effect of magnetization exchange due to cross relaxation (as described above) is recorded by the spins in the form of population changes. The final 90° pulse again generates transverse magnetization for signal detection during t_2 . Double Fourier transformation with respect to t_1 and t_2 results in the 2D NOESY spectrum⁶⁰. The effect of cross-relaxation between neighbouring protons during T_m is to transfer magnetization between them. This results in off diagonal intensities which are roughly proportional to $1/r^6$, where r is the proton-proton distance⁷⁰.

Peptide Synthesis

The procedure of solid phase peptide synthesis was first proposed by R.B. Merrifield in 1962⁷¹. He described the synthesis of a tetrapeptide by addition of successive benzyloxycarbonyl amino acids to a polystyrene resin⁷¹. In this type of sequential reaction, quantitative reactions at each step are crucial for success. The principle is simple, the peptide is elongated while it is attached to a stable, solid particle throughout all the synthetic steps, separated from soluble reagents and solvents by filtration, and finally detached from the solid support and

characterized⁷¹.

Unambiguous formation of a peptide bond between two structurally similar amino acids requires that the amino group of one and the carboxyl group of the other must be blocked from participating in the coupling reactions⁷². As well, there are many reactive groups in the side chains of naturally occurring amino acids which must be prevented from reaction. There is then a requirement for two different types of amino protecting groups, a temporary one, which will protect the amino terminal of the growing peptide and can be removed without removing a permanent protecting group on a functional group in the side chain of the amino acid. Traditionally, t-butoxycarbonyl and benzyloxycarbonyl groups were used as temporary and permanent amino protecting groups respectively (Fig.15)⁷¹. T-butoxycarbonyl groups are cleaved under relatively mild acidic conditions, where benzyloxycarbonyl groups are essentially inert. Employment of acid labile derivatives for both temporary and permanent amino group protection results in the latter requiring extremely strong acidic treatment at the end of the synthesis. Cleavage of these benzyloxycarbonyl amino acids requires extremely severe conditions (HF).

Fig.15 Protecting Group Combinations for the Merrifield Synthesis⁷¹.

A variation on the Merrifield technique has been developed by Atherton et al.⁷³ (Fig.16)

Fig.16 Protecting Group Combinations for the Fmoc Method of Peptide Synthesis.⁷³

This system uses a base labile α -amino protecting group, fluorenylmethoxycarbonyl (Fmoc), in combination with the acid labile t-butoxycarbonyl protecting group for amino groups on the side chains of amino acids. The Fmoc protecting group is base labile due to characteristics of the dibenzocyclopentadiene structure. Resonance stabilization of the derived cyclopentadiene anion imparts acidity to the 9-hydrogen atom enabling its removal by bases and initiation of the β -elimination reaction (Fig.17)⁷³.

Fig.17 Base Catalyzed Elimination of the Fmoc group.⁷³

One of the most serious side reactions in peptide synthesis is racemization at the C α atom of the amino acid. The Fmoc method of peptide synthesis requires much less severe conditions than the Merrifield technique, thus lessening the chance of side reactions and racemization.

Peptide bond formation between a carboxyl and amino group does not occur spontaneously under normal conditions. One of these two groups must be therefore converted into a more reactive form. Invariably, the carboxyl group of one amino acid is activated by augmenting the electrophilic character of the carbon of the carboxyl group⁷². The addition of an electron substituent to this group decreases the electron density of this carbon, permitting nucleophilic attack by a nucleophilic nonionized amino group. Pentafluorophenol esters in the presence of hydroxybenzotriazole as a catalyst are exceptionally reactive; due to the strength of the carbon-fluoride bond and absence of any other reactive centres the couplings are essentially free of side reactions⁷³.

The peptide is built from the carboxyl to the amino terminus. The first amino acid is esterified to the solid polyamide support via a benzyl ester linkage with a p-alkoxy substituent to enhance the acid lability of the ester bond. The second amino acid, with a protected (Fmoc) amino group, and an activated (pentafluorophenol ester) is then added. When coupling is complete, as determined by a qualitative test, the unreacted solvated reagents are removed. This is followed by deprotection of the newly formed amino terminus and sequential couplings of

other amino acids. After all couplings are completed, the peptide is cleaved from the resin, and fully deprotected by treatment with acid, giving the free peptide.

Molecular Modelling

Drug development has been greatly aided by new techniques in computational chemistry. Biological activity is critically dependant on 3-D interactions between molecules, and with the help of computer programs which use empirical energy functions macromolecular systems can be modeled. The empirical energy function of CHARMM⁷⁴ consists of the following terms: Bond potential (E_b) and bond angle potential (E_θ), which account for bond and angle deformations, dihedral angle (torsion) potential (E_ϕ), which is a four atom term based on the dihedral angle about an axis defined by the middle pair of atoms, E_ω is an improper torsion term designed to maintain chirality about a tetrahedron (i.e. C_α), and planarity about planar atoms (i.e. $C=O$). Nonbonded interactions are described by van der Waals interactions (E_{vdw}), electrostatic potential (E_{el}), and hydrogen bonding (E_{hb}). There are two constraints used to restrict the changes that occur in the structure: atom harmonic constraints (E_C) are used to avoid large displacements of atoms when minimizing, dihedral constraints ($E_{C\phi}$) are used to maintain certain local conformations. The total energy is then:

$$E = E_b + E_\theta + E_\phi + E_\omega + E_{vdw} + E_{el} + E_{hb} + E_C + E_{C\phi} \quad (\text{Eq.6.1})^{74}$$

CHARMM supports five different energy minimization algorithms, the simplest of these is a nonconvergent method (steepest descents). In this procedure

a displacement opposite to the potential energy gradient is added to the coordinates at each step. The step size is increased if a lower energy results, otherwise it is decreased⁷⁴. This procedure is normally used as a rough estimate of the energy minimum and is a nonconvergent method. The adopted basis Newton-Raphson (ABNR) is a convergent method in which the positions and forces from the previous steps are stored and generally stops when minimum energy has been achieved⁷⁴. The lowest energy of interaction of two structures can be calculated using van der Waals and electrostatic energies of interaction.

Experimental Design

In this study, the tetrapeptide isoleucine-lysine-lysine-lysine (IKKK) was synthesized using the Fmoc-polyacrilamide method of solid phase peptide synthesis. This sequence is highly conserved in retroviral RTs and represents a possible nucleotide triphosphate binding sequence on HIV RT. The interaction of the metabolically active nucleotide analogue, ddATP, with this sequence as well as di-L-lysine was determined via ¹³C relaxation times and 2D nOe spectroscopy. The results of these experiments together with molecular modelling data and the solution conformational equilibrium of ddATP were used to model possible interactions at the triphosphate binding site of HIV RT.

RESULTS and DISCUSSION

Peptide Synthesis

The synthetic sequence is given in Fig.18. Sequential couplings of amino acid residues to the growing peptide chain required increased acylation time for completion of reaction. The sterically hindered isoleucine pentafluorophenol ester reacted most slowly, with a total acylation time of 3 hours needed for completion.

Racemization is always a concern when synthesizing biological molecules. Normally, if racemization has occurred, the diastereomers can be separated with the use of reverse phase HPLC⁶⁷. As shown in Fig.19 the HPLC trace of the crude peptide contains many peaks, the peak at RT 2.23 being the desired product.

Fig.19 HPLC trace of Crude Tetrapeptide. 80% .1M NaPO₄, 20% EtOH
20mM Pentane Sulfonate, .8mL/min.

Fig.18 Synthetic Sequence

$N\alpha$ -Fmoc $N\epsilon$ -BOC-L-lysine pentafluorophenyl ester.

+

$N\epsilon$ -BOC-L-lysine resin ester

Fig.18 con't

Sequential couplings
performed with:

1. $\text{N}\alpha$ -Fmoc-L-lysine pentafluorophenyl ester
2. $\text{N}\alpha$ -Fmoc-L-isoleucine pentafluorophenyl ester

CLEAVAGE

95% TFA, 2Hrs, 25°

Spectral Analysis

The numbering scheme used for di-L-lysine, ddATP and isoleucine-lysine-lysine-lysine (IKKK) is given in Fig.20.

Fig.20 Numbering Scheme for Di-L-lysine, ddATP, and IKKK.

The ^1H NMR spectrum of di-L-lysine is given in Fig.21.

Based on literature values⁷⁶, the two triplets at ca. 3.8ppm and 4.0ppm correspond to the C_α protons of the two lysine residues. The large multiplet structure at ca. 2.9 ppm contains the resonances of both $\text{C}_\epsilon\text{H}_2$ protons. The signal due to both groups of C_β , C_δ and C_γ are at ca. 2.0 ppm, 1.7 ppm and 1.4 ppm respectively. The amino acid side chain resonances in a peptide are essentially identical to those

of the individual amino acid residues⁷⁶, and it should be noted the individual signals due to only one group of protons (i.e. C_3H_2 and C_2H_2) are indistinguishable.

Fig.21 200 MHz ^1H NMR of Di-L-lysine in D_2O

The ^{13}C NMR spectrum of di-L-lysine is given in Fig.22.

Utilizing the same numbering scheme, the most deshielded carbon at 180.9 ppm is the carboxyl group of the C-terminal residue C_{12} ⁷⁵; the line at 171.5 ppm corresponds to the carboxyl group of the N-terminal lysine, or C_6 . This is in good agreement with literature data which gives the chemical shift of the carboxyl group of lysine as 174.4 ppm, which is shifted downfield by 5.4 ppm for C_{12} and upfield 3.5 ppm for C_6 .⁷⁵ Generally, the side chain resonances of an amino acid residue will move upfield if the residue is on the amino terminal, and the resonances of the side chains in the carboxyl terminal residue will move downfield.⁷⁵ The chemical shifts of a nonterminal amino acid do not change appreciably from that of

individual amino acid residues. Based on this, the following assignments have been made: C₁, 57.3, C₂, 28.8, C₃, 23.4, C₄, 32.8, C₅, 41.5, C_{7,13}, 57.9, C₈, 28.8, C₉, 24.7, C₁₀, 33.2, and C₁₁, 41.5.

Fig.22 ^{13}C NMR of di-L-lysine in D_2O

A FABMS (fast atom bombardment mass spectrum) was performed on di-L-lysine (M.W. 276.2) (Fig.23). This spectrum shows a peak corresponding to $(\text{M}-1)^+$ as well as peaks of higher molecular weight which correspond to $(\text{M}+\text{Cl})^+$ at 333.2 m/z, $(\text{M}+\text{Na})^+$ at 297.2 m/z, and $(\text{M}+\text{Cl}_2+\text{Na})^+$ at 369.2 m/z.

711314 Scan 4 RT=0.52 100%=220290 mv 19 Jun 92 3:56
LRP +FAB

Fig.23 FABMS of di-L-lysine

Assignment of proton resonances for the tetrapeptide IKKK were made based on literature values as well as the NMR spectra of di-L-lysine. The C_α region ^1H NMR spectrum of the tetrapeptide is shown in Fig.24. The entire ^1H spectrum is given in Fig.25.

Fig.24 C_α region of the ^1H 400 MHz NMR spectrum of IKKK in D_2O .

Fig.25 ^1H NMR spectrum of IKKK in D_2O .

There are two quartets and a triplet which correspond to each of the C_α protons of the three lysine residues at 3.9 ppm, 4.0 ppm, and 4.1 ppm, as well as a doublet which corresponds to the C_α proton of isoleucine at 3.6 ppm. The

relative integration of these peaks is 1:94:1:1. The absolute value of $J_{AX}+J_{BX}$ for the $C_\alpha^1H - C_\beta^1H_2$ (ABX) spin systems of the three lysine residues as well as J_{AX} of the AX spin system of the isoleucine residue were calculated (Table 1).

Table 1 C_α^1H Chemical Shifts (δ_c from TMS ± 0.1) and ^3J ^1H - ^1H Coupling Constants^a (Hz ± 1)

Chemical Shift	J_{AX}	J_{BX}
4.12	7.2	7.2
4.02	7.2	6.8
3.86	6.8	6.4
3.61	5.6	---

^a-Coupling constants were determined by simulation of the X part of the ABX spectrum.

A COSY (correlation spectroscopy) spectrum (Fig.26) of the tetrapeptide reveals a connectivity between the doublet at 3.6 ppm (C_1H) with the peak at 1.7 ppm. This peak then must correspond to C_2H . This peak shows a connectivity with a doublet at 0.7 ppm which must then may be assigned to C_3H_3 . Unambiguous assignment for the remainder of protons is not possible due to the large amount of overlap, however some general assignments may be made based on the connectivities. The three lines corresponding to C_7H , $C_{13}H$, and $C_{19}H$ show

a connectivity to the multiplet centred at 1.5 ppm, which then must be due to C_8H , $C_{14}H$, and $C_{20}H$. This is connected to the multiplet centred at 1.2 ppm, which is due to the γ protons on the lysine side chain (C_9 , C_{15} , and C_{21}). From literature values, the resonances at ca. 2.6 ppm and 2.75 ppm can be assigned to the C_ϵ protons (C_{11} , C_{17} , and C_{23}). The C_δ protons $C_{10}H$, $C_{16}H$, and $C_{22}H$ are in the multiplet centred at 1.5 ppm.

Fig.26 COSY spectrum of IKKK in D_2O .

These assignments allow the assignment of ^{13}C lines using HETCOR (heteronuclear correlation spectroscopy) (Fig.27). C_1 , 59.9 ppm, C_{19} , 56.3ppm, C_7 and C_{13} cannot be unambiguously assigned, as the lines overlap at 56.0 ppm. These values are consistent with literature values for the chemical shifts of C_α in

peptides. The other resonances may be assigned as follows: C_8 , C_{14} , and C_{20} overlap at 28.8 ppm. Literature data suggests the shift of C_β of a nonterminal amino acid is zero, but that of a C-terminal is approximately 1.2 ppm downfield⁷⁵. However, the ^{13}C spectrum of di-L-lysine shows the resonances due to the C_β carbons overlap at 28.8 ppm (Fig.22), consistent with this result. The lines due to C_9 , C_{15} , and C_{21} appear at 24.6 ppm and 24.5 ppm, the line at 24.5 ppm is approximately one half the intensity of that at 24.6 ppm and is probably due to C_γ of the C-terminal lysine residue C_{21} ⁷⁵. The chemical shifts of the C_δ carbons, C_{10} , C_{16} , and C_{22} are 33.0 ppm and 32.9 ppm. As C_δ of the C-terminal lysine residue should be more deshielded⁷⁵, and the intensity of the line at 33.0 ppm is approximately one half of that at 32.9, the line at 33.0 ppm may be assigned to C_{22} -C-terminal. The ϵ carbons, C_{11} , C_{17} , and C_{23} , like the β carbons overlap at 41.6 ppm. The lines due to the carbons of isoleucine are assigned using literature values⁷⁵ as follows, C_2 at 38.9, C_3 at 26.5 ppm, C_4 at 13.0 ppm and C_5 at 16.5 ppm.

There are four peaks in the amide region of the ^{13}C NMR spectrum, corresponding to the four amide functional groups (Fig.28a). The most deshielded, at 179.4 ppm is C_{24} the carboxyl terminal residue, the most shielded amide carbon at 171.8 is C_6 , in the N-terminal isoleucine residue⁷⁵. The two resonances at 175.7 and 175.6 are due to C_{12} and C_{18} but cannot be unambiguously assigned.

Fig.27 HETCOR of IKKK in D₂O

Fig.28 ^{13}C Spectrum of IKKK in D_2O

Fig.28a Amide Region of ^{13}C NMR spectrum of IKKK in D_2O .

The FABMS of IKKK like that for di-L-lysine shows a peak which corresponds to $(\text{M}-1)^+$ at 516.5 m/z (Fig.29). As well there are peaks of higher molecular weights at 665.0 m/z, which corresponds to the tetrapeptide with pentane sulfonate as a counter ion.

Fig.29 FABMS spectrum of IKKK.

The results of the T_1 experiments are given in Tables 2,3 and Figs 30-33.

Statistically significant changes in T_1 values are shown in Fig.34.

Table 2

 T_1 Values for Di-L-Lysine in D_2O

^{13}C Line Assignments	^{13}C Chemical Shift (d_c from TMS $\pm$ .01)	^{13}C Chemical shift with ddATP (d_c from TMS $\pm$ .01)	T_1 in $D_2O(s)$ (.03 mmol) $\pm$		T_1 in D_2O (.03 mmol) with ddATP (.03 mmol) $\pm$	
C-7	57.92	57.52	.31	.02	.28	.03
C-1	55.32	56.33	.30	.01	.20	.1
C-5	41.67	41.74	.84	.03	.71	.02
C-11	41.48	41.74	.84	.03	.71	.02
C-10	33.23	35.24	.28	.02	.19	.01
C-4	32.81	33.37	.28	.02	.20	.01
C-2,8	28.80	28.89	.38	.01	.57	.1
C-3	24.66	24.65	.23	.005	.28	.01
C-9	23.37	24.03	.23	.02	.19	.05

Table 3

 ^{13}C T_1 Values for IKKK in D_2O

^{13}C Line Assignments	^{13}C Chemical Shift (d_c from $\text{TMS} \pm .01$)	^{13}C Chemical shift with ddATP (d_c from $\text{TMS} \pm .01$)	T_1 in $\text{D}_2\text{O(s)}$ (.038 mmol) $\pm$		T_1 in D_2O (.038 mmol) with ddATP (.032 mmol) $\pm$	
C-19	57.5	56.4	.18	.03	.14	.03
C-7,13	56.2	56.2	.13	.04	.23	.09
C-11	41.6	41.6	.47	.06	.33	.07
C-17	41.6	41.6	.47	.06	.33	.07
C-23	41.6	41.6	.47	.06	.33	.07
C-2	38.9	38.8	.12	.1	.32	.14
C-22	33.6	33.3	.17	.04	.10	.02
C-16	33.0	32.9	.19	.02	.05	.01
C-10	33.0	32.9	.19	.02	.05	.01
C-8	28.9	28.8	.10	.03	.32	.07
C-14	28.8	28.8	.10	.03	.32	.07
C-20	28.8	28.8	.10	.03	.32	.07
C-3	26.5	26.5	.73	.08	.10	.06
C-9	24.5	24.5	.25	.05	.20	.05
C-15	24.5	24.5	.25	.05	.20	.05
C-21	24.1	24.0	.46	.1	.20	.05
C-5	16.5	16.5	.67	.09	.16	.07
C-4	13.0	12.9	.69	.08	.67	.06

*C-1 could not be determined due to overlap.

Fig.30 T_1 Spectrum of di-L-lysine in D_2O

Fig.31 T_1 Spectrum of di-L-lysine with ddATP in D_2O

Fig.32 T_1 Spectrum of IKKK in D_2O

Fig.33 T_1 Spectrum of IKKK with ddATP in D_2O

Fig.34 Changes in ^{13}C Spin-lattice Relaxation Times for di-L-lysine and IKKK in Solution with ddATP.

T_1 values for the side chain carbons of di-L-lysine were greater than for the lysine residues in IKKK, as would be expected as the spin-lattice relaxation time tends to decrease with increasing molecular weight⁶⁵. Lysine residues in both di-L-lysine and IKKK showed very similar changes in T_1 values when complexed with ddATP. Decreased motion was seen for the C_δ and C_ϵ carbons, and increased motion for C_β carbons. This result is somewhat perplexing as decreased motion is expected for all carbons if a complex between ddATP and IKKK or di-L-lysine is being formed, leading to a corresponding increase in molecular weight. The association of ddATP to these peptides results in a statistically significant increase in motion for C_β , possibly due to chain perturbation. The error on the reported T_1 values are somewhat large, especially in the cases where the resonances of two carbons overlap. However, the errors are similar to those reported in the literature of approximately 10%⁶¹. C_β , $\text{C}_\gamma\text{H}_2$, and $\text{C}_\gamma\text{H}_3$ of the isoleucine residue in IKKK showed dramatic decreases in T_1 , with values changing by as much as 0.6 s. In the literature, changes in spin lattice relaxation time of as little as .06s have been

reported as significant. Changes in T_1 for all other carbons was not statistically different in the presence of ddATP.

A suppressed nOe spectrum reveals the dominant mechanism of relaxation for methylene carbons is dipole-dipole, 100% for both C_β C_γ of the lysine side chains and $C_\gamma H_2$ of the isoleucine side chain, 90% for both methyl groups of the isoleucine side chain, and 80% for C_β of the lysine side chain. This indicates changes in T_1 values are due to differences in dipolar relaxation rates which are intimately related to freedom of motion. Another mechanism of spin-lattice relaxation may account for the increase in T_1 for C_β when in solution with ddATP.

Fig. 35 shows a two dimensional NOESY spectrum of di-L-lysine complexed with ddATP. There is a correlation between H_3^1 and $C_\beta H$, as well as C_β indicating an interaction with a distance of less than $5^0 A^{69}$. Also evident is a change in chemical shift of H_3^1 , which is further upfield, indicating shielding.

In an aqueous environment, ddATP is dephosphorylated to ddADP, ddAMP, etc. In a slightly alkaline solution however, ddATP does not decompose significantly. The nature of the interaction study depends greatly on the integrity of ddATP. In order to determine the degree to which ddATP had become dephosphorylated, a proton NMR of ddATP was obtained immediately following the T_1 experiments. Comparison of this spectrum with a ^1H NMR spectrum of decomposed ddATP revealed that little to no decomposition had occurred over the acquisition of the T_1 spectra (Fig. 36).

Fig.35 NOESY spectrum of di-L-lysine with ddATP

Fig.36 ¹H NMR of Decomposed and Unaltered ddATP.

Molecular Modelling

Molecular models of IKKK and ddATP were constructed using *QUANTA* version 3.3 with the empirical energy function of CHARMM (Molecular Simulations Inc., Burlington Mass., U.S.A.). The lowest energy conformations of IKKK and ddATP were determined by a steepest descent followed by an adopted basis Newton-Raphson procedure with a dielectric constant of 80. A Boltzman jump procedure was then performed to search the conformational space for the geometries with the lowest energies (Figs 37-38) which were used to explore molecular interactions. The low energy conformation of IKKK was set as stationary while ddATP was used as a probe, rolling over a box of dimensions $20^{\circ}\text{\AA}$, flipping by 90° every $0.5^{\circ}\text{\AA}$ to calculate van der Waals plus electrostatic interaction energies.

Fig.37 Lowest Energy Conformation of ddATP

Fig.38 Lowest Energy Conformation of IKKK

Brick maps showing interaction energies of lower than -26,-30 and -35 kcal mol⁻¹ are represented in Figs.39-41 respectively. There is an electrostatic interaction between the positively charged amino groups of lysine as well as the N-terminal amino group of isoleucine with ddATP.

Fig.39

Fig.40

Fig.41

Figs.39-41 Brick Maps depicting interaction of ddATP and 1KKK with interaction energies lower than -26, -30 and -35 kcal mol⁻¹ respectively

Nucleotide Conformation

The ^1H NMR of ddATP is given in Fig.42. Proton resonance assignments were made based on literature values of ddA⁵⁷. The coupling constants are given in Table 4. The conformation of the furanose moiety was resolved based on eight coupling constants, which were determined for each temperature by means of spectral simulations (Fig.43a-e) (Table 4). The program PSEUROT relates experimental coupling constants to a two state equilibrium of the conformation of the furanose ring. The sugar moiety adopts an equilibrium between C3'-endo and C2'-endo conformations (Table 5). It should be noted the N-conformer predominates in the two state equilibrium in a ratio of 4:1 (Fig.44).

Comparison of the solution conformation of ddATP with related compounds (Table5) reveals a similarity in conformations. The S-conformer (C2'-endo) predominates in solution for the natural substrate, 2'-deoxyadenosine, which differs from the dominant N type conformer of the 3'-deoxy analogues⁸⁰.

From analysis of $^3J_{4,5}$ and $^3J_{4,5'}$ coupling constants it is evident the γ' rotamer predominates (68%) (Table 6). This differs from ddA, in which the population of the γ' and γ rotamers are essentially equal⁷⁹.

The overall preferred conformation is given in Fig.44

Table 4 $^2J_{HH}$, $^3J_{HH}$ and $^3J_{HP}$ Coupling Constants (± 0.1 Hz) and Chemical Shifts^a for ddATP.

Temp.(K)	278	283	288	293	299
$^3J_{1'2'}$	2.9	3.0	2.8	3.2	3.0
$^3J_{1'2''}$	6.9	6.9	5.6	6.1	6.5
$^2J_{2'2''}$	-14.0	-14.0	-14.1	-13.5	-14.1
$^3J_{2'3'}$	7.8	7.4	8.0	7.3	8.0
$^3J_{2'3''}$	3.4	3.6	3.4	3.7	3.6
$^3J_{2''3'}$	10.2	10.2	9.6	9.6	9.9
$^3J_{2''3''}$	6.9	6.9	6.7	7.4	7.4
$^2J_{3'3''}$	-12.4	-12.5	-12.8	-12.8	-12.3
$^3J_{3'4'}$	8.9	8.9	8.3	8.7	8.0
$^3J_{3'4''}$	6.5	6.3	6.0	6.4	6.0
$^3J_{4'5'}$	2.3	2.6	2.5	2.6	2.8
$^3J_{4'5''}$	4.3	4.6	4.7	4.7	4.7
$^2J_{5'5''}$	-11.0	-11.4	-11.6	-11.4	-11.7
$^3J_{5'P}$	4.4	5.1	5.0	5.2	5.4
$^3J_{5''P}$	6.3	5.7	6.1	6.3	6.3

^a δ (283K)=H_{1'}(5.9ppm), H_{2'}(2.0ppm), H_{2''}(2.2ppm), H_{3'}(1.7ppm), H_{3''}(1.8ppm), H_{4'}(4.1ppm), H_{5'}(3.8ppm), H_{5''}(3.8ppm)

Table 5
Comparison of the conformation of ddATP with related compounds.

Compound	P_N	ϕ_N	P_S	ϕ_S	X_N	rms (Hz)
ddATP	2.3	36.2	160.0	35.2	.8	.34
2'-deoxyadenosine	9.5	35.3	162.6	34.5	.25	.26 ⁸⁰
3'-deoxyadenosine	6.4	32.9	162.0	35.0	.75	.18 ⁸⁰
ddA	1.4	34.2	160.0	34.2	.7	.45 ⁷⁹
ddl	.1	34.1	161.0	34.1	.7	.40 ⁷⁹

Table 6 Population of γ rotamers of ddATP.

Temp. (K)	% γ'	% γ	% γ''
278	68.9	0	33.1
283	63.0	1.8	35.2
288	62.4	0.7	36.9
293	62.1	1.3	36.6
299	60.3	3.8	35.9

Fig.42 Experimental ^1H NMR of ddATP in D_2O .

Fig. 43a.
Experimental and Simulated ^1H NMR Spectra of ddATP in D_2O at 278K.
Expansion of H_5, H_5' (3.95-3.5 ppm) and $\text{H}_2, \text{H}_2', \text{H}_3, \text{H}_3'$ (2.2-1.5 ppm)

Fig. 43b.

Experimental and Simulated ^1H NMR Spectra of ddATP in D_2O at 283K.
Expansion of H_5 , H_5' (3.95-3.5 ppm) and H_2 , H_2' , H_3 , H_3' (2.2-1.5 ppm)

Fig. 43c.

Experimental and Simulated ^1H NMR Spectra of ddATP in D_2O at 288K.
Expansion of H_5 , H_5' (3.95-3.5 ppm) and H_2 , H_2' , H_3 , H_3' (2.2-1.5 ppm)

Fig. 43d.

**Experimental and Simulated ^1H NMR Spectra of ddATP in D_2O at 293K.
Expansion of H_5 , $\text{H}_{5'}$ (3.95-3.5 ppm) and H_2 , $\text{H}_{2'}$, H_3 , $\text{H}_{3'}$ (2.2-1.5 ppm)**

Fig. 43e.

**Experimental and Simulated ^1H NMR Spectra of ddATP in D_2O at 299K.
Expansion of H_5 , $\text{H}_{5'}$ (3.95-3.5 ppm) and H_2 , $\text{H}_{2'}$, H_3 , $\text{H}_{3'}$ (2.2-1.5 ppm)**

Fig.44 Principal Solution Conformation of ddATP.

Conclusions

In solution, the freedom of motion of molecules interacting with other molecular species is restricted. This results in a longer correlation time⁷⁰. Excluding molecules with very large molecular weights, this increase in the correlation time normally results in a decrease in the NMR spin-lattice relaxation time (T_1) for molecules undergoing primarily dipolar relaxation (Fig.10)⁶⁰. Measurements of T_1 have been reported in studies of the binding of sulfonamides and penicillin to bovine serum albumin⁶⁴, lipid-protein interactions⁶⁵, cellulose-water interactions⁶⁶, calcium-parvalbumin binding⁶⁷, and bilirubin-dilysine interactions⁶⁸.

Retroviral RTs conserve a sequence motif of a hydrophobic residue (i.e. I) followed by at least two basic residues (i.e. KK)⁷⁹ (Fig.45). Mutation of the hydrophobic amino acid, isoleucine, to the hydrophilic amino acid, serine, results in a decrease in polymerase activity to 50% of wildtype⁷⁹. Mutation of the second lysine residue to arginine, an evolutionary conservative mutation, abolishes HIV RT activity. This region has been implicated in AZT resistant mutants of HIV RT indicating it plays a role in nucleotide binding⁷⁹.

		a.a. 60	a.a. 70
1) HIV	326	<u>NPYNTPVFA</u> <u>IKK</u> DSTKRRK	
2) MuTV-pol	42	SPNHTPVFVIKKK	SGKRR
3) VISA	199	HTCNTPIFCIKKK	SGKRR
4) ELAV	243	NPYNSPIFVIKK	RSKRR
5) RSV	55	SCNHTPVFVIKK	ASGSYR
6) MoMuLV	217	SPNHTPLLVPKKPGTMDYRP	
7) HTLV-I	63	GPQNPVFPVKK	AMQTR
8) CAEV	49	HTCNTPIFCIKKK	SGKRR
9) BLV	49	GPQNPVFPVKK	PMGARR
10) HTLV-II	148	GPQNPVFPVKK	PMGARR
11) MPLV	71	SPNHTPIFVIKKK	SGKRR

Fig.45 Amino Acid Sequence Comparison between HIV and 10 other Retroviral RTs. The Sequence under Study is Underlined⁷⁹.

Experimental T_1 data suggest there is both an electrostatic and hydrophobic interaction of ddATP with IKKK and di-L-lysine. The decrease in spin-lattice relaxation time for C_β and C_ϵ of the lysine residues, which are closest to the positively charged amino groups of the lysine side chain leads one to speculate the negatively charged phosphates are interacting with the positively charged amino groups, as indicated by molecular modelling. The spectra were recorded at pH adjusted to 7.8, so as to insure protonation of the amino groups of the lysine side chain ($pK_a=10.8$)⁷².

The changes in spin-lattice relaxation times for two of the carbons on the side chain of isoleucine were extremely large, indicating there is a hydrophobic association with ddATP within an aqueous environment. Hydrophobic interactions are commonly seen in protein folding as a result of inability to hydrogen bond with the polar solvent⁷². A study done by Painter *et al.*, on the initial binding of 2'-deoxynucleotides to HIV RT has revealed via K_d (dissociation constant) calculations that the binding of dNTP to HIV RT is driven mainly by hydrophobic interactions, with electrostatic interactions being of secondary importance³⁵. The Gibbs free energies of binding in standard buffer and in a solution of ionic strength 1M were compared. In a solution of ionic strength 1M, electrostatic interactions are effectively eliminated⁷⁸. The difference in free energy of binding was 14%, which gives an estimate of the contributions of electrostatic interactions to binding. The large decreases in spin-lattice relaxation time seen in the hydrophobic side

chain of isoleucine are either a result of a hydrophobic interaction with ddATP or a result of an interaction between the positively charged N-terminal amino group with ddATP as predicted by molecular modelling. The proximity of the isoleucine side chain to the ribose moiety of ddATP may be enhanced by the electrostatic interactions discussed above. The NOESY spectrum of di-L-lysine and ddATP reveals a connectivity between H_{α} of ddATP and H_{β} as well as H_{γ} of the lysine residues (Fig.35), which supports this theory. Although the T_1 experiments were performed using 1:1 ratios of ddATP and the tetrapeptide, the ratio of the two interacting species remains unclear.

An electrostatic interaction is consistent with results obtained in the binding of MgATP to adenylate kinase⁷⁷. In this study it was found the Mg^{2+} -ATP portion (i.e. Mg^{2+} - PO_4 chelate of ATP) bound near Lys-21 and Lys-27 of adenylate kinase. Interestingly, in adenylate kinase there is an isoleucine residue (Ile-28) adjacent to the proposed triphosphate binding site which forms part of the hydrophobic adenine ribose binding pocket. This supports a hydrophobic interaction of the isoleucine residue with ddATP⁷⁷. As well, the conformation of enzyme bound ATP was reported to be in a conformation similar to that of ddATP when in solution with IKKK (C3'-endo, $\gamma=60^\circ$).

Future Directions

In order to construct an effective nucleoside analogue with limited cellular toxicity, the triphosphate binding geometry of HIV RT must be known. Although this modelling supports the existence of an electrostatic interaction between a conserved sequence on HIV RT with ddATP, a substrate analogue, this does not provide a complete picture of the overall binding and recognition of ddNTP to HIV RT. Proteins in vivo recognize substrates due to their 3 dimensional structures, with many amino acids from different parts of the primary sequence coming together to form a substrate binding pocket. Analysis by NMR of the whole enzyme with ddATP would certainly provide more information about the residues required for the binding of triphosphates to RT.

EXPERIMENTAL

Peptide Synthesis

All chemicals were obtained from Sigma, and used without further purification. All solvents were of reagent grade and distilled prior to use. A bubbler type apparatus was used for sequential synthetic steps (Fig.46).

Fig. 46 Bubbler Apparatus

N-Fmoc-N-Boc-L-Lysine p-alkoxy benzyl resin ester (.5g,.5-1 mmol/ g resin) was swollen for 30 min in dimethylformamide (DMF, 5mL) under argon prior to the acylation steps. After swelling, the resin was washed with DMF (5X15mL) followed by deprotection (15 mL 20% piperdine in DMF, 2x6min). A ninhydrin colour test was performed to ensure the resin was completely

deprotected, with free primary amino groups.

N-Fmoc-N-Boc-L-Lys-OPfp (.3g,.5mmol) and hydroxybenzotriazole (HOBt, .4g, 3mmol) dissolved in DMF (4mL) were added to the deprotected resin. The mixture was agitated by bubbling of argon through the reaction vessel. After 45 min a ninhydrin colour test was performed to determine the completeness of the acylation step. The colour test was negative, indicating a complete reaction. Agitation was continued for a further 30 min to ensure a complete reaction. The resin was then washed with DMF (10x15mL) and deprotected (15 mL 20% piperidine in DMF, 2x6min). The deprotected resin was then washed with DMF (10x15mL) to remove all traces of piperidine. A sample of the deprotected resin gave a positive ninhydrin colour test. A third lysine residue, and an isoleucine residue were added to the peptide chain by an identical procedure as outlined above. The acylations were performed by first using N-Fmoc-N-Boc-L-Lys-OPfp (.3g, .5mmol) for 60 min, followed by deprotection and N-Fmoc-L-Ile-OPfp (.4g,.77mmol) for 60 min again followed by deprotection.

Cleavage

The alkoxybenzyl linkage is cleaved under acidic conditions similar to those required to remove t-butyl-based protecting groups.

The peptide resin was transferred to a sintered glass funnel and washed with DMF, t-amyl alcohol, 6M acetic acid, t-amyl alcohol, DMF and ether and dried in high vacuum. The peptide was cleaved from the resin using aqueous

trifluoroacetic acid (50mL, 95:5;v/v, 2h, RT) in a 100mL round bottomed flask with gentle stirring. The resin was then filtered through a sintered glass filter, and washed with trifluoroacetic acid (4x10mL). The filtrate was evaporated under high vacuum to give the free peptide (.8g). The crude product was purified by gel filtration chromatography (Sephadex G-10) in .1M NH_4Cl . The first 3 10 mL fractions contained the peptide, as determined by A_{210} . The fractions were pooled, and an NMR of the peptide revealed some further impurities. This mixture was then further purified on an analytical reverse phase column (Perkin Elmer US5 C-18) (80% .1 M phosphate buffer containing 20 mM pentane sulfonate as a counter ion and 20% ethanol, .8mL/min) The HPLC trace consisted of four main peaks, the desired product had a retention time of 2.28 min. The mass spectrum of the purified peptide showed a peak at 517m/e corresponding to $(M-1)^+$. ^1H , and ^{13}C spectra of the peptide were recorded on a Bruker AMX 400 spectrometer.

^1H NMR: C_αH 4.1(1H,t), 4.0(1H,q), 3.9(1H,q), 3.6(1H,d), $\text{C}_\epsilon\text{H}$ 2.8(K)(6H,m), $\text{C}_\beta\text{H}(K)$ 1.5(6H,m), $\text{C}_\gamma\text{H}(K)$ 1.2(6H,m), $\text{C}_\delta\text{H}(k)$ 1.4(6H,m), $\text{C}_\beta(I)$ 1.7(1H,m), $\text{C}_\gamma\text{H}_3(I)$ 0.7(3H,d), $\text{C}_\gamma\text{H}_2(I)$ 1.4 (2H,m), $\text{C}_\delta\text{H}(I)$ 0.5(3H,t)

^{13}C NMR:Amide: 179.4, 175.8, 175.7, 171.8, aliphatic: 59.9, 56.2, 56.1, 56.0, 41.6, 38.9, 33.0, 32.9, 28.8, 26.5, 24.6, 24.5, 16.5, 13.0.

Ninhydrin

The ninhydrin colour test indicates the presence of primary amino functions on the resin beads. A blue colour indicates a positive test, a straw yellow colour

a negative test. A small resin sample was removed (5mg) and transferred to a sintered glass funnel and washed under suction with DMF, t-amyl alcohol, 6M acetic acid, t-amyl alcohol, DMF and ether. A few resin beads are transferred to a small test tube and 2-3 drops of the ninhydrin reagent (1g ninhydrin dissolved in 10mL ethanol) were added. The tubes were placed in a 120° for 5 min and the colour compared with a blank sample.

NMR

The ^{13}C T_1 experiments were performed using a Varian Gemini 200 NMR spectrophotometer with the samples contained in 5mm XR55 NMR tubes. The inversion recovery Fourier transform method was used. A 180° pulse was applied nonselectively to all signals to invert the magnetization. Immediately after the pulse, the magnetization was allowed to recover to its equilibrium value over a preselected period of time (.01,.05,.1,.5,1,3,5,10)s. The T_1 values were calculated using the software installed in the Varian instrument. The reported error for the ^{13}C experiments was somewhat large, but within 10%.

T_1 values were determined for Di-L-Lysine in D_2O (10.4 mg, .03 mmol, 20 mM MgCl_2 , pH 8). Identical conditions were used to determine T_1 values for a 1:1 mixture of di-L-Lysine (10.4mg, .03mmol) and ddATP (14.3mg, .03mmol), the tetrapeptide IKKK (20mg, .038 mmol), and a 1:1 mixture of IKKK (20mg, .038mmol) and ddATP (15mg, .032mmol).

REFERENCES

1. Gonda, M.A., Braun, M.J., Clements, J.E., Pyper, J.M., and Wong-Staal, F.W. 1986. *Proc. Natl. Acad. Sci. USA* 83:4007-11.
2. Hasse, A.T. 1986. *Nature* 322:131-36.
3. Wong-Staal, F. 1990. *Virology*, ed. B.N. Fields, D.M. Knipe, R.M. Chanock, J. Melnick, B. Roizman, and R. Slope, 2:1529-93. New York: Raven. 2nd ed.
4. Fenner, F. 1975. *Intervirology* 6:1-12.
5. Weiss, R., Teich, N., Varmus, H., Coffin, J., eds. 1984. *RNA Tumor Viruses*, Vols. 1,2. Cold Spring Harbor: Cold Spring Harbor Lab. 1292 pp., 2344 2nd ed.
6. Baum, R. 1992. *Chemistry and Engineering* 8:26-32.
7. Yashwantrao, N., Vaishnav, V. and Wong-Staal, F. 1991. *Annual Review of Biochemistry* 60:577-630.
8. Yarchoan, R., Berg, G., Brouwers, P., Fischl, M., and Spitzer, A. 1987. *Lancet* 1(8525):132-35.
9. Jones, K.A., Kadonaga, J.T., Luciw, P.A., and Tjian, R. 1986. *Science* 232:755.
10. Nabel, G., and Baltimore, D. 1987. *Nature* 326:711.
11. Klatzman, D., Champagne, E., Chamaret, S., Greust, J. and Greaves, H. 1984. *Nature* 312:767-68.
12. Matthews, T.J., and Bolognesi, D.P. 1988. *Sci. Amer.* 259:120.
13. Majumdar, C., Abbotts, J., Broder, S., and Wilson, S.H., 1988. *J. Biol. Chem.* 263:15657-15665.
14. Hopkins, S., Furman, P.A., and Painter, G.R., 1989. *Biochem. Biophys. Res. Comm.* 163:106-110.
15. Painter, G.R., Wright, L.L., Hopkins, S. and Furman, P.A., 1991. *J. Biol. Chem.* 266:19362-19368.

- 16.Kuchta, R.D., Mizrahi, V., Benkovic, P.A, Johnson, K.A., and Benkovic, S.J., 1987. *Biochemistry*, 26:8410-8417.
- 17.Molina, A.J. and Arnold, E. 1991. *Biochemistry* 30:6351-6361.
- 18.Mitsuya, H. and Broder, S. 1986. *Proc. Natl. Acad. Sci. USA* 83:1911.
- 19.Yarchoan, R., Mitsuya, H., Thomas, R., Pluda, J., Hartman, M., Perno, C., Marczyk, K., Allain, J.P., Johns, D., and Broder, S. 1989. *Science* 245:412.
- 20.Furman, P., Fyfe, J., St.Clair, M., Weinhold, K., Rileout, J., Freeman,G., Lehrman, S., Bolognesi, D., Broder, S.,Mitsuya, H. and Barry, D.W. 1986. *Proc. Natl. Acad. Sci.* 83:8333.
- 21.Schinazi. R., Eriksson,B.F.H., and Hughes, S.H. 1987. *Antimicrob. Agents Chemother.* 31:1972.
- 22.Lightfoote, M., Coligan, J.E., Folks, T.M., fauci, A.S., martin, M.A. and Venkatesan, S. 1986. *J. Virol.* 60:771-775.
- 23.Hizi, A., Hughes, S.H., and Shaharabany, M. 1990. *Virology* 175:575-580.
- 24.Hizi, A., Barber, A. and Hughes, S.H. 1989. *Virology* 170:326-329.
- 25.Goff, S.P. 1990. *J. Aquired Immune Defic. Syndr.* 3:817-831.
- 26.Lowe, D.M., Aitken, A., Bradley, C., Darby G.K., Larder, B., Powell, K., Purifoy, D.J.M., Tisdale, M. and Stammers, D.K. 1988. *Biochemistry* 27:8884-8889.
- 27.Chen, N., Painter, G.R., and Furman, P.A. 1991. *Biochem. Biophys. Res. Comm.* 174:785-789.
- 28.Horwitz, J.P., Chua, J., and Noel, M. 1964 *J. Org Chem.* 29:2076; Robins, M.J., and Robins, R.J. 1964. *J. Am. Chem. Soc.* 86:3585.; Horwitz, J.P., Chua, J., Donatti, J.T. and Noel, M. 1966 *J. Org Chem.* 32:817.; Ono, K., Ogasawara, A., and Matsukage, A. 1979. *Biochem. Biophys. Res. Comm.* 88:1255.; Doering, a.M., Jansen, M., and Cohen, S.S. 1966. *J. Bacteriol.* 92:565.
- 29.Mitsuya, H., and Broder, S. 1987. *Nature* 325:773.

30. McGowan, J. J. 1990. *Rev. Infect. Dis.* 12:513.
31. Fischl, M.A., and Richman, D.D. 1987. *N. Engl. J. Med.* 317:185.
32. Yarchoan, R., Mitsuya, H., Meyer, C.E. and Broder, S. 1989. *N. Engl. J. Med.* 321:726.
33. Zimmerman, T.P., Mahony, W.B., and Prus, K.L. 1987. *J. Biol. Chem.* 262:5748.
34. Pahlic, M., and Tyson, J.J. 1983. *Exp. Cell res.* 144:159-164.
35. Johnson, M.A., Ahluwalia, G., Connelly, M., Cooney, D., Broder, S., Johns, D., and Fridland, A. 1988. *J. Biol. Chem.* 263:15354-15357.
36. Johnson, M.A., and Fridland, A. 1989. *Mol. Pharmacol.* 36:291-295.
37. Mitsuya, H., Yarchoan, R., and Broder, S. 1990. *Science* 249:1533-1543.
38. Van Roey, P., Salerno, J., Chu, C.K., and Schinazi, R.F. 1989. *Proc. Natl. Acad. Sci. USA* 86:3929-3933.
39. Cheng, Y.C., Dutschman, G.E., Bastow, K.F., Sarngadharan, M.G., and Ting, R.Y.C. 1987. *J. Biol. Chem.* 262:2187.
40. Griffin, J.L. 1990. *N. Engl. J. Med.* 322:1098.
41. Izuta, S., Saneyoshi, M., Sakurai, T., Suzuki, M., Kojima, K., and Yoshida, S. 1991. *Biochem. Biophys. Res. Comm.* 179:776-783.
42. Larder, B.A., and Kemp, S.D. 1989. *Science* 246:1155-1158.
43. Larder, B.A., Cheseboro, B., and Richman, D.D. 1990. *Antimicrob. Agents. Chemother.* 34:436.
44. Larder, B.A., Purifoy, D.J.M., Powell, K.L., and darby, G. 1987. *Nature* 327:716.
45. Larder, B.A., Purifoy, D.J.M., and Kemp, S.D. 1989. *Proc. Natl. Acad. Sci. USA* 86:4803.

46. McClure, M.A., Johnson, M.S., Feng, D.F., and Doolittle, R.F. 1988. *Proc. Natl. Acad. Sci. USA* 85:2469-2474.
47. Barber, A.M., Hizi, A., Maizel, J.V., and Hughes, S.H. 1990. *AIDS Res. and Hum. Retrovir.* 6:1061-1071.
48. Basu, A., Nanduri, V.B., Gerard, G.F. and Modk, J. 1988. *J. Biol. Chem.* 261:1648-1653.
49. Kohlstaedt, L.A., Wang, J., Friedman, J.M., Rice, P.A., and Steitz, T.A. 1992. *Science* 256:1783-1790.
50. Boyer, P.L., Ferris, A.L., and Hughes, S.H. 1992. *J. Virol.* 66:7533-7537.
51. Arnold, E., Molina, A.J., Nanni, R.G., Williams, R.L., Lu, X., Ding, J., Clark, A.D., Zhang, A., Ferris, A.L., Clark, P., Hizi, A., and Hughes, S.H. 1992. *Nature* 357:85-89.
52. Furfine, E.S., and Reardon, J.E. 1991. *J. Biol. Chem.* 266:406-412.
53. Sundaralingham, M. 1975. in Structure and Conformation of Nucleic Acid and Protein-Nucleic Acid Interactions, eds. Sundaralingam, M. and Rao, S.T. University Press, Baltimore, pp 487-536.
54. Altona, C., and Sundaralingam, M. 1973. *J. Am. Chem. Soc.* 93:2333.
55. Haasnoot, C.A.G., de Leeuw, F.A.A.M., de Leeuw, H.P.M., and Altona, C. 1981. *Organic Magnetic Resonance* 15:43-52.
56. Huggins, M.L. 1953. *J. Am. Chem. Soc.* 75:4123.
57. Plavec, J. Koole, L.H., Oberg, B., and Chattopadhyaya, J. 1991. *J. Org. Chem.*
58. de Leeuw, F.A.A.M., and Altona, C.J. 1983. *J. Comp. Chem.* 4:438. and PSEUROT: QCPE Program No. 463.
59. Hruska, F.E., wood, D.J., and Ogilvie, K.K. 1974. *Can. J. Chem.* 52:3353.
60. Abraham, R.J., Fisher, J., and Loftus, P. 1988. in Introduction to NMR Spectroscopy John Wiley & Sons, N.Y., N.Y.

61. Pavia, D.L., Lampman, G.M., and Kriz, G.S. 1979. in Introduction to Spectroscopy Saunders College Publishing, Philadelphia, PA, U.S.A.
62. Karplus, M. 1963. *J. Am. Chem. Soc.* 85:2870.
63. Zhu, X.X., and Brown, G.R. 1990. *Can. J. Appl. Spec.* 35:121-127.
64. Jardetzky, O., and Wade-Jardetzky, N.G. 1965. *Mol. Pharmacol.* 1:214-230.
65. Stoffel, W., Zierenberg, O., Tunggal, B.D., and Schreiber, E. 1974. *Physiol. Chem.* 355:1381-1390.
66. Froix, M.F. and Nelson, R. 1976. *Macromolecules* 9:428-430.
67. Nelson, D.J. 1978. *Inorg. Chem. Acta.* 27:L71-L74.
68. Zhu, X.X., Brown, G.R. and St-Pierre, L.E. 1987. *Can. J. Spectr.* 32:49-55.
69. Wurthrich, K. 1989. *Acc. Chem. Res.* 22:36.
70. Borgias, B.A., and James, T.L. 1990. in Biological Magnetic Resonance, Vol. 9. eds. Berliner, L.J. and Reuben, J. Plenum Press, N.Y., N.Y.
71. Merrifield, R.B. 1963. *J. Am. Chem. Soc.* 85:2149.
72. Creighton, T.E. 1984. Proteins, structures and Molecular Principles, W.H. Freeman and Company, N.Y., N.Y.
73. Atherton, E. and Sheppard, R.C. 1989. Solid Phase Peptide Synthesis, IRL Press at Oxford University, England.
74. Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S.S., and Karplus, M. 1983. *J. Comput. Chem.* 4:187-217.
75. Atherton, E., Cameron, L.R., and Sheppard, R.C. 1988. *Tetrahedron*, 44:843-857.
76. Wurthrich, K. 1976. NMR in Biological Research: Peptides and Proteins, North-Holland Publishing Company, Amsterdam, Netherlands.
77. Fry, D.C, Kuby, S.A., and Mildvan, A.S. 1985. *Biochemistry*, 24:4860-4694.

78. Record, M., Lohman, T., and deHaseth, P. 1976. *J. Mol. Biol.* **107**:145-158.
79. Boyer, P.L., Ferris, A.L., and Hughes, S.H. *J. Virol.* **66**:1031-1039.
80. Haasnoot, C.A.G., deLeeuw, F.A.A.M., and Altona, C. 1981. *Org. Mag. Res.* **15**:43-52.

END

2 0 1 0 7 - 9 4

FIN