https://theses.gla.ac.uk/

Theses Digitisation:

https://www.gla.ac.uk/myglasgow/research/enlighten/theses/digitisation/

This is a digitised version of the original print thesis.

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This work cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Enlighten: Theses
https://theses.gla.ac.uk/
research-enlighten@glasgow.ac.uk

# MUTAGENESIS STUDIES OF THE VARICELLA-ZOSTER VIRUS THYMIDYLATE SYNTHASE

by

PATRICK THOMAS HARRISON©

A Thesis Presented for the Degree of
Doctor of Philosophy

in

The Faculty of Science,
University of Glasgow

Institute of Virology,
Church Street,
Glasgow G11 5JR

January 1992

ProQuest Number: 11011431

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

ProQuest 11011431

Published by ProQuest LLC (2018). Copyright of the Dissertation is held by the Author.

All rights reserved.
This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

ACKNOWLEDGEMENTS
SUMMARY
ABBREVIATIONS

## TABLE OF CONTENTS


Page No

1. INTRODUCTION 1

1.1 **Herpesviruses** 1
1.1.1 Virion structure
1.1.2 Genome structure
1.1.3 Molecular and biological properties
1.1.4 Herpesvirus taxonomy
1.1.4.1 Alphaherpesvirinae
1.1.4.2 Betaherpesvirinae
1.1.4.3 Gammaherpesvirinae

1.2 **Molecular biology of varicella-zoster virus** 6
1.2.1 Genome structure and organization
1.2.2 Transcriptional analysis
1.2.3 VZV gene products involved in DNA precursor synthesis
1.2.4 Regulation of gene expression
1.2.5 VZV DNA replication

1.3 **Biology of varicella-zoster virus** 14
1.3.1 Varicella: disease
1.3.2 Varicella: pathogenesis
1.3.3 Zoster: disease
1.3.4 Zoster: pathogenesis
1.3.5 Latency

1.4 **Control of infection** 21
1.4.1 VZV immune serum globulin and zoster immune plasma
1.4.2 Vaccination
1.4.3 Antiviral compounds
1.4.4 Antiviral resistance

1.5 **Thymidylate synthase: an introduction** 26
1.5.1 The thymidylate cycle
1.5.2 TS from different sources
1.5.2.1 Bacteria
1.5.2.2 Bacteriophage
1.5.2.3 Mammals
1.5.2.4 Herpesviruses
1.5.2.5 Protozoa
1.5.2.6 Yeast
1.5.2.7 Plants
1.5.3 Amino acid sequence homology

1.6 **Catalytic mechanism of TS** 35
1.6.1 Identification of catalytic residues
1.6.2 Ordered binding of substrate and cofactor
1.6.3 Conformational changes associated with catalysis: ternary complex formation
1.6.4 Stereospecific reaction mechanism

1.6.5 Reaction mechanism based on X-ray crystallographic data
1.6.5.1 Model of Matthews and co-workers
1.6.5.2 Model of Stroud and co-workers
1.6.5.3 Summary

1.7 **Structural features of TS** 44
1.7.1 X-Ray crystallography
1.7.2 General structural features
1.7.3 Specific features of the large domain
1.7.3.1 ß-kink of strands i, ii and iii
1.7.3.2 Dimer-interface
1.7.3.3 The J-helix
1.7.3.4 Active site cavity

1.8 **Conformational changes during ligand binding: segmental accomodation** 48

1.9 **Structural plasticity and covariant accomodation** 49

1.10 **Models of VZV TS** 51

1.11 **Mutational analysis of TS** 52
1.11.1 Site-directed mutagenesis
1.11.2 Saturation site-directed mutagenesis

1.12 **TS inhibitors** 58
1.12.1 Nucleoside analogues requiring activation
1.12.2 Cofactor analogues
1.12.3 Bisubstrate analogues
1.12.4 Nucleoside analogues not requiring activation
1.12.5 *In vivo* regulators
1.12.6 Anti-sense oligonucleotides
1.12.7 Oligopeptide inhibition of TS

1.13 **Nucleoside and nucleoside analogue metabolism** 64
1.13.1 Mammalian systems
1.13.2 Bacterial systems

1.14 **Resistance to TS inhibitors** 69
1.14.1 Mammalian systems
1.14.2 Bacterial systems

1.15 **Aims of the project** 72

**2. MATERIALS AND METHODS** 73

2.1 **Bacteria and bacteriophage** 73
2.1.1 Bacterial strains
2.1.2 Bacteriophage

2.2 **Plasmids**
2.2.1 Plasmids provided 74
2.2.2 Plasmid constructs

2.3 **Bacterial growth media** 76
2.3.1 Liquid media
2.3.2 Solid media

2.4 **Chemicals, drugs, radiochemicals and enzymes** 76
2.4.1 Chemicals
2.4.2 Drugs
2.4.3 Radiochemicals
2.4.4 Enzymes

2.5 **Commonly used buffers** 78

2.6 **Growth and maintenance of bacteria and bacteriophage** 78
2.6.1 Bacteria
2.6.1.1 Laboratory stocks of *E.coli*
2.6.1.2 New strains of *E.coli*
2.6.1.3 Overnight cultures
2.6.1.4 Glycerol stocks
2.6.1.5 Antibiotics
2.6.2 Bacteriophage
2.6.2.1 Titration of helper phage
2.6.2.2 Growth of helper phage

2.7 **Bacterial manipulation** 80
2.7.1 Transformation
2.7.1.1 Calcium chloride
2.7.1.2 DMSO-PEG4000
2.7.2 Generation of *thyA* strains of *E.coli*

2.8 **DNA Isolation** 81
2.8.1 Mini-prep plasmid DNA preparation
2.8.2 Midi-Prep Plasmid DNA preparation
2.8.3 Maxi-prep plasmid DNA preparation
2.8.4 ssDNA preparation from phagemid vectors
2.8.4.1 Midi-scale ssDNA preparation
2.8.4.2 Maxi-scale ssDNA preparation
2.8.5 Plasmid DNA sequencing preparations

2.9. **DNA Manipulations** 85
2.9.1 Restriction enzyme digests
2.9.2 Removal of 5'-phosphate groups
2.9.3 Isolation of DNA fragments from agarose gels
2.9.3.1 Electroelution
2.9.3.2 Low melting point agarose
2.9.3.3 Use of NA45 ion-exchange membrane
2.9.4 Conversion of 5'-overhangs to blunt ends
2.9.5 DNA ligation
2.9.6 Gapped duplex DNA (gdDNA) synthesis

2.10 **Oligonucleotides** 88
2.10.1 Synthesis and purification
2.10.2 Synthesis of spiked oligonucleotides
2.10.3 Radioactive 5'-end labelling of oligonucleotides
2.10.4 Oligonucleotide sequences

2.10.4.1 Sequencing oligos
2.10.4.2 Site-directed mutagenesis oligos
2.10.4.3 Spiked oligos

2.11 **Mutagenesis** 92
2.11.1 Chemical mutagenesis
2.11.1.1 Ethyl methanesulphonate
2.11.1.2 Hydroxylamine
2.11.1.3 Sodium bisulphite
2.11.2 Growth in a mutator strain
2.11.3 Oligonucleotide directed mutagenesis
2.11.4 Misincorporation mutagenesis

2.12 **Mutant screening** 96
2.12.1 Growth complementation assays
2.12.1.1 Thymidylate synthase
2.12.1.2 Thymidine kinase
2.12.2 Drug resistance screening protocols
2.12.2.1 Method 1
2.12.2.2 Method 2
2.12.2.3 Method 3
2.12.2.4 Method 4
2.12.2.5 Method 5
2.12.3 Slot blot hybridization

2.13 **Mutant characterization** 99
2.13.1 DNA Sequencing
2.13.1.1 Sequencing solutions
2.13.1.2 Sequencing reactions
2.13.1.3 Polyacrylamide gel electrophoresis
2.13.2 TS Extracts
2.13.2.1 Extract preparation
2.13.2.2 Protein concentration determination
2.13.3 TS Assays
2.13.3.1 Tritium release assay
2.13.3.2 Debromination assay
2.13.4 Characterization of FUdR resistant TS variants
2.13.5 TK extracts
2.13.6 TK assays

2.14 **Computing and molecular modelling** 106

3. **RESULTS AND DISCUSSION** 107

3.1 **Drug resistance studies**
3.1.1 Initial experiments 107
3.1.2 New drug screening protocols
3.1.2.1 The use of FUdR rather than FU
3.1.2.2 A *thyA* derivative of the *E.coli* strain SØ928
3.1.2.3 Co-expression of VZV TK in SØ928*thyA*
3.1.2.4 Drug evaluation
3.1.3 New Mutagenesis Protocols
3.1.3.1 ssDNA and plasmid constructs

3.1.3.2 Strand selection procedure
3.1.4 Misincorporation mutagenesis
3.1.4.1 Use of the Klenow fragment of DNA polymerase I
3.1.4.2 Use of reverse transcriptase
3.1.5 Spiked oligonucleotide mutagenesis
3.1.6 Screening clones for drug resistance generated by misincorporation mutagenesis
3.1.6.1 Success of drug selection scheme
3.1.6.2 Wide scale screening using methods 4 and 5
3.1.6.3 Summary and statistical analysis
3.1.7 Screening clones generated by spiked oligo mutagenesis
3.1.7.1 Statistical analysis
3.1.7.2 Summary
3.1.8 Site-directed mutagenesis studies
3.1.8.1 Mutant construction and enzyme activities
3.1.8.2 Molecular modelling
3.1.8.3 Summary

3.2 **Mutants created by random mutagenesis** 133
3.2.1 Mutant enzyme activities
3.2.2 Modelling studies
3.2.2.1 I41L
3.2.2.2 N54K
3.2.2.3 I96M
3.2.2.4 S102F
3.2.2.5 D174A/174H175
3.2.2.6 V179E
3.2.2.7 C183/Q188K
3.2.2.8 C187G
3.2.2.9 F189L
3.2.2.10 S197T/C198R
3.2.3 Summary

3.3 **Studies of Key Structural Regions of TS** 141
3.3.1 ß-Strand i
3.3.1.1 Mutant construction and enzyme activities
3.3.1.2 Predicted role of G31 in *E.coli* TS
3.3.1.3 A glycine residue at postition 48 is not essential for VZV TS activity
3.3.1.4 The additional residue in ß-strand i of yeast TS and a VZV TS mutant is accomodated by a ß-bulge
3.3.1.5 Identification of a covariant insertion adjacent to the ß-bulge in ß-strand i
3.3.1.6 Summary

3.3.2 Carboxy terminus
3.3.2.1 Mutant construction and

enzyme activities
3.3.2.2 The C-terminal interactions of *L.casei* and *E.coli* TS are essentially identical
3.3.2.3 The C-terminal interactions of VZV and *L.casei* are significantly different
3.3.2.4 Comparison of the C-terminal interactions of VZV and *L.casei* at a molecular level
3.3.2.5 Differences in activity of the VZV and *L.casei* TS C-terminal deletion mutants explained in molecular terms
3.3.2.6 The C-terminal interactions in VZV TS are unique
3.3.2.7 TS activity and modelling of the mutant A300E/L301Δ
3.3.2.8 Summary

3.4 **Thymidine Kinase** 155
3.4.1 Site-directed mutagenesis of VZV TK
3.4.1.1 Mutant construction and enzyme activities
3.4.1.2 Summary
3.4.2 Evolution of the herpesvirus thymidine kinase
3.4.2.1 Herpesvirus TKs are genetically and biochemically distinct from their cellular counterparts
3.4.2.2 Channel catfish virus TK: the most divergent herpesvirus TK sequence known?
3.4.2.3 Evolutionary source of herpesvirus TK is the cellular deoxycytidine kinase

4. **FINAL DISCUSSION** 163

4.1 **Drug resistance studies** 163
4.1.1 Original aims and rationale
4.1.2 A drug resistant variant of VZV TS requires more than one nucleotides substituition
4.1.3 Possibilities for future study

4.2 **VZV TS as a possible target for antiviral chemotherpy: identification of an exploitable difference for rational drug design** 166

4.3 **Plastic and covariant accomodation: the ability of a protein to tolerate amino acid substitutions** 167
4.3.1 Globins, *lac* and lambda repressor, and T4 lysozyme
4.3.2 Relationship of our studies of VZV TS to previous studies in other proteins

4.4 **Conservation of the TS primary structure** 172
4.4.1 The primary structure of TS is highly conserved
4.4.2 Why is the primary structure of TS so highly conserved

5. APPENDIX 175

A reappraisal of the *thy1* gene product of *Dictyostelium discoideum*

REFERENCES

ACKNOWLEDGEMENTS

I thank Professor John H. Subak-Sharpe for providing me with opportunity to work and make use of the facilities at the Institute of Virology.

I am indebted to Dr. Russell Thompson for his supervision over the past four years and for giving me the opportunity to explore some of my more unorthodox ideas. I am also grateful for his assistance with photography and tireless proof-reading of my thesis.

I would like to thank Dr. Graham Darby for allowing me to conduct part of my studies at the Wellcome Research Laboratories and Dr. Peter Collins and Dr. Denise Lowe for the support, encouragement and supervision during this period.

I am grateful to Professor Neil Isaacs for allowing me access to the computer facilities in the Chemistry Department of Glasgow University, and to Dr. Robert Stroud and Dr. Celia Schiffer for generating the model structures of VZV TS and providing the co-ordinates of *Escherichia coli* TS.

I thank Dr. Andrew Davison for access to unpublished sequencing data, Dr. John McLauchlan for synthesizing the large number of oligonucleotides used in this project and Simon Cook for assistance and patience with computing.

Special thanks to Jim Scott for invaluable help, assistance and patience over the past few years, without which this project would have been infinitely more difficult.

I would like to acknowledge the friendship and am grateful for the encouragement I have received from other members of Lab 201, past and present.

I am also thankful to all those people who have shared

the good bits, and endured the not-so-good bits, and put up with my moans when most other folk wouldn't. These include Valgerdur, Martina, Janice, Ruth, Gordon, John M., John E., Joe. and Jane. Apologies, but no less thanks to those I've missed.

The 'Hungry Hairy Hippos' of the Wellcome rugby XV also deserve a mention as they'll probably never get mentioned anywhere else.

**This thesis is dedicated to my parents. I would like to acknowledge their considerable sacrifices, constant support and financial assistance over many years. Without this, my studies would have been as near to impossible, as I believe exists.**

The author was the recipient of a C.A.S.E. research studentship from the Science and Engineering Research Council, and the Wellcome Research Laboratories. Unless otherwise stated, all the results described in this thesis were obtained by the author's own efforts.

## SUMMARY

Varicella-zoster virus (VZV), the causative agent of chickenpox and shingles, encodes approximately 70 different proteins including the enzyme thymidylate synthase (TS). TS catalyses the reductive methylation of deoxyuridylate to form thymidylate using $N^5,N^{10}$-methylene-tetrahydrofolate as a cofactor. The enzyme plays a pivotal role in the provision of an essential precursor for DNA synthesis. As such it has been used as a target for antimetabolites which are effective against a variety of infectious and proliferative diseases.

The work presented in this thesis describes the establishment of two random mutagenesis protocols suitable for the generation of complete mutant DNA libraries of the VZV TS gene. Use of these libraries, in conjunction with a novel screening protocol, has shown that a single nucleotide change is not sufficient to give rise to a VZV TS variant resistant to the potent inhibitors 5-fluoro-2'-deoxyuridine-monophosphate (FdUMP) and 5-ethynyl-2'-deoxy-uridine-monophosphate (EYdUMP). This is in contrast to a variant of human TS which contains a single nucleotide substitution that has been proposed to confer resistance to FdUMP. The introduction of the equivalent nucleotide substitution into the VZV TS gene by site-directed mutagenesis (which resulted in the equivalent amino acid change) did not result in a drug resistant viral enzyme.

The mutant library was used to isolate variants of VZV TS that lack catalytic activity. Molecular modelling of the effects of the amino acid substitutions that gave rise to these inactive variants, extended our understanding of the plasticity of protein structure.

As a complementary approach to investigating the structure and function of TS, site-directed mutagenesis was used to study the effects of amino acid changes in two specific regions of the protein; one buried and one exposed to solvent. In the buried region it was shown that the enzyme is more flexible to amino acid substitutions than previously expected, and the mechanism of accomodation of at least one of the changes was identified.

A study of the solvent exposed C-terminal region identified structural and functional differences between VZV TS and *Lactobacillus casei* TS. The two enzymes responded differently to amino acid changes, and the VZV enzyme relied more heavily on hydrophobic contacts to maintain a functional conformation. Phylogenetic comparisons and molecular modelling suggested that the VZV TS may be unique in this regard. Exploitation of the difference between the VZV TS and the human enzyme may allow the rational design of VZV-specific inhibitors and such compounds may have a role as anti-varicella drugs.

It was shown that the expression of the active VZV thymidine kinase (TK) was necessary for the phosphorylation of 5-ethynyl-2'-deoxyuridine to a form that would inhibit TS. This led to the identification of an amino acid essential for VZV TK activity in a region of the enzyme not previously thought to play a role in catalysis.

The evolutionary source of the herpesvirus TKs has been hitherto obscure. A comparison of the conserved amino acid regions of herpesvirus TKs with cellular deoxycytidine kinase revealed a significant degree of homology. In addition, the herpesvirus TKs and deoxycytidine kinase share common biochemical properties. Taken together, this led to the proposal that the evolutionary source of the herpesvirus TKs was the cellular deoxycytidine kinase.

## ABBREVIATIONS

| | |
|---|---|
| Ac | acetate |
| ACV | acyclovir |
| AMP | ampicillin |
| APS | ammonium persulphate |
| araA | adenosine arabinoside |
| araC | cytosine arabinoside |
| BAT | 5'-bromoacetamido-2'-deoxythymidine |
| BUdR | 5-bromo-2'-deoxyuridine |
| BVaraU | 5-bromo-ß-D-arabinofuranosyldeoxyuridine |
| BVdU | 5-bromovinyl-2'-deoxyuridine |
| BrdUMP | 5-bromo-2'-deoxyuridylate |
| BSA | bovine serum albumin |
| CAT | 5'-chloroacetamido-2'-deoxythymidine |
| CB3717 | $N^{10}$-propargyl-5,8-dideazatetrahydrofolate |
| CM | chloramphenicol |
| dCK | deoxycytidine kinase |
| ddC | dideoxycytidine |
| ddNTP | dideoxynucleotide-triphosphate |
| DHFR | dihydrofolate reductase |
| dsDNA | double stranded DNA |
| dUTPase | deoxyuridylate pyrophosphatase |
| DTT | dithiothreitol |
| EDTA | sodium ethylenediamine tetra-acetic acid |
| EMS | ethyl methanesulphonate |
| EtBr | ethidium bromide |
| EtOH | ethanol |
| EYaraU | 5-ethynyl-ß-D-arabinofuranosyldeoxyuridine |
| EYUdR | 5-ethynyl-2'deoxyuridine |
| FIAC | 1-(2-deoxy-2-fluoro-ß-D-arabinofuranosyl)-5-methyluracil |
| FMAU | 1-(2-deoxy-2-fluoro-ß-D-arabinofuranosyl)-5-iodo-cytosine |
| FdUMP | 5-fluoro-2'-deoxyuridylate |
| FU | 5-fluorouracil |
| FUdR | 5-fluoro-2'-deoxyuridine |
| FUR | 5-fluorouridine |
| gp | glycoprotein |
| G-3-P | glyceraldehyde-3-phosphate |

| | |
|---|---|
| HEPES | 4-(2-hydroxyethyl)-1-piperazine ethane |
| IAT | 5'-iodoacetamido-2'-deoxythmidine |
| IE | immediate early |
| IUdR | 5-iodo-2'-deoxyuridine |
| LB | Luria-Bertani medium |
| MOPS | 3-(*N*-morpholino)-propanesulphonic acid |
| MTX | methotrexate |
| $NO_2$-dUMP | 5-nitro-2'-deoxyuridylate |
| OD | optical density |
| ORF | open reading frame |
| PABA | para-aminobenzoic acid |
| PEG | polyethylene glycol |
| PLP | pyridoxal phosphate |
| PYaraU | 5-propargyl-ß-D-arabinofuranosyl-2'-deoxyuridine |
| PYUdR | 5-propynyl-2'-deoxyuridine |
| rpm | revolutions per minute |
| RR | ribonucleotide reductase |
| RT | reverse transcriptase |
| SDS | sodium dodecylsulphate |
| (S)-HPMPA | (S)-9-(3-hydroxy-2-phosphonylmethoxy-propyl)-adenine |
| ssDNA | single stranded DNA |
| TEMED | *N*,*N*,*N'*,*N'*-tetramethylethylenediamine |
| TET | tetracycline |
| TFT | trifluorothymidine |
| TK | thymidine kinase |
| TMP | trimethoprim |
| Tris | tris(hydroxymethyl)aminomethane |
| TS | thymidylate synthase |
| UV | ultraviolet |
| wt | wild type |
| α-thio dNTP | 2'-dNTP-5'-O-(1-thiotriphosphate), sodium salt |
| ß-ME | ß-mercaptoethanol |
| Ø | phenol |
| Y | pyrimidine |

**HERPESVIRUSES**

| | |
|---|---|
| BHV-1 | bovine herpesvirus type 1 |
| BHV-2 | bovine herpesvirus type 2 |
| CCV | channel catfish virus |
| EBV | Epstein-Barr virus |
| EHV-1 | equine herpesvirus type 1 |
| EHV-4 | equine herpesvirus type 4 |
| FHV | feline herpesvirus |
| HCMV | human cytomegalovirus |
| HHV-6 | human herpesvirus type 6 |
| HHV-7 | human herpesvirus type 7 |
| HSV-1 | herpes simplex virus type 1 |
| HSV-2 | herpes simplex virus type 2 |
| HVA | herpesvirus aotus |
| HVT | turkey herpesvirus |
| HVS | herpesvirus samirii |
| ILTV | infectious laryngotracheitis virus |
| MCHV | murine cytomegalovirus |
| MDV | Marek's disease virus |
| MHV | marmoset herpesvirus |
| PRV | pseudorabies virus |
| VZV | varicella-zoster virus |

## AMINO ACIDS

| Symbol | 3-letter | Meaning | Codons |
|---|---|---|---|
| A | Ala | Alanine | GCA,GCC,GCG,GCT |
| C | Cys | Cysteine | TGC,TGT |
| D | Asp | Aspartate | GAC,GAT |
| E | Glu | Glutamate | GAA,GAG |
| F | Phe | Phenylalanine | TTC,TTT |
| G | Gly | Glycine | GGA,GGC,GGG,GGT |
| H | His | Histidine | CAC,CAT |
| I | Ile | Isoleucine | ATA,ATC,ATA |
| K | Lys | Lysine | AAA,AAG |
| L | Leu | Leucine | TTG,TTA,CTA,CTC,CTG,CTT |
| M | Met | Methionine | ATG |
| N | Asn | Asparagine | AAC,AAT |
| P | Pro | Proline | CCA,CCC,CCG,CCT |
| Q | Gln | Glutamine | CAA,CAG |
| R | Arg | Arginine | CGA,CGC,CGG,CGT,AGA,AGG |
| S | Ser | Serine | TCA,TCC,TCG,TCT,AGC,AGT |
| T | Thr | Threonine | ACC,ACC,ACG,ACT |
| V | Val | Valine | GTA,GTC,GTG,GTT |
| W | Trp | Tryptophan | TGG |
| Y | Tyr | Tyrosine | TAC,TAT |

# 1. INTRODUCTION

This introduction will review aspects of the biology and molecular biology of herpesviruses, with special reference to varicella-zoster virus. In addition, it will provide some background information on the enzyme thymidylate synthase.

## 1.1 Herpesviruses

Membership in the family Herpesviridae is based primarily, though not exclusively, on virion architecture (Roizman, 1990). Indeed, herpesviruses have been defined as viruses of eukaryotes, having single, linear, double-stranded DNA genomes of greater than, or equal to, 100kb. They are replicated and packaged into icosahedral nucleocapsids of 100nm diameter, with 162 prismatic capsomers within the nuclei of infected cells. These are then enclosed in protein teguments and glycoprotein and lipid envelopes to give their infectious extracellular forms (Honess & Watson, 1977; Roizman, 1982; Honess, 1984). Nearly 100 herpesviruses have been characterized, including seven human herpesviruses (Roizman, 1990; Frenkel *et al.*, 1990).

### 1.1.1 Virion structure

The herpesvirus virion is 120 to 200nm in diameter and consists of four structural components, the core, capsid, tegument and envelope (reviewed in Dargan, 1986). The precise number of viral polypeptides in herpesvirus virions is not known but estimates generally vary from 18 to 33. The core, which lies at the centre of the virion, consists of the viral genome. The widely accepted proposal that the DNA was wound around a cylindrical or toroidal proteinaceous plug (Furlong *et al.*, 1972; Nazerian, 1974; Nii & Yasuda, 1975; Irmiere & Gibson, 1983) has recently been challenged by Booy and co-workers (1991). They analysed the organization of DNA within the herpes simplex virus type 1 (HSV-1) capsid by cryoelectron microscopy and image

reconstruction and demonstrated the presence of a uniformly dense ball of DNA. This ball of DNA extends radially as far as the inner surface of the icosahedral shell, thus negating the presence of a protein plug. The absence of a candidate protein to form this plug, as reflected by the identical protein profiles of A- and C-capsids (Booy *et al.*, 1991), the similarity to the observed DNA organisation of several large dsDNA bacteriophages (Lepault *et al.*, 1987; Black, 1989) and biophysical studies of DNA packaging (Rau *et al.*, 1984; Richter & Dubochet, 1990) give further weight to this proposal. The suggestion by Friedman and co-workers (1975) that the toroid/cylinder morphology may represent a developmental intermediate in mature capsid formation, with a uniformly staining species of particle representing the mature capsid, seems to reconcile the two sets of observations.

Capsid morphology is a characteristic feature of the Herpesviridae. The capsid exhibits 2-, 3- and 5-fold symmetry and comprises 162 capsomers. There are 150 hexameric prisms and 12 pentameric prisms that are arranged in the form of an icosahedron (Wildy *et al.*, 1960). Recent studies using the techniques of cryoelectron microscopy and image reconstruction have significantly increased our knowledge of herpesvirus capsid structure (Schrag *et al.*, 1989; Baker *et al.*, 1990; Booy *et al.*, 1991).

In 1974, Roizman and Furlong used the term tegument to describe the structures between the capsid and the envelope. This region has no distinct features in thin section but appears fibrous on negative staining (Morgan *et al.*, 1968). The tegument is frequently distributed assymetrically and its overall bulk seems to be virus rather than host determined (McCombs *et al.*, 1971).

When visualised by electron micoscopy, the virus envelope has a trilaminar appearance (Epstein, 1962). It appears to be derived from the infected cell nuclear membrane and contains a variety of virus specific glycoproteins (Stannard *et al.*, 1987).

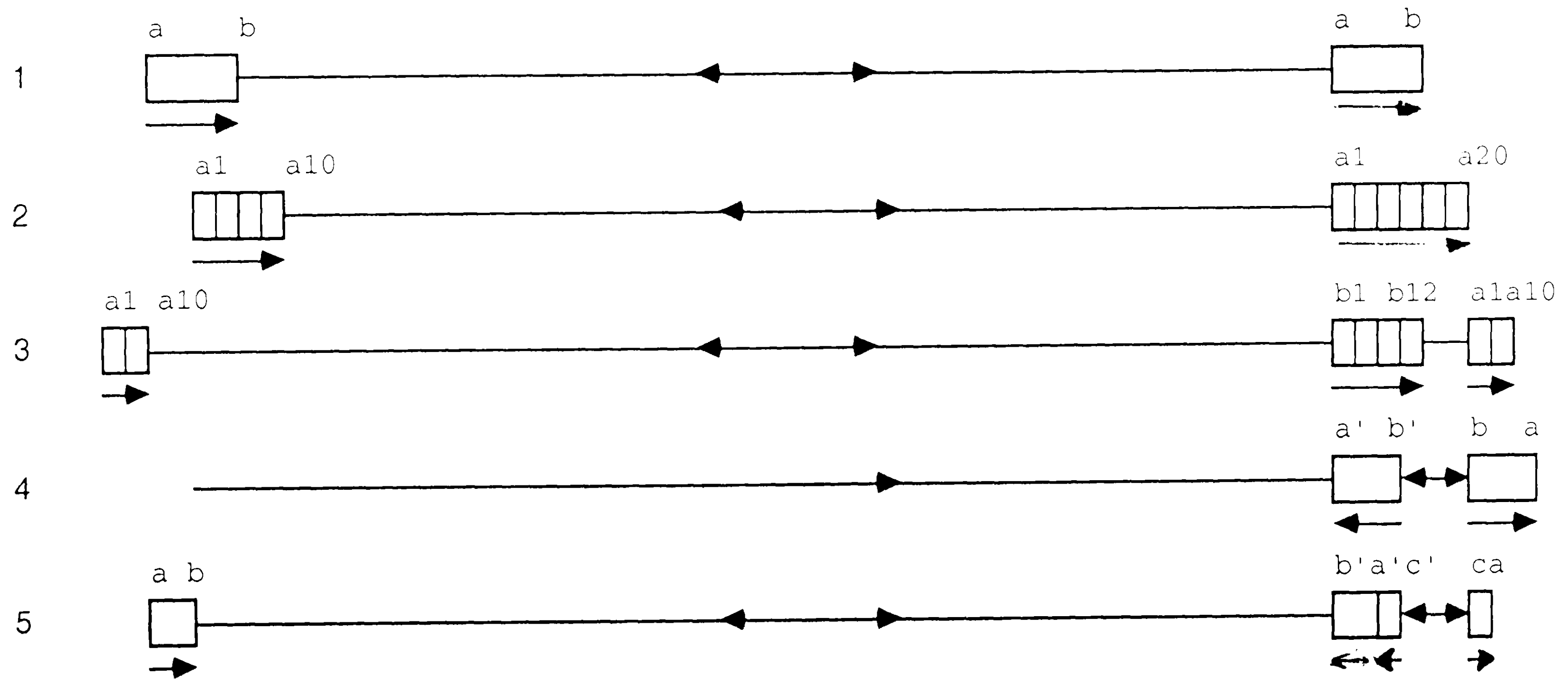



Fig. 1.1. **Herpesvirus Genome Structure**.
The genomic arrangements of representative herpesviruses from the five groups delineated in terms of the presence and location of reiterations in excess of 100bp. In each case the repeat sequences are shown by rectangles and their relative orientation by arrows and the letters abc or their complement a'b'c'. The sequences a1-a10 or a20, and b1-b12 signify multiple tandem repeats. The use of the letters does not denote similarity of sequence, size or function in these regions between different genomes. The representative virus for each group (and other viruses that fall within each group) are as follows. 1 CCV (HHV-6); 2 HVS (HVA); 3 EBV; 4 VZV (EHV-1, PRV); 5 HSV-1 (HCMV, HSV-2). Arrows indicate the relative orientations of various regions of the genome. A single arrow indicates that only one orientation is observed; two arrows indicate that both orientations are seen.

### 1.1.2 Genome structure

The complete DNA sequence of at least six herpesviruses have been determined. These are, Epstein Barr virus (EBV; Baer *et al.*, 1984), varicella-zoster virus (VZV; Davison & Scott, 1986), herpes simplex virus type 1 (HSV-1; McGeoch *et al.*, 1988), human cytomegalovirus (HCMV; Chee *et al.*, 1990), channel catfish virus (CCV; Davison, 1992) and equine herpesvirus type 1 (EHV-1; E. Telford, personal communication). Moreover the genome layout of many more herpesviruses is known. When analysed in terms of the presence and location of reiterations in excess of 100bp, they can be arranged into five groups (Roizman *et al.*, 1981; Dargan, 1986; McGeoch, 1989; Chee & Barrell, 1990; see Fig. 1.1).

### 1.1.3 Molecular and biological properties

Many molecular and biological properties are common to all herpesviruses. Viral replication commences with the attachment of the viral envelope to receptors on the plasma membrane of the host cell, followed by fusion with the membrane and subsequent release of the capsid.

Virus DNA is replicated in the nucleus and unit length DNA is cleaved from concatemers and spooled into preformed, immature capsids which mature by the acquisition or processing of proteins that bind to the surface of the capsid. Virus DNA is also transcribed in the nucleus and the mRNA is translated in the cytoplasm (Roizman, 1991).

The virus develops the ability to become infectious when the capsids are enveloped by budding through the inner lamella of the nuclear membrane. The virus accumulates in the perinuclear space and cisternae of the endoplasmic reticulum, prior to the release of virus particles via transport to the cell surface through modified endoplasmic reticulum structures bounded by cytoplasmic membranes.

In terms of common biological properties, each herpes-virus has its own natural and experimental host range, with both warm and cold-blooded vertebrates and invertebrates

being hosts for herpesviruses. Some herpesviruses can induce neoplasias in both their natural hosts and experimental animals.

Herpesvirus transmission is usually by contact between moist mucosal surfaces. Some herpesviruses can be transmitted transplacentally, via breast milk or by transfusions. Herpesviruses can remain latent in their primary hosts for the lifetime of these hosts, the cells harbouring the latent virus vary depending upon the virus (Honess, 1984; Clements, 1990; Roizman, 1990. 1991).

### 1.1.4 Herpesvirus taxonomy

The similarities, and the differences, in the molecular and biological properties of the herpesviruses, have been exploited for the purpose of classification, with the result that the members of the Herpesviridae are divided into three subfamilies (Roizman, 1982, 1991; Matthews, 1982).

#### 1.1.4.1 Alphaherpesvirinae

The host range of the Alphaherpesvirinae is variable in that some viruses have a very narrow host range, whereas others can have a very wide host range. For example HSV-1, naturally infects only one organism, namely humans, whereas other members, such as pseudorabies virus (PRV), can infect a wide range of organisms (Trainer & Karstad, 1963; Kirkpatrick *et al.*, 1980; Gustafson, 1981). Alpha herpesviruses have a relatively rapid growth cycle (less than 24h), their rapid spread in tissue culture is associated with a high level of cell destruction. In addition, latent infections are frequently, though not exclusively, established in sensory and autonomic ganglia. The DNA genome ranges in length from 120 to 180kb, with sequences from the termini present in an inverted form internally. The genome packaged in virions may consist of two or four isomeric forms. This subfamily includes the type species HSV-1, as well as HSV-2, VZV, EHV-1 and PRV (Roizman, 1991).

### 1.1.4.2 **Betaherpesvirinae**

The Betaherpesvirinae have a restricted host range. *In vivo* the host range is restricted to the species or genus to which the host belongs; *in vitro* replication is usually restricted to a specific cell type, although exceptions do exist. The Betaherpesvirinae are also distinguished by a relatively slow reproductive cycle (longer than 24h), with slowly progressing lytic foci observed in tissue culture. This subfamily includes the type species human (HCMV) and murine (MCMV) cytomegaloviruses which take their name from the observation that infected cells frequently become enlarged (cytomegalia). Other members include the more recently characterized human herpesviruses 6 (HHV-6; Lawrence *et al.*, 1990) and 7 (HHV-7; Frenkel *et al.*, 1990). Latent infections are predominantly established in the secretory glands, lymphoreticular cells and the kidneys. The DNA genome ranges in size from 180 to 250kb and sequences from either or both termini may be present in an inverted form internally.

### 1.1.4.3 **Gammaherpesvirinae**

The members of this subfamily only infect their natural host and cannot infect experimental animals outside the family or order. The length of the replication cycle is variable. These viruses are specific for B- or T-lymphocytes, and infection is frequently arrested at a prelytic stage, with persistence and minimal expression of the viral genome (latent infection), or at a lytic stage, causing cell death without the production of infectious virus progeny. *In vitro* they replicate in lymphoblastoid cells. The DNA genome varies in length from 100 to 170kb with both ends of the molecule containing reiterated sequences that are not reiterated internally (Roizman, 1991). Members of the Gammaherpesvirinae include the type species EBV, herpesvirus saimiri (HVS) and herpesvirus ateles (HVA).

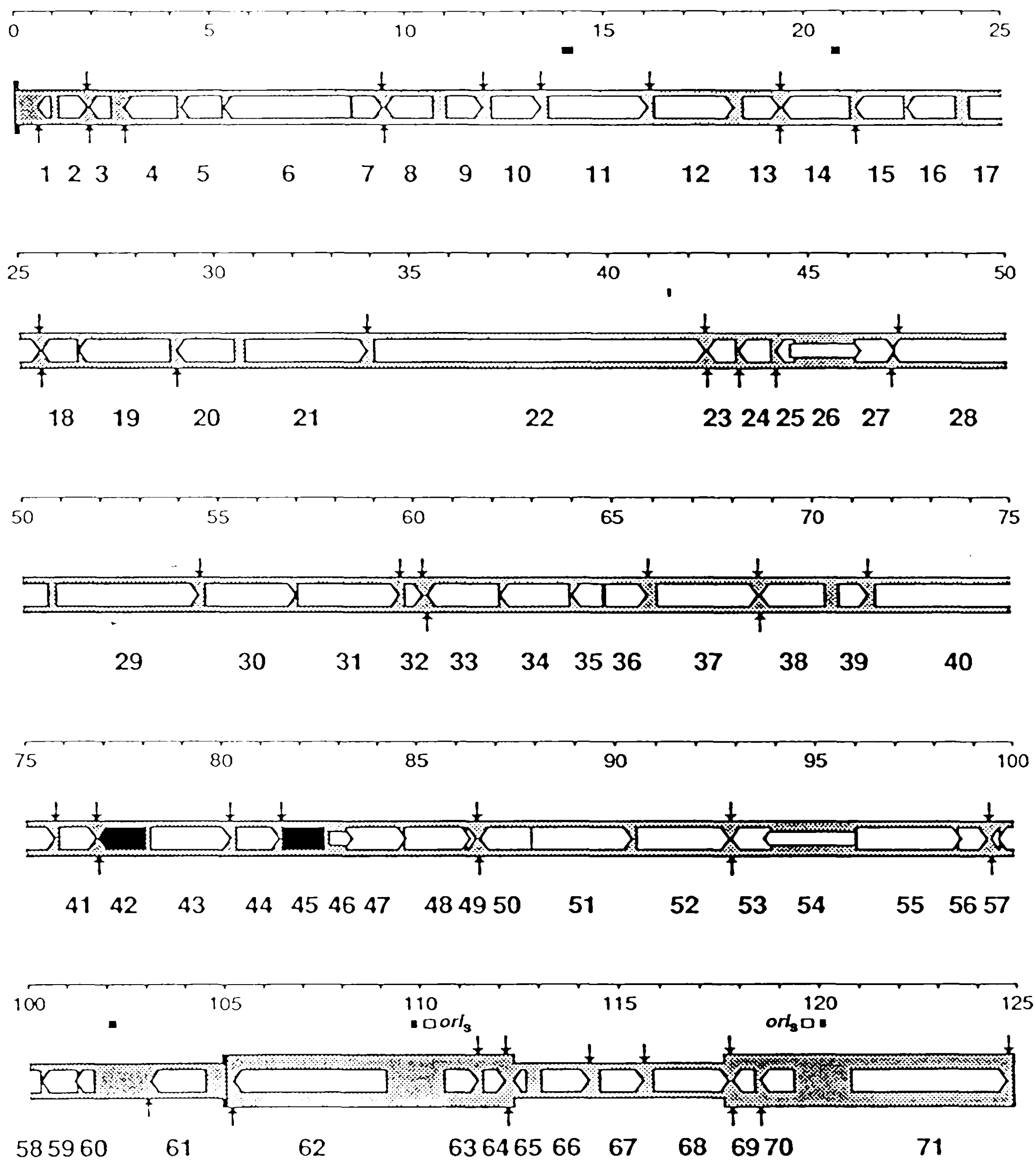



Fig. 1.2. **Genome layout of VZV.**
VZV gene arrangement as deduced from the DNA sequence. ATG-initiated open reading frames (ORFs) predicted to encode proteins are shown as open arrows. ORFs 42 and 45 are shown in black and are predicted to form a spliced message. Vertical arrows indicate potential polyadenylation sites in the appropriate strand. The origins of viral DNA replication are indicated by open boxes; the reiterations are shown by filled boxes (see text for details). The scale in kb is also shown (adapted from Davison, 1991).

## 1.2 Molecular biology of varicella-zoster virus

VZV is an extremely cell-associated virus that grows relatively rapidly in tissue culture cells. However, a method for preparing high titre cell-free virus is unavailable, nearly 40 years after VZV was first isolated (Weller, 1953). This, together with the lack of a suitable animal model, has, until recently, resulted in a very limited understanding of VZV genetics and gene expression. A major step to bringing the understanding of the molecular biology of VZV into line with that of HSV, was the determination of the complete genome sequence by Davison and Scott (1986).

### 1.2.1 Genome structure and organization

The genome of the Dumas strain of the VZV (Dumas *et al.*, 1981) consists of 124,884bp (Davison & Scott, 1986). Of the 71 predicted ATG-initiated open reading frames (ORFs), three are duplicated and two (ORFs 42 and 45) are probably expressed as a spliced mRNA. Thus, there are potentially 67 unique genes, containing from 71 to 2,763 codons, that are arranged compactly on both strands of the DNA and do not extensively overlap (Davison & Scott, 1986). The locations of 48 predicted polyadenylation signals are shown in Fig. 1.2. These were identified by the presence of the polyadenylation signal sequence, AATAAA (or ATTAAA), which in most cases, was located upstream of the consensus sequence, YGTGTTYY, a conserved element located 25 to 40bp downstream from the polyadenylation signal sequence (McLauchlan *et al.*, 1985). This particular analysis suggested that whilst some polyadenylation signals serve a unique gene, others may serve up to four genes in so-called 3' co-terminal families. This type of gene arrangement has also been recorded for HSV (reviewed in Wagner, 1985).

The genome of VZV possesses five regions, R1 to R5, which comprise tandem direct reiterations of fairly short sequences with a partial copy at the 3' end (Davison, 1991). The reiterations comprise individual elements that are made up from multiples of 3bp, thus the

reading frame of the ORFs within which R1, R2 and R3 lie, is not disrupted. Reiteration R1 has the most complex arrangement of its individual elements and varies in different strains of VZV (Kinoshita *et al.*, 1988). Reiterations R2, R3 and R4 have less complex patterns, but also vary in size in different virus isolates (Straus *et al.*, 1983; Hondo *et al.*, 1987). R5, which exists as only one single and one partial copy of the reiterated element in the Dumas strain of VZV (Davison & Scott, 1986) has been shown to exist in the form of multiple copies in other strains (Hondo & Yogo, 1988). These observations suggest that the genome of VZV is not of a unique size, and may vary in length between the approximate limits of 124,000 to 126,000bp (Davison, 1991). The role of the reiterated regions is unknown at present. The fact that R1, R2 and R3 lie within protein coding regions, and the reiterations do not introduce alterations in the reading frame, suggests that the virus possesses the potential to encode repeated amino acid sequences within proteins, the functional significance of this unknown.

### 1.2.2 **Transcriptional analysis**

Maps of VZV transcription units have come from just two groups (Ostrove *et al.*, 1985; Maguire *et al.*, 1986; Reinhold *et al.*, 1988). Both groups used specific VZV DNA fragments to map transcripts in VZV infected cells by Northern blot hybridization, and in the most recent investigation ssDNA probes were used to determine the location and orientation of the mRNAs (Reinhold *et al.*, 1988). The results are in accordance with the gene layout predicted from the DNA sequence (Davison & Scott, 1986). In addition to general transcript analysis, the precise 5' and 3' ends of the mRNAs corresponding to genes 36 and 61 (Davison & Scott, 1986; Stevenson *et al.*, 1992) and the 5' end of the mRNA corresponding to gene 62 (McKee *et al.*, 1990) have been determined. More recently, several transcripts have been mapped to genes 12 to 15, including a 1.1kb transcript that maps to gene 13 (Ling *et al.*, 1991),

| Gene | Codons | Function or Properties | Gene | Codons | Function or Properties |
|---|---|---|---|---|---|
| 1 | 108 | | 35 | 258 | |
| 2 | 238 | | 36 | 341 | Thymidine kinase |
| 3 | 179 | | 37 | 841 | Glycoprotein (gpIII) |
| 4 | 452 | Transcriptional activator | 38 | 541 | |
| 5 | 340 | Poss. membrane protein | 39 | 240 | Multiply hydrophobic protein |
| 6 | 1083 | Component of DNA helicase-primase complex | 40 | 1396 | Major capsid protein |
| 7 | 259 | | 41 | 316 | Capsid protein |
| 8 | 396 | dUTPase | 42+45 | 747 | |
| 9 | 302 | | 43 | 676 | |
| 10 | 410 | Poss. tegument protein | 44 | 363 | |
| 11 | 819 | | 46 | 199 | |
| 12 | 661 | | 47 | 510 | Probable protein kinase |
| 13 | 301 | Thymidylate synthase | 48 | 551 | DNase |
| 14 | 560 | Glycoprotein (gpV) | 49 | 81 | Myristylated virion protein |
| 15 | 406 | Multiply hydrophobic protein | 50 | 435 | Multiply hydrophobic protein |
| 16 | 408 | Associated with DNA polymerase | 51 | 835 | $ori_S$-binding protein |
| 17 | 455 | Host shutoff virion protein | 52 | 771 | Component of DNA helicase-primase complex |
| 18 | 306 | Ribonucleotide reductase small sub-unit | 53 | 331 | |
| 19 | 775 | Ribonucleotide reductase large sub-unit | 54 | 769 | Virion protein |
| 20 | 483 | Capsid protein | 55 | 881 | Component of helicase-primase complex |
| 21 | 1038 | | 56 | 244 | |
| 22 | 2763 | Tegument protein | 57 | 71 | |
| 23 | 235 | | 58 | 221 | |
| 24 | 269 | Poss. virion protein | 59 | 305 | Uracil-N glycosylase |
| 25 | 156 | | 60 | 159 | |
| 26 | 585 | Prob. virion protein | 61 | 467 | Poss.transcriptional repressor |
| 27 | 333 | | 62/71 | 1310 | Transcriptional activator |
| 28 | 1194 | DNA polymerase | 63/70 | 278 | Homologue of HSV-1 IE gene |
| 29 | 1204 | ssDNA-binding protein | 64/69 | 180 | Virion protein |
| 30 | 770 | Prob. virion protein | 65 | 102 | Tegument protein |
| 31 | 868 | Glycoprotein (gpII) | 66 | 393 | Protein kinase |
| 32 | 143 | | 67 | 254 | Glycoprotein (gpIV) |
| 33 | 605 | Required for DNA packaging | 68 | 623 | Glycoprotein (gpI) |
| 34 | 579 | Virion protein | | | |

Table 1.1. **VZV gene products**.
Functions or properties of VZV proteins (see text for further details). (Adapted from Davison, 1991).

the gene that encodes TS (Thompson *et al.*, 1987). The 5' and 3' ends of the mRNA for gene 14 have been determined. In addition, the transcripts of genes 12 and 13 are 3' co-terminal (Ling *et al.*, 1991) in agreement with the prediction of Davison and Scott (1986).

### 1.2.3 VZV gene products involved in DNA precursor synthesis

Prior to the genomic sequencing of VZV relatively few virus-induced proteins had been described and no localization of a protein to a specific gene had been reported. Of the 67 unique genes encoded by VZV, functions and/or properties have been assigned to at least 44 as shown in Table 1.1. Functions were initially assigned by identifying homologies to other herpesvirus proteins and/or those of other organisms. Several of these gene products are directly involved in the synthesis of DNA precursors.

The thymidine kinase (TK) expressed by VZV produces thymidylate (dTMP) by the salvage pathway, and like most herpesvirus TKs has a broad substrate specificity (Kit, 1985). In addition to catalysing the phosphorylation of thymidine (TdR) it can also catalyse the phosphorylation of deoxycytidine (CdR), thymidylate (dTMP) and a variety of nucleoside analogues (Fyfe *et al.*, 1978; Cheng *et al.*, 1981; Larder *et al.*, 1983; Chen *et al.*, 1979; Kit *et al.*, 1974). VZV is unique amongst the human herpesviruses as it encodes TS, the enzyme that catalyses the *de novo* synthesis of thymidylate (Thompson *et al.*, 1987; see Section 1.5). Ribonucleotide reductase (RR), which catalyses the conversion of ribonucleoside diphosphates to their corresponding deoxyribonucleotides is encoded by VZV genes 18 and 19 (Davison & Scott, 1986). In the case of HSV-1, this enzyme is essential for growth in non-dividing cells or at elevated temperatures (39.5°C; Goldstein & Weller, 1988). In addition, it has been shown to be a successful target enzyme for antiviral chemotherapy in both VZV and HSV (Cameron *et al.*, 1988; Spector *et al.*, 1989). VZV also specifies deoxyuridine triphosphatase (dUTPase)

(Davison & Scott, 1986), an enzyme that catalyses the conversion of deoxyuridine triphosphate (dUTP) to its corresponding monophosphate (dUMP) with the concomitant release of pyrophosphate. Whether dUTPase plays a role in providing dUMP as a substrate for TS and/or helps to lower the levels of dUTP pools to prevent the incorporation of uracil into viral DNA remains unresolved.

### 1.2.4 **Regulation of gene expression**

The understanding of the regulation of gene expression during VZV replication is still rudimentary, partly due to the difficulty of inducing co-ordinate infections. It is still not clear whether VZV follows a distinct cascade pattern during which three classes of genes, immediate early, early and late, are sequentially expressed, as found in other herpesviruses such as HSV-1. Immediate early (IE) genes in HSV have been defined as those which are transcribed by the unmodified host RNA polymerase in the absence of *de novo* protein synthesis (Honess & Roizman, 1974; Clements *et al.*, 1977; Jones & Roizman, 1979). Functional IE gene products are absolutely required for the expression of HSV-1 early and late genes (Preston, 1979; Watson & Clements, 1980). Early genes are expressed after protein synthesis but are not dependent on DNA replication. The expression of both immediate early and early genes is a prerequisite for late gene expression, which requires DNA replication for maximum efficiency (Jones & Roizman, 1979; Pedersden *et al.*, 1981).

From studies of HSV-1 infected cells, it has been shown that immediate early gene transcription is stimulated by the virion protein Vmw65, a protein that is important for virus replication after infection at low multiplicities or *in vivo* (Ace *et al.*, 1989). Five immediate early genes are transcribed upon initial infection, and at least three of the immediate early proteins are important transactivators required for maximal expression of early and late genes (Everett, 1984a; O'Hare & Hayward, 1985; Sacks *et al.*, 1985). Two of these proteins, Vmw175 (IE175 gene product)

and Vmw63 (UL54 gene product) appear to be essential for virus growth in tissue culture (Deluca *et al.*, 1985; Dixon & Schaffer, 1980; Preston, 1979b). The third transcriptional activator, Vmw110 (IE110 gene product) is not essential for virus replication in tissue culture (Stow & Stow, 1986; Sacks & Schaffer, 1987).

Sequencing of the VZV genome revealed the presence of potential homologues to four of the five HSV-1 IE genes (Davison & Scott, 1986). In addition, four polypeptides have been identified as immediate early proteins by the classic inhibitor approach (Lopetegui *et al.*, 1985; Shiraki & Hyman, 1987), suggesting that the control of gene expression between the two viruses may be similar. A number of recent studies by several groups have confirmed some functional similarities between the control of gene expression in both viruses, but also discovered significant differences.

The protein encoded by VZV gene 62 shares amino acid homology to functionally important regions of Vmw175 of HSV-1 (McGeoch *et al.*, 1986), and, like Vmw175, can function as a transcriptional transactivator (Everett, 1984a; Inchauspe *et al.*, 1989). The protein can complement the growth of HSV-1 mutants with temperature sensitive lesions in, or deletions of the IE175 gene, that normally can only be propagated on cell lines that constitutively express Vmw175 (Felser *et al.*, 1988). When both copies of the IE175 gene from HSV-1 were replaced by the VZV gene 62, under the control of the IE175 promoter, the resulting recombinant virus could be propagated on normal cell lines (Disney & Everett, 1990). The gene product of VZV gene 62 can autoregulate expression from its own promoter (Disney *et al.*, 1990) but not the IE175 promoter (Disney & Everett, 1990). This is similar to Vmw175 and suggests that the two proteins play very similar roles in the control of virus gene expression.

The product of VZV gene 4 and Vmw63 of HSV-1 share 29% amino acid identity (McGeoch *et al.*, 1988). Vmw63 is essential for HSV-1 replication (Sacks *et al.*, 1985) and

functions as a trans-regulatory protein (Everett, 1986; Rice & Knipe, 1988; Sekulovich *et al.*, 1988). The VZV gene 4 protein product can transactivate at least one early and one late gene in a synergistic manner with the gene product of VZV gene 62 (Inchauspe *et al.*, 1989). However, it has also been shown that the gene products of the two VZV immediate early genes, 4 and 62 can transactivate the thymidine kinase promoters of both the VZV and HSV, but that their HSV-1 homologues, Vmw175 and Vmw63, can only transactivate the HSV-1 thymidine kinase promoter (Inchauspe & Ostrove, 1989). Thus, despite the functional simlilarities of these transcriptional activator proteins, their precise mechanisms of action differ. This is probably caused by a combination of differences in the cis acting sequences of the genes upon which they act (Inchauspe & Ostrove, 1989) and differences in the architecture of the individual proteins.

The gene product of VZV gene 63 displays a short region of amino acid sequence homology to the HSV-1 IE protein Vmw68 (Davison & McGeoch, 1986). No function has been assigned to either protein, but an HSV-1 mutant lacking the carboxy-terminal third of Vmw 68 exhibits poor growth in tissue culture, shows reduced expression of at least one late gene and is non-neurovirulent in mice (Post & Roizman, 1981; Sears *et al.*, 1985). Expression of gene 63 has been detected in human trigeminal ganglia (Vafai *et al.*, 1988), but no further studies of the function of its gene product have been reported.

VZV gene 61 is located at the position equivalent to HSV-1 IE110 which encodes the transactivating protein Vmw110 (McGeoch *et al.*, 1988). Despite the fact that the VZV protein is just over half the size of Vmw110 and bears only local amino acid homology, at a region that has been proposed as a putative zinc-finger DNA binding motif, this lead to the tentative proposition that the two proteins could have similar functions (Perry *et al.*, 1986). This particular zinc-finger motif has since been found in several non-herpesvirus proteins (Freemont *et al.*, 1991).

The VZV gene 61 product can function as trans-repressor (Inchauspe *et al.*, 1989; Nagpal & Ostrove, 1991) and possesses no trans-activational activity (Cabriac *et al.*, 1990). The IE110 gene possesses several upstream elements essential for immediate early gene expression, such as the so-called "TAATGARAT" sequence (Mackem & Roizman, 1982). However, analysis of the upstream region of the VZV gene 61 revealed no convincing examples of such elements (Stevenson *et al.*, 1992). In addition, it is not clear from studies of VZV immediate early gene expression (Lopetegui *et al.*, 1985; Shiraki & Hyman, 1987) if any of the proteins identified correspond to the gene 61 protein. Indeed, the counterpart of gene 61 in another alphaherpesvirus, PRV, is an early gene (Cheung, 1991), and the transcription of the gene 61 homologue in EHV-1 has not been detected under immediate early conditions (Gray *et al.*, 1987; E. Telford and A. J. Davison, personal communication). These data suggest that $V_{mw}$110 and the VZV gene 61 protein are not functional homologues.

The discovery that $V_{mw}$65, the product of UL48 gene of HSV-1, could transactivate the VZV gene 62 promoter (McKee, 1990; McKee *et al.*, 1990), in an analogous manner to the way in which it transactivates HSV-1 immediate early genes, led to obvious interest in VZV gene 10, which lies at the equivalent position to the HSV-1 gene UL48 (McGeoch *et al.*, 1988). The predicted gene product of gene 10 shows 35% homology with $V_{mw}$65 (Dalrymple *et al.*, 1985), but recent analysis has shown that, in contrast to $V_{mw}$65, it cannot stimulate transcription from the gene 62 promoter (McKee, 1990). Neither can it form a complex analogous to the immediate early complex that is formed when $V_{mw}$65 interacts with oligonucleotides representative of promoter elements of HSV-1 immediate early genes or VZV gene 62 (McKee *et al.*, 1990).

In summary, there are certain similarities between the control of gene expression in VZV and HSV-1, but also significant differences. These differences between the two viruses may, in part, explain some of their major

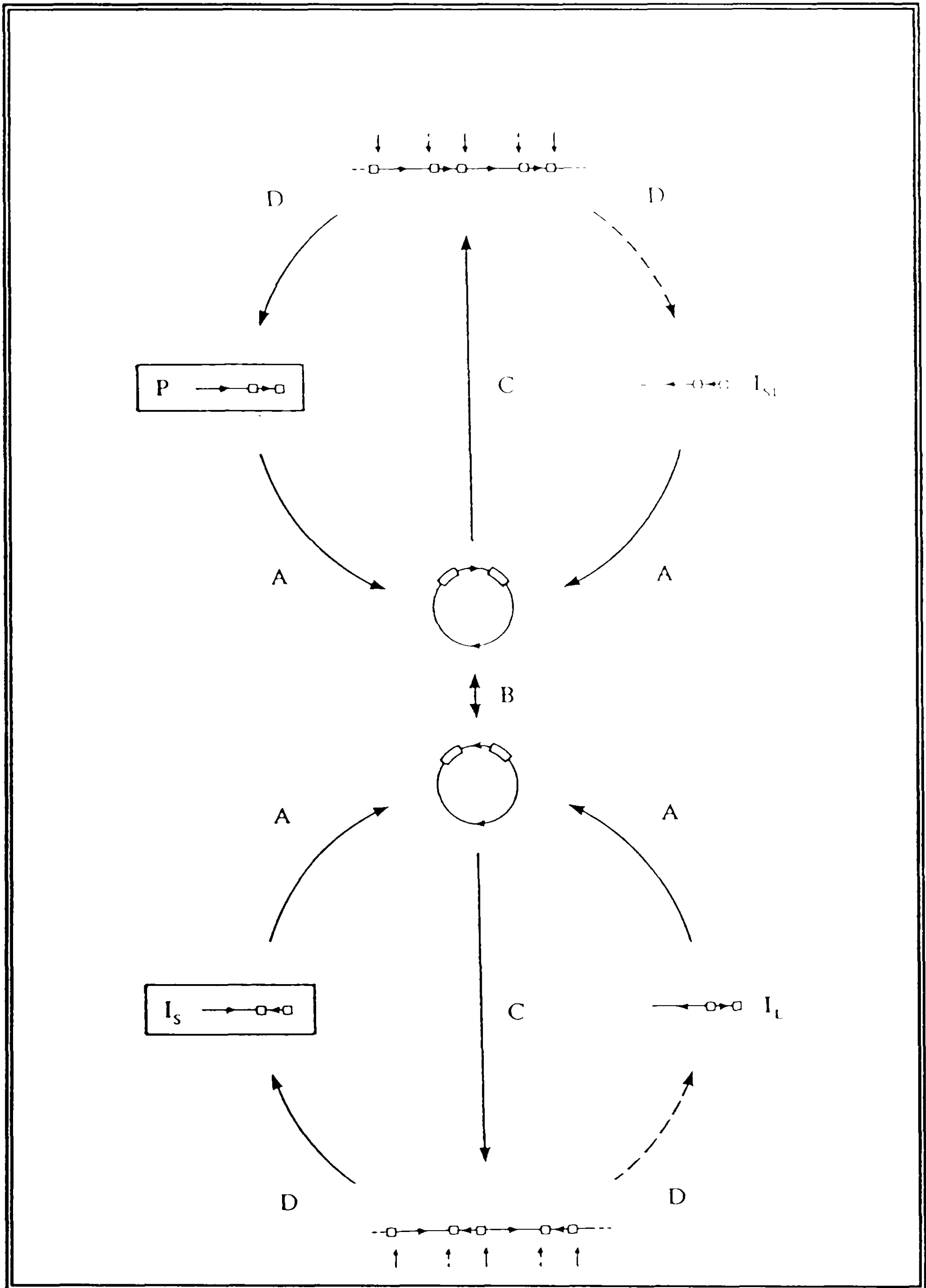



Fig. 1.3. **VZV DNA replication.**
Step A represents the circularization of one of the four possible genome arrangements and step B indicates the period during which limited replication occurs and segment inversion may take place. At step C, DNA replication occurs by a rolling circle mechanism, and in the final step D, concatemers are cleaved to produce unit length genomes for encapsidation (adapted from Davison, 1984).

biological differences. The view is now emerging that HSV-1 may not be the best model on which to base studies of the control of VZV gene expression, and other alphaherpesviruses such as PRV and EHV-1, may serve as better models for future studies (Stevenson *et al.*, 1992).

### 1.2.5 **VZV DNA replication**

A model for VZV DNA replication, based on several structural features of the viral genome and experimentally supported views of PRV and HSV-1 DNA replication, has been proposed by Davison (1984). Linear DNA, in one of its four isomeric forms, circularises, probably by the direct ligation of the termini to produce a novel L-S joint identical to the normal L-S joint in the linear genome. After circularisation, limited DNA replication occurs and the short segment can be inverted by intramolecular recombination between inverted repeats. DNA replication proceeds by the generally accepted rolling-circle method to produce head-to-tail concatemers, as has been shown for PRV and HSV-1 (Ben-Porat & Rixon, 1979; Jacob *et al.*, 1979). The final stage involves the specific cleavage of concatemers to generate unit length genomes for encapsidation. Cleavage usually occurs at the novel L-S junction, but occasionally at the normal L-S junction generating the major and minor genome arrangements (Davison, 1984; see Fig. 1.3).

Two copies of a putative origin of DNA replication were detected in TRs/IRs (Davison & Scott, 1985), on the basis of homology to the HSV-1 oris sequence (Stow & McMonagle, 1983). A 259bp fragment containing the putative palindromic origin of replication showed activity in plasmid replication assays, originally used to identify the HSV-1 origin of DNA replication, strongly suggesting that this was the real origin of VZV DNA replication (Stow, 1982; Stow & Davison, 1986). Deletions made in the palindromic region of the putative origin sequence dramatically reduced replicative ability, confirming that this sequence was the origin of replication (Stow &

Davison, 1986).

A set of seven HSV-1 genes which encode proteins necessary and sufficient for viral DNA synthesis have recently been identified (McGeoch *et al.*, 1988; Wu *et al.*, 1988; Challberg & Kelly, 1989). VZV possesses seven genes whose predicted gene products show homology to those of HSV-1 (McGeoch *et al.*, 1988), only one of which, the gene product of gene 51, has been characterized in any detail. This protein, which has extensive amino acid homology with its HSV-1 counterpart, the gene product of the UL9 gene, has been shown to bind to the VZV origin of DNA replication (Stow *et al.*, 1990).

## 1.3 **Biology of varicella-zoster virus**

Varicella-zoster virus is the causative agent of two distinct clinical syndromes, chickenpox and shingles. Chickenpox, or varicella, is a relatively benign disease in children, but can be severe, and in some cases life threatening. in infected adults or immunocompromised individuals.

### 1.3.1 **Varicella: disease**

After an incubation period of 10 to 20 days following initial infection a characteristic rash, the most typical feature of varicella, starts to erupt. Succesive crops of new lesions erupt over the next two to four days, appearing first on the scalp and trunk and then spreading to the extremities. The crusts from these lesions are shed between one and three weeks later. Vesicles also appear on all mucosal surfaces, which rupture and are thought to be responsible for the spread of virus by the respiratory route (Gelb, 1990).

The illness is generally benign with only mild fever whilst new lesions are erupting, and pruritis which persists throughout the vesicular stage (Broughton, 1966). The most frequent complication is bacterial superinfection of skin lesions (Bullowa & Wishik, 1935) but administration

of appropriate antibiotic treatment is usually successful. Varicella pneumonia is the commonest serious complication in adults with a mortality rate of 10% to 40% (Guess, 1986). A recent study of pregnant women with primary varicella reported an incidence of symptomatic pneumonia of almost 10% with one fatality (Paryani & Arvin, 1986). Neonatal varicella is seen in infants infected *in utero* or shortly after birth and have insufficent maternal antibody to combat the disease and has a fatality rate exceeding 30% (Brunell, 1983; Preblud *et al.*, 1985). Varicella infections in the mother during the first trimester of pregnancy can occasionally result in congenital defects (Brunell, 1983; Paryani & Arvin, 1986; Batik & Stevens, 1989).

Immunocompromised patients, whether leukaemic, suffering other malignacies, receiving immunosuppressive chemotherapy or suffering from the acquired immunodeficiency syndrome (AIDS), have significantly increased complications and severe infections of varicella (Krugman, 1977; Jura *et al.*, 1989). In addition, a variety of neurological defects are associated with varicella, not least Reye syndrome. Little is known about this association but varicella preceeds 20% to 30% of Reye syndrome cases and salicylate (aspirin) use has been linked to an increased risk of the disease (Hurwitz *et al.*, 1982; Remington *et al.*, 1986).

### 1.3.2 **Varicella: pathogenesis**

Infection is initiated when the virus invades the mucosa of the upper respiratory tract and oropharynx, or alternatively via the conjunctiva. Viral replication commences at the primary site of inoculation and virus dissemination, termed primary viremia, is via the bloodstream and the lymphatic system. When taken up by the cells of the reticuloendothelial system, the virus undergoes multiple cycles of replication and extensive secondary viremia develops. These events take place during the prodromal or incubation period of the disease (Gelb, 1990). Three to five days prior to the onset of clinical varicella, simultaneous increases in cellular and viral TK

activity are seen in serum samples (Källander *et al.*, 1989).

Secondary viremia is followed by vesicle eruption with fresh lesions appearing in successive crops, due to the process of cyclic viremia (Feldman & Epp, 1979). Viremia is usually terminated after three days by specific humoral and cell mediated responses, after which no fresh lesions erupt (Gelb, 1990).

### 1.3.3 **Zoster: disease**

Shingles, or zoster, is the clinical syndrome associated with the reactivation of VZV from its site of latency (see Section 1.3.5). The rash associated with zoster is usually preceeded by pain and paresthesia in the involved dermatome which ranges from mild itching, or tingling, to severe pain. The term 'zoster sine herpete' has been applied to describe those prodromal symptoms where there is no subsequent vesicle eruption, but a typical antibody response is observed (Lewis, 1958).

When vesicle eruption does occur, it is characteristically unilateral, does not cross the mid-line and is limited to the cutaneous innervation of a single sensory ganglion. Lesions are typically varicelli-form, but evolve less rapidly and appear more often as grouped vesicles rather than the discrete lesions seen with varicella. The areas supplied by the trigeminal nerve and thoracic ganglia are most commonly involved in zoster (Burgoon *et al.*, 1957). New lesions erupt over a period of several days, and crusts, taking as long as a week to form, may persist for two to three weeks. The severity and duration of the cutaneous eruptions are roughly proportional to patient age (Brown, 1976). When zoster involves the opthalmic division of the trigeminal nerve, a variety of eye lesions can occur (Guess, 1986). The Ramsey-Hunt syndrome, characterised by facial palsy and zoster of the external ear, results from the involvement of the geniculate ganglia (Denny-Brown *et al.*, 1944).

The commonest complication of zoster is post-herpetic

neuralgia, occuring in 50% of infected individuals over 60 years of age. This is especially common with opthalmic zoster and, although difficult to treat, often resolves spontaneously (de Moragas & Kierland, 1957). True post-herpetic neuralgia, where pain persists for more than one month, has an estimated incidence of approximately 10% (Watson & Evans, 1986). In approximately half of these patients pain resolves within two months, but in 20% to 30% of those affected it can persist for over a year, and in a minority of cases for over a decade (de Moragas & Kierland, 1957).

As with varicella, herpes zoster is more severe in immunocompromised individuals or those with malignancies. Chronic zoster and persistent infection of the central nervous system are common complications in these individuals (Gallagher & Merigan, 1979; Horten *et al.*, 1981; Ryder *et al.*, 1986).

### 1.3.4 Zoster: pathogenesis

Although the pathogenesis of herpes zoster is poorly understood, the relationship between the distribution of skin lesions and the anatomy of the nervous system was first appreciated by Bright in 1831. The current model of zoster pathogenesis is based on clinical and epidemiological data, and analogy to HSV infections (Hope-Simpson, 1965). Reactivation of VZV is sporadic and infrequent, and although the mechanism of reactivation is unclear, many of the conditions correlated to reactivation are associated with immunosuppression (Gelb, 1990). This led to the proposal that it is the deterioration of host defences, below the level of containment, that allows VZV to begin active replication in the sensory ganglia (Hope-Simpson, 1965). The resultant virus produced is almost immediately contained by either antibody or cell mediated immunity. However, the release of antigen boosts antibody levels sufficiently to mark the event (Gershon *et al.*, 1982; Luby *et al.*, 1977). Similar boosts in VZV-specific antibody are seen following exposure to varicella and reinfection has

been suggested in the partially immune host (Arvin *et al.*, 1983; Gershon *et al.*, 1986; Oxman 1981; Weigle & Grose, 1984). When host defences deteriorate below the point where VZV replication can be fully contained, extensive viral multiplication occurs in ganglia, resulting in neuronal necrosis and intense inflammation. The infectious virus then spreads down the sensory nerve, infects the skin and produces the characteristic clustered vesicles of herpes zoster. Zoster vesicles, like those of varicella, contain cell-free VZV that may be shed to the environment.

Occasionally, the immune system can curtail the formation of cutaneous lesions but not the necrosis and inflammatory response in the ganglia, resulting in pain without associated skin lesions, (zoster sine herpete) (Lewis, 1958; Easton, 1970; Juel-Jenson, 1972; Luby *et al.*, 1977).

Formal proof that herpes zoster is caused by the same virus that caused the initial varicella infection only came in the last decade. Straus and co-workers (1984) used DNA restriction enzyme analysis to show that a single VZV isolate caused both varicella and subsequent zoster infections in a child with Wiskott-Aldrich syndrome.

Recent invesigations have suggested that more than one strain of VZV can be present during the course of a VZV infection. This was shown in a study of VZV isolated from zoster lesions from a boy previously inoculated with the vaccine strain of VZV. Although zoster may be caused by the reactivation of the vaccine strain (see Section 1.4.2), Gelb and co-workers (1987) were able to show that the virus recovered from the zoster lesions was wild type VZV as it differed from the DNA restriction enzyme profile of the vaccine strains of VZV. In a second study, Hondo and co-workers (1987) reported different strains of virus isolated from zoster lesions and the corresponding spinal ganglia. However, the strains described by Hondo and co-workers (1987) vary predominately in the regions of reiteration. Although these regions are relatively stable, even after extensive passaging in tissue culture, variants do arise

(Straus *et al.* 1981; Ecker *et al.* 1984; Zweerink *et al.* 1984). In addition, the occasional loss of restriction sites in this case (Hondo *et al.*, 1987) could be the result of a single point mutation or small deletion. This is indicative of only minor differences that could theoretically arise during viral DNA replication and therefore not reflect strain differences.

### 1.3.5 **Latency**

The processes and mechanisms of herpesviral latency are now becoming reasonably well understood, especially for HSV and EBV (Stevens, 1989; Kosz-Vnenchak *et al.*, 1990). Despite this, the understanding of the latent infection of VZV is still at an early stage. However, it is now clear that although there are some similarities between the latent infections of VZV and HSV, there are also some distinct differences.

In HSV latent infections, it is universally accepted that the virus resides in sensory neurons (Kennedy *et al.*, 1983; McLennon *et al.*, 1980; Cook *et al.*, 1974). Detectable transcription only occurs in a single region of the HSV genome to produce abundant copies of the so-called latency associated transcripts (LATs) which are anti-sense, relative to the $V_{MW}110$ mRNA (Stevens *et al.*, 1987). Stevens and co-workers (1987) postulated that the retention of $V_{MW}110$ mRNA within the nucleus could effectively block translation of a viral protein important in initiating productive infection, and thereby play a role on the maintenance of the latent state.

The HSV-1 IE protein $V_{MW}175$ has been implicated in the establishment and maintenance of the latent state (Stevens *et al.*, 1987; Green *et al.*, 1981; Puga & Notkins, 1987). Its role in latency is not yet fully understood, however, it can serve as a transcriptional activator both *in vivo* and *in vitro* (Beard *et al.*, 1986). In addition, $V_{MW}175$ is essential for the reactivation of HSV in an *in vitro* model system (Russell & Preston, 1986; Russell *et al.*, 1987). A second HSV-1 IE protein, $V_{MW}110$, which is dispensable for

growth in tissue culture, has also been shown to be essential for the reactivation of HSV-1 in the *in vitro* model system (Russell & Preston, 1986; Russell *et al.*, 1987). In addition, $V_{MW}110$ can act synergistically with $V_{MW}175$ in the process of transactivation (Everett, 1984a).

It is now clear that sensory ganglia harbour latent VZV as has been demonstrated by *in situ* and Southern blot analysis (Croen *et al.*, 1988; Gilden *et al.*, 1983, 1986; Hyman *et al.*, 1983). This has been confirmed by explantation and *in vitro* cultivation experiments of such ganglia which have demonstrated the production of VZV encoded IE mRNAs and proteins (Vafai *et al.*, 1988). In contrast to HSV, infectious VZV could not be detected in the ganglia by this procedure (Vafai *et al.*, 1988).

A detailed analysis of ganglia latently infected with VZV has shown that transcription occurs from five separate regions of the viral genome (Croen *et al.*, 1988). By using strand specific probes, Croen and co-workers (1988) were able to demonstrate the transcription of VZV genes 4 (transcriptional activator), 29 (major DNA binding), 31 (gpII), and 62 (transcriptional activator), whilst showing that genes 10 (tegument protein), 14 (gpV) and 67 (gpIV) were not transcribed during latent infection. In addition, explantation and *in vitro* cultivation experiments have detected a transcript from VZV gene 63 (homologue of HSV-1 IE68 gene), and seven major protein species from ganglia latently infected with VZV (Vafai *et al.*, 1988). Moreover, VZV lacks a homologue of $V_{MW}110$ (see Section 1.2.4) and transcripts covering a region of the VZV genome equivalent to the region of HSV-1 where the LATs have been mapped do not appear to have been detected (Vafai *et al.*, 1988; Croen *et al.*, 1988; Croen & Straus, 1991). Clearly, this situation is different to the HSV latent infection.

Further differences between the latent infections of VZV and HSV have been reported. VZV replicates efficiently in both neuronal and non-neuronal cells (Nagaskima *et al.*, 1975; Croen *et al.*, 1988). The virus remains latent, and subsequently reactivates from non-neuronal cells that

support its replication. This facilitates its spread to neuronal and non-neuronal cells, resulting in the destruction of large parts of the ganglion (Head & Campbell, 1900; Esiri & Tomlinson, 1972). During this process, multiple sensory nerves become infected and convey the virus to all quadrants of the cutaneous dermatome, the resulting inflammation and necrosis inciting severe and chronic neuropathy.

In contrast, HSV reactivates within a neuronal cell and spreads readily to the skin. However, spread within the ganglion is inefficient, as the surrounding satellite cells are less permissive for productive HSV replication (than VZV), so that cutaneous involvement and neuropathy are limited.

The relative infrequency of VZV reactivation compared to that of HSV may reflect its non-neuronal site of latency, thus rendering it less subject to the neural triggers that provoke HSV reactivation. There may also be a variety of physical, metabolic or immune obstacles to the spread of VZV from its latent site to the neural conduit through which it is transported to the skin.

In summary, VZV latent infections, do not produce LATs, but at least five regions of the genome are transcriptionally active. The virus also lacks a functional homologue of the HSV-1 $V_{MW}110$ protein. VZV resides at many sites in ganglia and reactivates infrequently, possibly from non-neuronal cells. In HSV-1 latent infections, the only detectable transcripts are the LATs, the virus resides almost exclusively in neuronal cells, and it reactivates frequently from single neurons. These observed differences may, in part, explain the differences in HSV-1 and VZV reactivation from latency.

## 1.4 Control of infection

As stated above, varicella is usually a benign and ubiquitous disease in normal children that requires no specific therapy, itching being controlled locally by

calamine lotion or systemically by antihistamines. However, in immunocompromised individuals, healthy adults, pregnant women and cases of zoster, treatment, either preventative or prophylactic, is desirable and several approaches exist.

### 1.4.1 VZV immune serum globulin and zoster immune plasma

The administration of either VZV immune serum globulin, a high-titre VZV antibody preparation, or VZV convalescent plasma, within three days of exposure to VZV, has been shown to be effective in preventing or attenuating subsequent varicella in normal (Brunell *et al.*, 1969) and immunocompromised children (Gershon *et al.*, 1974; Zaia *et al.*, 1983; Geiser *et al.*, 1975; Balfour *et al.*, 1977). However, VZV immune serumglobulin (and probably zoster immune plasma) has no role to play in the treatment of established disease (Brunell, 1981; Brunell *et al.*, 1969; Gershon *et al.*, 1974).

### 1.4.2 Vaccination

The Oka vaccine strain of VZV was produced by attenuating the virus in tissue culture (Takahashi *et al.*, 1974, 1975). The virus has been tested extensively in Japan and appears safe and effective, even in immunocompromised children. It is currently licenced in Japan and several European countries. In healthy children its efficacy approaches 100% (Arbeter *et al.*, 1986a; Asano *et al.*, 1985; Weibel *et al.*, 1984), with the immune response lasting for several years (Arbeter, 1986a). In adults lower rates of seroconversion and break-through infections are seen (Gershon *et al.*, 1986).

Although efficacious the vaccine is not without its problems. Zoster may be caused by reactivation of the latent vaccine strain (Gelb *et al.*, 1987; Hayakawa *et al.*, 1984; Williams *et al.*, 1985) or wild-type virus. The true longevity of protection is still to be determined, with waning protection in adults, after childhood immunization, possibly resulting in increases of morbidity and mortality.

However, the Oka vaccine strain is responsive to at least one antiviral compound, acyclovir (Preblud *et al.*, 1984a).

The effectiveness of the vaccine strain in high risk children is sufficient to warrant vaccination of these individuals (Gershon *et al.*, 1986). A good immune response is seen when the Oka vaccine strain is combined with the vaccine against mumps, measles and rubella. In addition, both universal vaccination with the combined vaccine, and the Oka vaccine strain alone, has been proposed (Brunell, 1977; Arbeter *et al.*, 1986). However, for reasons stated above, this still remains a controversial proposition.

### 1.4.3 Antiviral compounds

Antiviral therapies are probably inappropriate or unnecessary in routine varicella of immunocompetent children, but should be considered when complications occur. They should also be a consideration in persons with acquired or congenital cellular immune deficiences, or in pregnant women (Straus, 1987; Straus *et al.*, 1988).

Human leukocyte, or alpha-interferon, has successfully been used to treat children infected with varicella who have cancer (Arvin *et al.*, 1982). However, it must be given parenterally and has several associated toxicities. As a consequence, most therapies are currenly based on two nucleoside analogues. Vidarabine (araA; 1-ß-D-arabinofuranosyladenosine; adenosine arabinoside) has proved efficacious for both varicella and zoster despite some toxicity (Johnson *et al.*, 1975; Whitley *et al.*, 1981, 1982a, b). However, it has been largely superceded for most applications by acyclovir (ACV, 9-(2-hydroxyethoxymethyl)guanine). Although originally identified as a potent inhibitor of HSV-1 (Elion *et al.* 1977; Schaeffer *et al.*, 1978), ACV is also effective against VZV and seems to be superior to araA for most applications (Shepp *et al.*, 1986). ACV, along with many other nucleoside analogues that are active against herpesviruses, must first be phosphorylated by the virus encoded thymidine kinase (see Sections 3.1 and 3.4).

Many other compounds have been shown to have anti-varicella activity. These include 5-bromovinyl-2'-deoxyuridine (BVdU), the 5-bromo-, 5-chloro- and 5-iodo-vinyl-derivatives of ß-D-arabinofuranosyldeoxyuridine (BVaraU, CVaraU, IVaraU respectively), 1-(2-deoxy-2-fluoro-ß-D-arabinofuranosyl)-5-methyluracil (FMAU) and 1-(2-deoxy-2-fluoro-ß-D-arabinofuranosyl)-5-iodocytosine (FIAC; Machida, 1986; Machida *et al.*, 1982). The therapeutic index, the ratio between the $IC_{50}$ for tissue culture cells and the concentration of drug to give 50% plaque reduction for VZV, of BVaraU is $>3x10^6$ whilst ACV has a therapeutic index of 520, an approximate 10,000 fold difference (Machida *et al.*, 1986).

Additional compounds, which have been shown to possess anti-varicella activity include 5-ethynyl-2'-deoxyuridine (EYUdR) and 5-propynyl-2'-deoxyuridine (PYUdR) and their corresponding ß-D-arabinofuranosyl-derivatives EYaraU and PYaraU (Lacey *et al.*, 1991). Interestingly, several of these compounds can inhibit VZV TS, when converted to the monophosphate form, including EYUdR (Simon Lacey, personal communication). Whether their antiviral activities are mediated through this enzyme remains unresolved. One of these recently described compounds, BVaraU, is currently under-going clinical trials as a potential anti-varicella drug (Lacey *et al.*, 1991).

A number of highly potent antiviral compounds that do not require activation by the the viral TK have recently been described, including (S)-9-(3-hydroxy-2-phosphonylmethoxypropyl)adenine ((S)-HPMPA) which is highly active against both wild-type and $TK^-$ strains of VZV (DeClercq *et al.*, 1986). In this system, it has a therapeutic index in excess of 25,000, compared with ACV (>400) and BVdU (>15,000) (DeClercq *et al.*, 1986). Similar values were obtained in a separate study for ACV (520) and BVdU (18,000) (Machida *et al.*, 1986). In an analysis of twenty different VZV isolates, the mean $IC_{50}$ value for (S)-HPMPA was 0.0018µM, giving a mean therapeutic index of 29,000 (Baba *et al.*, 1987). Another compound that is active

against VZV, and is not activated by TK (at least for HSV-1), is cyclopentenylcytosine (Marquez *et al.*, 1990).

Recently the mechanism of action of (S)-HPMPA has been elucidated. The enzyme 5-phosphoribosyl-1-pyrophosphate synthetase catalyses the formation of the diphosphate derivative (Balzarini & DeClercq, 1991) and the triphosphate derivative is presumably formed by cellular kinases (see Larder & Darby, 1984). The triphosphate form, which may be thought of as an ATP analogue, is predicted to interfere with the viral DNA polymerase (Balzarini & DeClercq, 1991).

Another drug that is active against VZV is 9-[4-hydroxy-2-(hydroxymethyl)butyl] guanine, which inhibits the VZV DNA polymerase (Abele *et al.*, 1988). Improved prodrug forms of this compound have also been described (Lake-Bakaar *et al.*, 1989). In addition, the 9-(3-hydroxypropoxy)- and 9-[3-hydroxy-2-(hydroxymethyl)propoxy]-guanine analogues have been shown to be more active *in vitro* against VZV than ACV (Harnden *et al.*, 1990).

### 1.4.4 **Antiviral resistance**

Strains of VZV resistant to one or more of a broad range of antivirals have been isolated *in vitro*. These include strains resistant to ACV (Biron *et al.*, 1982; Shiraki *et al.*, 1986; Sawyer *et al.*, 1988), BVaraU and araA (Sakuma, 1984), BVaraU (Sakuma *et al.*, 1985), BUdR and IUdR (Shigeta *et al.*, 1986) and BVaraU and PYaraU (Lacey *et al.*, 1991). Recently, ACV resistant clinical isolates have been observed, but to date have been confined to individuals with AIDS (Jacobson *et al.*, 1990; Linnemann *et al.*, 1990). In the cases where the mode of resistance has been characterized at a molecular level, lesions in the virus encoded TK have been shown to confer drug resistance. These lesions take the form of single point mutations in the TK gene. In many cases, resistance is associated with the expression of a form of VZV TK with no detectable enzyme activity; in some cases the expressed TK has modified activity. Mutations in the VZV DNA polymerase

have been shown to confer resistance to ACV (Gaillard *et al.*, 1990), a mode of resistance which is also seen in HSV-1 (Larder & Darby, 1984).

The variety of methods for attacking the diseases caused by VZV hold promise for control of the virus in the future. This is particularly so because of the availability of the broad array of antiviral compounds and the fact that wide-spread resistance has not been seen to the same levels as with ACV treatment of HSV infections.

## 1.5 Thymidylate synthase: an introduction

Efficient DNA synthesis requires the presence of adequate pools of the four 2'-deoxyribonucleotides, dATP, dCTP, dGTP and dTTP. The first three of these are synthesized by direct reduction of their corresponding ribonucleotides in a reaction catalysed by ribonucleotide reductase. In eukaryotes and certain prokaryotes, reduction occurs at the diphosphate level and is followed by phosphorylation of the dNDPs to dNTPs by the enzyme nucleoside diphosphate kinase. In other prokaryotic organisms reduction of ribonucleotides occurs at the triphosphate level. Ribonucleotide reductase also mediates the synthesis of a fourth 2'-deoxyribonucleotide, namely dUTP.

As dTTP has no ribonucleotide counterpart, it cannot be synthesized in this way. Instead, it is synthesized *de novo*, commencing with the formation of thymidylate (dTMP) from deoxyuridylate (dUMP) in a reaction catalysed by the enzyme thymidylate synthase (TS). dUMP is provided by the hydrolysis of dUTP, catalysed by deoxyuridylate triphosphate dephosphohydrolase (dUTPase), or by the deamination of deoxycytidylate (dCMP). dTMP can also be formed by a salvage pathway where thymidine kinase (TK) catalyses the phosphorylation of thymidine. The phosphorylation of dTMP to dTDP is catalysed by thymidylate kinase and nucleoside diphosphate kinase catalyses the phosphorylation of dTDP to form dTTP (Neuhard & Nygaard, 1987; O'Donovan & Neuhard, 1970).

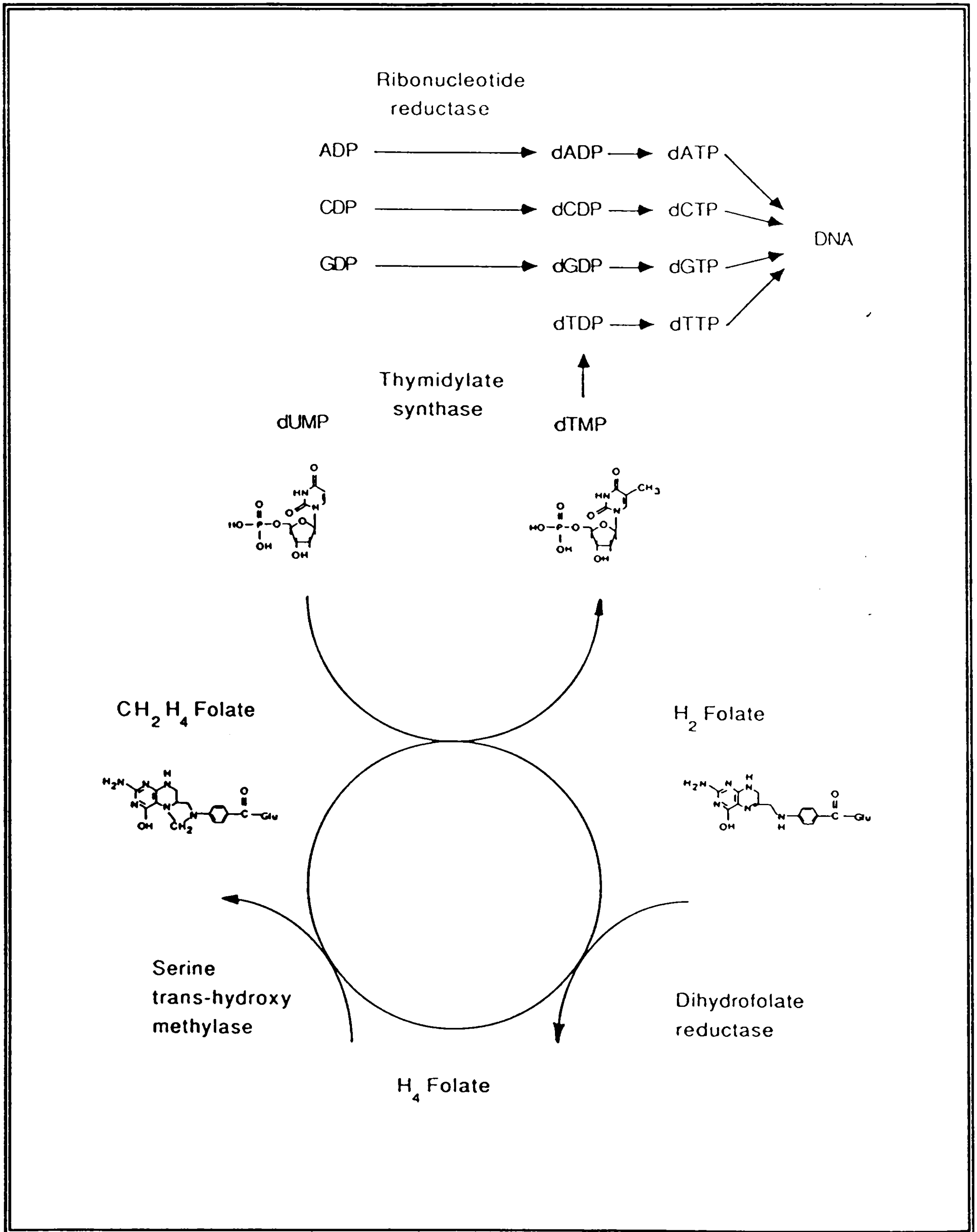



Fig. 1.4. **Thymidylate cycle.**
The role of TS in the formation of DNA precursors. The following abbreviations are used: $H_2$folate - dihydrofolate; $H_4$folate - tetrahydrofolate; $CH_2H_4$folate - $N^5N^{10}$methylenetetrahydrofolate. Standard abbreviations are used for nucleotides (adapted from Lewis & Dunlap, 1981).

### 1.5.1 The Thymidylate Cycle

The thymidylate cycle involves three enzymes, thymidylate synthase, dihydrofolate reductase and serine trans-hydroxymethylase (Fig. 1.4; Lewis & Dunlap, 1981). Thymidylate synthase (methylenetetrahydrofolate: deoxyuridine-5'-monophosphate C-methyltransferase: EC 2.1.1.45) catalyses the *de novo* synthesis of thymidylate via the reductive methylation of deoxyuridylate with the concomitant conversion of the cofactor, $N^5,N^{10}$-methylenetetrahydrofolate to dihydrofolate. This reaction is the only one in Nature that results in the production of dihydrofolate, an oxidised form of folate that cannot be used in any biological process. To regenerate the reduced folate pools, dihydrofolate is converted to tetrahydrofolate by dihydrofolate reductase (DHFR), a reaction essential for the maintenance of these pools and thus cell viability. The enzyme serine trans-hydroxymethylase completes the thymidylate cycle by catalysing the formation of $N^5,N^{10}$-methylenetetrahydrofolate. The pivotal roles played by both TS and DHFR in the synthesis of DNA precursors has lead to their recognition as target enzymes for the treatment of both infectious diseases and cancer (see Sections 1.4.3 and 1.12)

### 1.5.2 TS from different sources

TS is a ubiquitous enzyme found in all cellular organisms and some large DNA viruses. The characterization of TS, its relationship with DHFR and its interaction with other proteins is discussed below for a variety of different organisms.

#### 1.5.2.1 Bacteria

The first studies of TS were performed on the enzyme isolated from *Lactobacillus casei* (Friedkin *et al.*, 1962; Blakley, 1963). The isolation of a methotrexate resistant *L.casei* strain, that over-produced TS and DHFR, meant that large amounts of highly purified enzyme were available for study (Crusberg *et al.*, 1970) and has resulted in the

*L.casei* enzyme becoming the most intensively studied TS from all sources (Santi & Danenberg, 1984). The over-production of both TS and DHFR suggests that the genes encoding these two enzymes are closely linked.

The genes encoding TS and DHFR are also closely linked in *Bacillus subtilis* and *Staphylococcus aureus*. *B.subtilis* has two TS loci, *thyA* and *thyB* (Piggot & Hoch, 1985). The *thyB* gene encodes a heat-sensitive form of TS (Neuhard *et al.*, 1978) and can complement the growth of both an *Escherichia coli thyA* strain (Rubin *et al.*, 1978) and a *B.subtilis thyA thyB* strain (Iwakura *et al.*, 1988). This gene is located within the same operon as *dfrA*, the gene encoding DHFR, with the termination codon of *thyB* overlapping the initiation codon of *dfrA* (Iwakura *et al.*, 1988).

The nucleotide sequence of the gene for DHFR and the partial sequence of the gene for TS of *S.aureus* (Burdeska *et al.*, 1990) are identical to the complete nucleotide sequences of the genes for TS and DHFR in *Tn*4003, a transposon that confers resistance to trimethoprim (a DHFR inhibitor) in *S.aureus* (Rouch *et al.*, 1989). In both cases, the two genes are separated by only 44 nucleotides on the same strand, suggesting they form part of the same operon.

By way of contrast the genes for TS (*thyA*) and DHFR (*frd*) are widely separated on the *E.coli* chromosome (Bachmann, 1983). However, the two enzymes have been proposed to form part of a multienzyme complex that is responsible for channelling newly synthesized dNTPs into replicating DNA (Mathews *et al.*, 1978; 1988).

The gene coding for TS in *Lactococcus lactis* can complement the growth of an *E.coli thyA* strain, and its predicted amino acid sequence is the most widely diverged TS sequence known (Ross *et al.*, 1990a; see Section 1.5.3 and Appendix). This gene has also been proposed as a safe selectable marker for use in the generation of strains whose ultimate fate will be in food systems or for release into the environment (Ross *et al.*, 1990b).

### 1.5.2.2 Bacteriophage

TS and DHFR from bacteriophage T4 have been extensively characterized by Mathews and co-workers (1978, 1988). The two enzymes form part of a multi-enzyme aggregate, comprising both T4 and *E.coli* proteins and termed the dNTP synthetase complex (Moen *et al.*, 1988). This complex produces dNTPs in a manner tightly co-ordinated with DNA replication (Allen *et al.*, 1980; Chiu *et al.*, 1982; Mathews *et al.*, 1988; Moen *et al.*, 1988). In addition to their catalytic roles, the two enzymes act as structural components of the T4 virion base plate (Capco *et al.*, 1973), an example of enzymes with a stable protein fold being recruited into a structural role.

Analysis of mutants that lack both TS and DHFR genes has confirmed their role as structural components of the T4 virion, and also shows that the products of these two genes are not essential for bacteriophage T4 viability (Wang & Mathews, 1989).

### 1.5.2.3 Mammals

In mammalian cells, TS gene expression is tightly regulated with respect to the proliferative state of the cell (Conrad, 1971; Navalgund *et al.*, 1980). In human, mouse and Chinese hamster, TS enzyme and mRNA levels are low in quiescent cells but increase approximately 20-fold on stimulation to re-enter the cell cycle (Ayusawa *et al.*, 1986; Jenh *et al.*, 1985; Conrad & Ruddle, 1972). In human and mouse cells, the increase in mRNA levels alone is not sufficient to explain the increase seen in TS activity, thus implying some form of post-transcriptional control mechanism, possibly at the level of mRNA stability (Ayusawa *et al.*, 1986; Jenh *et al.*, 1985). In mouse cells, this control mechanism, at least in part, is at the level of polyadenylation (DeWille *et al.*, 1989). The mouse TS mRNA completely lacks a 3' untranslated region with poly-adenylation commencing at the termination codon (Jenh *et al.*, 1986), whereas the human TS mRNA has a 3' untranslated region of approximately 500 nucleotides (Kaneda *et al.*,

| Herpesvirus | % G+C content | | |
|---|---|---|---|
| | total | unique regions | repeat regions |
| VZV (1) | 46.0 | 44.3 | 59.9 |
| HVS (2) | 45.5 | 35.8 | 70.6 |
| HVA (2) | 47.0 | 37.5 | 74.5 |
| Herpesvirus aotus (3) | ~40.0 | ~49.0 | ~69.0 |

Table 1.2. **A+T rich herpesviruses**.
The G+C% content of the four herpesviruses known to encode TS. Genome compostion data taken from; 1 Davison & Scott, 1986; 2 Fleckenstein *et al.*, 1978; 3 Rüger *et al.*, 1980).

1990). It is still unkown whether these differences in mRNA primary sequence play a role in post-transcriptional regulation.

The TS genes from both mice and humans are 12kb and 16kb in length respectively (Deng *et al.*, 1986; Kaneda *et al.*, 1990). Both genes lack TATAA and CCAAT elements, but contain potential binding sites for other transcription factors such as Spl (Jolliff *et al.*, 1991; Kaneda *et al.*, 1990), features that are characteristic of 'house-keeping' gene promoters (Bird, 1986).

The use of vectors that contain the TS gene, constructed such that its intron sequences have been removed, so-called 'mini-genes', has shown that the inclusion of introns 1 and 2, or 5 and 6, in a mouse TS mini-gene leads to increased levels of transcription compared to one that contains no introns (Deng *et al.*, 1989). In human TS mini-genes, intron 1 may play a role in post-transcriptional regulation (Kadena *et al.*, 1990). Similar stimulatory effects have been previously described in other genes (Buchman & Berg, 1988; Chung & Perry, 1989; Hamer & Leder, 1979). TS is also thought to be a component of a so-called 'metabolon' multi-enzyme complex, that catalyses the synthesis of dNTPs (reviewed in Mathews *et al.*, 1988).

#### 1.5.2.4 **Herpesviruses**

TS has so far been found in four herpesviruses, HVS (Honess *et al.*, 1986), VZV (Thompson *et al.*, 1987), HVA (Richter *et al.*, 1988) and herpesvirus aotus 2 (Bob Honess, personal communication), all of which have A+T-rich genomes (see Table 1.2; Davison & Scott, 1986; Fleckenstein *et al.*, 1978; Rüger *et al.*, 1980). This observation, coupled with the fact that TS has not been found in any of the G+C-rich viruses studied to date, led to speculation that TS may form part of a virus specific replication complex. This could influence the supply of thymidylate directed towards DNA replication and contribute to the nonselected bias in base substitutions that have produced the A+T-rich coding sequences of some subgroups of herpesviruses (Honess, 1984;

| | 122 G 127<br>\|----\| |
|---|---|
| VZV | gdL**G**pi**YG**fqwrh |
| HVS | gdL**G**pv**YG**fqwrh |
| HVA | gdL**G**pv**YG**fqwrh |
| Human | gdL**G**pv**YG**fqwrh |
| Mouse | gdL**G**pv**YG**fqwrh |
| *C.albicans* | gdL**G**pv**YG**fqwrh |
| *P.carinii* | gdL**G**pi**YG**fqwrh |
| *S.cerevisiae* | gdL**G**pv**YG**fqwrh |
| *C.fasiculata* | mdL**G**pv**YG**fqwrh |
| *L.amazonensis* | mdL**G**pv**YG**fqwrh |
| *L.major* | mdL**G**pv**YG**fqwrh |
| *L.tropica* | mdL**G**pv**YG**fqwrh |
| *P.chabaudi* | ndL**G**pi**YG**fqwrh |
| *P.falciparum* | ndL**G**pi**YG**fqwrh |
| *E.coli* | gdL**G**pv**YG**kqwra |
| *B.subtilis* | geL**G**pv**YG**sqwrs |
| *L.lactis* | gtL**G**qr**YG**atvkk |
| *L.casei* | gdL**G**lv**YG**sqwra |
| *Tn*4003 | gnL**G**nv**YG**kqwrd |
| Phage T4 | geL**G**pi**YG**kqwrd |
| Phage phi3t | gtL**G**ha**YG**fqlgk |

Fig. 1.5. **TS sequence alignment.**
Alignment of the 21 TS sequences to illustrate the hypervunerable site of the protozoan enzymes. These are *Crithidium fasciculata, Leishmania amazonensis*, *L.major*, *L.tropica*, *Plasmodium falciparum* and *P. chadaudi*. For full details see text. The location of helix G is shown with numbers referring to VZV TS. For full organism names and the convention for lower case, upper case and **bold**, see legend to Fig. 1.6.

Honess *et al.*, 1986). However, a virally encoded TS gene is not a prerequisite for an A+T-rich genome, as several other herpesviruses with such a genome composition lack TS (Lawrence *et al.*, 1990; Neipel *et al.*, 1991; Bob Honess, personal communication).

No studies to formally investigate the requirement for the virally encoded TS in virus DNA replication have been reported but, the pattern of TS gene expression in HVS and HVA has been linked to their ability to induce tumours in marmosets and other New World primates (Fleckenstein & Desrosiers, 1982). It has been proposed that the continuous unregulated expression of a herpesvirus TS enzyme in persistantly infected lymphoid cells could interfere with the activity of cellular TS, disrupt nucleotide pools, induce mutations and thus contribute to the neoplastic phenotype (Bodemer *et al.*, 1986).

1.5.2.5 **Protozoa**

In protozoa, TS exists on the same polypeptide as DHFR. The DHFR domain forms the amino terminus, TS the carboxy terminus and a junction peptide links the two domains to form a single protein, whose size is species dependent, usually falling in the range of 110 to 140 kDa (Ivanetich & Santi, 1990). One immediately apparent biological advantage of a bifunctional DHFR/TS is metabolic channelling of dihydrofolate as is seen with the *Leishmania tropica* enzyme. Dihydrofolate produced by the TS component of the bifunctional enzyme is channeled to DHFR faster than it is released into the medium, thus the net rate of the sequential reactions is governed by the rate of TS (Meek *et al.*, 1985). This arrangement also results in both activities being co-ordinately regulated, resulting in optimal dTMP synthesis when required.

Another novel feature of protozoan DHFR/TS is the proposition that enzyme activity may be regulated *in vivo* by proteolysis at a 'hypervulnerable' site (Grumont *et al.*, 1986; Garvey & Santi, 1985). Determination of the sequence of the amino-terminal region of a 20 kDa fragment released

upon proteolysis of DHFR/TS from *L.tropica* suggests cleavage immediately before residue 334 (117 in VZV). Cleavage at this site does not completely disrupt the enzymes tertiary structure and, although TS activity is lost, DHFR activity is preserved (Garvey & Santi, 1985). Interestingly, protozoan DHFR/TS sequences have either the amino acid methionine or asparagine at this position, whilst in TS sequences from all other sources a glycine residue is seen (see Fig. 1.5). This is the only amino acid change in this region that differentiates the protozoan TS-DHFR sequences from those of TS from other sources.

The TS activity in *Scenedesmus obliquus*, is inseparable from that of DHFR. The two activities co-purify as a dimer of relative molecular mass of 100 kDa suggesting that they exist as a bifunctional enzyme as seen in other protozoa (Bachmann & Follmann, 1987).

1.5.2.6 **Yeast**

TS specific mRNA production peaks in *Saccharomyces cerevisiae* at the time during the cell cycle when TS levels are at their highest, a point that immediately preceeds maximal rates of DNA synthesis during S-phase (Storms *et al.*, 1984). Yeast cells lack thymidine kinase and are impermeable to thymine, thymidine and thymidylate, thus their only source of thymine nucleotides is from thymidylate produced *de novo* by TS (Bisson & Thorner, 1977). The *S.cerevisiae* temperature sensitive mutant, *cdc-21*, becomes quiescent when shifted to the non-permissive temperature due to the rapid cessation of both nuclear and mitochondrial DNA synthesis (Hartwell, 1973; Newlon & Fangman, 1975). The availability of *S.cerevisiae* strains permeable to dTMP (Wickner, 1975), has permitted the isolation of a dTMP auxotrophic mutant, *tmp-1*, that rapidly loses viability in the absence of exogenous thymidylate (Bisson & Thorner, 1977; Brendel & Langjahr, 1974). Recent sequence analysis has confirmed that both *tmp-1* and *cdc-21* mutants of *S.cerevisiae* contain lesions

in the same gene, that encoding TS (Taylor *et al.*, 1987).

TS is thought to be a component of a multi-enzyme complex found in yeast, similar to those involved in dNTP synthesis seen in other organisms, although only some of its components have been identified (Jazwinski & Edelman, 1984; Mathews *et al.*, 1988). However, the introduction of a large insertion into the TS gene of *S.cerevisiae* has no effect on viability, other than to create a thymidylate auxotrophic phenotype, implying that *S.cerevisiae* TS plays no significant structural role in the complex of replication enzymes (Taylor *et al.*, 1987).

1.5.2.7 **Plants**

Early analysis of TS and DHFR in higher plants was limited to a handful of investigations that described the preliminary measurements of TS activity (Ohyama, 1976; Vandiver & Fites, 1979) and the purification of DHFR from carrot and soybean cells (Albani *et al.*, 1985; Reddy & Rao, 1976). More recently, two groups have produced conflicting reports concerning the nature of TS and DHFR expressed in carrot cells. Nielson and Cella (1988) reported that, in contrast to other eukaryotic cells, the levels of TS did not vary significantly, in the carrot cells growth cycle, except for a slight decrease in the late stationary phase. The same authors described the purification of TS activity and found it inseparable from that of DHFR activity. The fact that both activities were ascribed to a protein complex of 185 kDa and that DHFR levels vary in concentration with TS during the growth cycle (Albani *et al.*, 1985), led to the proposal that TS and DHFR from carrot cells may exist as a bifunctional enzyme, similar to, albeit larger than, those found in protozoa (Ivantiech & Santi, 1990). In contrast, Goodman and co-workers described a 285 kDa complex comprised of five polypeptides. TS activity has been assigned in this case to a 70 kDa polypeptide whilst DHFR has been located to a 45 kDa polypeptide within the complex (Toth *et al.*, 1987). Further characterization of these activities by workers in

```
                                           17      A      31  -1               42       i     54                 68    B
                                           |-------------|  |^|                |^^^^^^^^^^^|                   |-------
VZV                  mgdlsewtkvpgftltgelqYlkqvddILryGvrkr.........DRTGiGTLSlFG.mqaRYnLrne.FPLLTTKR.VfwraVveEL
HVS                         msthteeqhgehqYlsqvqhILnyGsfkn.........DRTGtGTLSiFG.tqsRFsLene.FPLLTTKR.VfwrgVveEL
HVA                               meelhaehqYlsqvkhILncGnfkh.........DRTGvGTLSvFG.mqsRYsLekd.FPLLTTKR.VfwrgVveEL
Human        mpvagselprrplppaaqerdaeprpphgelqYlgqiqhILrcGvrkd.........DRTGtGTLSvFG.mqaRYsLrde.FPLLTTKR.VfwkgVleEL
Mouse        mlvvgsel.....qsdaqqlsaeaprhgelqYlrqvehILrcGfkke.........DRTGtGTLSvFG.mqaRYsLrde.FPLLTTKR.vfwkgVleEL
C.albicans                  mtvspntaeqaYldlckrIIdeGehrp.........DRTGtGTkSlFappqlRFdLsndtFPLLTTKK.VfskgIihEL
P.carinii                         mvnaeeqqYlnlvqyIInhGedrp.........DRTGtGTLSvpapsplKFsLrnktFPLLTTKR.VflrgVieEL
S.cerevisiae                  mtmdgknkeeeqYldlckrIIdeGefrp.........DRTGtGTLSlFappqlRFsLrddtFPLLTTKK.VftrgIilEL
C.fasciculata      NH2 -223- mkyvphnaeerqYlelidrIMktGlvke.........DRTGvGTISlFG.aqm.FsLrdnqLPLLTTKR.VfwrgVceEL
L.amazonensis      NH2 -224- ckyvprnheerqYlelidrIMktGiake.........DRTGvGTLSlFG.aqmRFsLrdnrLPLLTTKR.VfwrgVceEL
L.major            NH2 -224- ckyvprnheerqYlelidrIMktGivke.........DRTGvGTISlFG.aqmRFsLrdnrLPLLTTKR.VfwrgVceEL
L.tropica          NH2 -224- ckyvprnheerqYlelidrIMktGivke.........DRTGvGTISlFG.aqmRFsLrdnrLPLLTTKR.VfwrgVceEL
P.chabaudi         NH2 -280- nsikykhhpeyqYlniiydIMhGnkqd.........DRTGvGvLSkFG.ymmKFnLsey.FPLLTTKK.LfvrgIieEL
P.falciparum       NH2 -315- nslkykyhpeyqYlniiydIMmnGnkqs.........DRTGvGvLSkFG.yimKFdLsqy.FPLLTTKK.LflrgIieEL
E.coli                               mkqYlelmqkVLdeGtqkn.........DRTGtGTLSiFG.hqmRFnLqdg.FPLVTTKR.chlrsIihEL
B.subtilis                           mkqYkdfcrhVLehGekkg.........DRTGtGTIStFG.yqmRFnLreg.FPMLTTKK.LhfksIahEL
L.lactis                 mtyadqvFkqniqnILdnGvfsenarpkyk...DgqmanSkyvtG.sfvtYdLqkgeFPI.TTlRpIpiksaikEL
L.casei                        mleqpYldlakkVLdeGhfkp.........DRThtGTySiFG.hqmRFdLskg.FPLLTTKK.VpfglIksEL
Tn4003                        mynpfdeaYhglceeILeiGnrrd.........DRThtGTISkFG.hqlRFdLtkg.FPLLTTKK.VsfklVatEL
Phage T4                             mkqYqdlikdIfenGyetd.........DRTGtGTIalFG.sklRWdLtkg.FPaVTTKK.LawkaciaEL
Phage phi3T            mtqfdkqYnsiikdIInnGisdeefdvrtkwdsDgTpahTLSvms.kqmRFdnse..vPILTTKK.VawktaikEL
```

```
              82    86  C  93              99 D 103                                                             122 G  127
              -----|   |------|             |---|                                                                |----|
VZV           LWFIrGsTd.skeLaakd.......IHIWdiygsskflnrngfhkrht...........................gDLGPIYGfQWRhFg
HVS           LWFIrGsTd.skeLsaag.......VHIWdangsrsfldklgfydrde..........................gDLGPVYGfQWRhFg
HVA           LWFIrGsTd.skeLaasg.......VHIWdangsrsyldklglfdree..........................gDLGPVYGfQWRhFg
Human         LWFIkGsTn.akeLsskg.......VKIWdangsrdfldslgfstree..........................gDLGPVYGfQWRhFg
Mouse         LWFIkGsTn.akeLsskg.......VRIWdangsrdfldslgfsarqe..........................gDLGPVYGfQWRhFg
C.albicans    LWFVaGsTd.akiLsekg.......VKIWegngsrefldklglthrre..........................gDLGPVYGfQWRhFg
P.carinii     LWFIrGeTd.slkLrekn.......IHIWdangsreyldsiglttkrqe.........................gDLGPIYGfQWRhFg
S.cerevisiae  LWFLaGdTd.anlLseqg.......VKIWdgngsreyldkmgfkdrkv..........................gDLGPVYGfQWRhFg
C.fasciculata IWFLrGeTn.ahvLadkd.......IHIWdgngsrefldsrgltenke..........................mDLGPVYGfQWRhFg
L.amazonensis LWFLrGeTn.aqlLadkd.......IHIWdgngsrefldsrgltente..........................mDLGPVYGfQWRhFg
L.major       LWFLrGeTs.aqlLadkd.......IHIWdgngsrefldsrgltenke..........................mDLGPVYGfQWRhFg
L.tropica     LWFLrGeTs.aqlLtdkd.......IHIWdgngsrefldsrgltenke..........................mDLGPVYGfQWRhFg
P.chabaudi    LWFIrGeTn.gntLlekn.......VRIWeangtrefldnrklfhrev..........................nDLGPIYGfQWRhFg
P.falciparum  LWFIrGeTn.gntLlnkn.......VRIWeangtrefldnrklfhrev..........................nDLGPIYGfQWRhFg
E.coli        LWFLqGdTn.iayLhenn.......VtIWdewaden..........................gDLGPVYGkQWRaWp
B.subtilis    LWFLkGdTn.vryLqeng.......VRIWnewaden..........................gELGPVYGsQWRsWr
L.lactis      MWiyqdqTselsvLeeky.....gVKyWgewgigd..........................gtiGqrYGatvKkYn
L.casei       LWFLhGdTn.irfLlqhr.......nHIWdewafekwvksdeyhgpdmtdfghrsqkdpefaavyheemakfddrvlhddafaakygDLGlVYGsQWRaWh
Tn4003        LWFIkGdTn.iqyLlkyn.......nnIWnewafenyvqsddyhgpdmtdfghrsqqdpefneqykeemkkfkerilnddafakkygnLGnVYGkQWRdWe
Phage T4      IWFLsGsTn.vndLrliqhdsliqgktVWdenyenqakdlgyhs..........................gELGPIYGkQWRdF.
Phage phi3T   LWiwqlksndvteLnkmg.......VHIWdqwkqed..........................gtiGhaYGfQlgkkn
```

```
               147          H          159  165  v    173     I     183    iv      192                     195       iii         210
               |------------|    |^^^|  |------|  |^^^^^^^^^|                      |^^^^^^^^^^^^|  |----
VZV            aeykdcqsnylqqgiDQLqtVIdtIKtnPesRRmIissWNpkdip.lMvLPPCHtLcQFYVang...........eLscqVYQRSgDMgLGVPFN
HVS            aeykgvgrdykgegvDQLkqLIdtIKtnPtdRRmLmcaWNvsdip.kMvLPPCHvLsQFYVcdg...........kLscqLYQRSaDMgLGVPFN
HVA            aeyqglkhnyggegvDQLkqIIntIHtnPtdRRmLmcaWNvldvp.kMALPPCHvLsQFYVcdg...........kLscqLYQRSaDMgLGVPFN
Human          aeyrdmesdysgqgvDQLqrVIdtIKtnPddRRiImcaWNprdlp.lMALPPCHaLcQFYVvns...........eLscqLYQRSgDMgLGVPFN
Mouse          aeykdmdsdysgqgvDQLqkVIdtIKtnPddRRiImcaWNpkdlp.lMALPPCHaLcQFYVvng...........eLscqLYQRSgDMgLGVPFN
C.albicans     aeykdcdsdytgqgfDQLqdVIkkLKtnPydRRiImsaWNppdfa.kMALPPCHvFcQFYVnfptslpdpnnpkqaktakpkLscILYQRScDMgLGVPFN
P.carinii      aeyidcktnyiggqvDQLanIIqkIRtsPydRRlIlsaWNpadle.kMALPPCHmFcQFYVhipsnnh..........rpeLscqLYQRScDMgLGVPFN
S.cerevisiae   akyktcddaytgqgiDQLkqVIhkLKtnPydRRiImsaWNpadfd.kMALPPCHiFsQFYVsfpkegeg...........sgkprLscILYQRScDMgLGVPFN
C.fasciculata  adykgfdanyd.egvDQIktIVetLKtn..dRRlLvtaWNpcalh.kMAVrPCHlLgQFYVntqtk...........eLscmLYQRScDMgLGVPFN
L.amazonensis  aeyrgleanydgegvDQIkfIVetIKanPndRRlLftaWNpcalh.kMALPPCHlLaQFYVnteks...........eLscmLYQRScDMgLGVPFN
L.major        adykgfeanydgegvDQIklIVetIKtnPndRRlLvtaWNpcalq.kMALPPCHlLaQFYVntdts...........eLscmLYQRScDMgLGVPFN
L.tropica      adykgfeanydgegvDQIklIVetIKtnPndRRlLvtaWNpcalq.kMAVPPCHlLaQFYVntdts...........eLscmLYQRScDMgLGVPFN
P.chabaudi     aeytdmhadykdkgvDQLknIInlIKndPtcRRiIlcaWNvkdld.qMALPPCHiLcQFYVfdg...........kLsciMYQRScDLgLGVPFN
P.falciparum   aeytnmydnyenkgvDQLknIInlIKndPtsRRiLlcaWNvkdld.qMALPPCHiLcQFYVfdg...........kLsciMYQRScDLgLGVPFN
E.coli         tp........dgrhiDQIttVLnqLKndPdsRRiIvsaWNvgeld.kMALaPCHaFfQFYVadg...........kLscqLYQRScDVfLGLPFN
B.subtilis     ga........dgetiDQIsrLIedIKtnPnsRRlIvsaWNvgeid.kMALPPCHcLfQFYVsdg...........kLscqLYQRSaDVfLGVPFN
L.lactis       i..................IgkLLegLaknPwnRRnIinlWqyedfeetegLlPCafqtmFdVrrekdgqi...........yLdatLIQRSnDMlVahhiN
L.casei        ts.......kgdtiDQLgdVIeqIKthPysRRlIvsaWNpedvp.tMALPPCHtLyQFYVndg...........kLslqLYQRSaDIfLGVPFN
Tn4003         dk........ngnhyDQLksVIqqIKtnPnsRRhIvsaWNpteid.sMALPPCHtMfQFYVqeg...........kLncqLYQRSaDIfLGVPFN
Phage T4       ...........ggvDQIieVIdrIKklPndRRqIvsaWNpaelk.yMALPPCHmFyQFnVrng...........yLdlqwYQRSvDVfLGLPFN
Phage phi3T    rs.......lngekvDQVdyLLhqLKnnPssRRhItmlWNpddld.aMALtPCvyetQWYVkqg...........kLhleVraRSnDMaLGnPFN
```

```
              J             230        237       ii       251      K    258           267    270
              ----------------|        |^^^^^^^^^^|      |--------|                  |---|
VZV           IAgYaLLTyIVAhvtgLktGDLIhtmGdaHIYlnHIDaLKvQLaRsPkpf.PcLkiirnv....tdIndF........kwdDfqLdgYnp.hppLKmeMAL
HVS           IASYsLLTcMIAhvtnLvpGEFIhtiGdaHIYvdHIDaLKmQLtRtPrpf.PtLrfarnv....scIddF........kadDiiLenYnp.hpiIKmhMAV
HVA           IASYsLLTcMIAhvtdLvpGEFIhtlGdaHVYvnHVDaLteQLtRtPrpf.PtLkfarkv....asIddF........kanDiiLenYnp.ypsIKmpMAV
Human         IASYaLLTyMIAhitgLkpGDFIhtlGdaHIYlnHIEpLKiQLqRePrpf.PkLrilrkv....ekIddF........kaeDfqIegYnp.hptIKmeMAV
Mouse         IASYaLLTyMIAhitgLqpGDFVhtlGdaHIYlnHIEpLKiQLqRePrpf.PkLkilrkv....etIddF........kveDfqIegYnp.hptIKmeMAV
C.albicans    IASYaLLTkMIAhvvdMdcGEFIhtlGdaHVYldHIDaLKeQFeRiPkqf.PkLvikeerkneiksIddF........kfeDfeIvgYep.yppIKmkMsV
P.carinii     IASYaLLTcMIAhvcdLdpGDFIhvmGdcHIYkdHIEaLqqQLtRsPrpf.PtLslnrsi....tdIedF........tldDfnIqnYhp.yetIKmkMsI
S.cerevisiae  IASYaLLTrMIAkvvdMepGEFIhtlGdaHVYkdHIDaLKeQitRnPrpf.PkLkikrdv....kdIddF........kltDfeIedYnp.hprIqmkMsV
C.fasciculata IASYaLLTiLIAkatgLrpGELVhtlGtaHVYsnHVEaLKeQLqRvPvaf.PvLvfkker....efLedY........estDmeVvdYvp.yppIKmeMAV
L.amazonensis IASYaLLTiLIAkatgLrpGELVhtlGdaHVYrsHIDaLKaQLeRvPhaf.PtLvfkeer....qfLedY........elmDmeVidYvp.hppIKmeMAV
L.major       IASYaLLTiLIAkatgLrpGELVhtlGdaHVYrnHVDaLKaQLeRvPhaf.PtLifkeer....qyLedY........eltDmeVidYvp.hpaIKmeMAV
L.tropica     IASYaLLTiLIAkatgLrpGELVhtlGdaHVYrnHVDaLKaQLeRvPhaf.PtLifkeer....qyLedY........eltDmeVidYvp.hpaIKmeMAV
P.chabaudi    IASYsIFTyMIAqvcnLqpaEFIhvlGnaHVYnnHVEsLKvQLnRtPypf.PtLklnpei....knIedF........tisDftVqnYvh.hdkIsmdMAa
P.falciparum  IASYsIFThMIAqvcnLqpaqFIhvlGnaHVYnnHIDsLKiQLnRiPypf.PtLklnpdi....knIedF........tisDftIqnYvh.hekIsmdMAa
E.coli        IASYaLLvhMmAqqcdLevGDFVwtgGdtHlYsnHMDqtHlQLsRePrpl.PkLiikrkp....esIfdY........rfeDfeIegYdp.hpgIKapVAI
B.subtilis    IASYaLLTmIIAhvtgLepGEFIhtfGdvHIYqnHIEqVnlQLeRdvrpl.PqLrfarkv....dsIfnF........afeDfiIedYdp.hphIKgaVsV
L.lactis      amqYvaLqmMIAkhfswkvGkFfyfvnnlHIYdnqfEqanelMkRtasekePrLvlnvp.....dgtnfF........dikpeDfeLvdYepvkpqLKfdLAI
L.casei       IASYaLLThLVAhecgLevGEFIhtfGdaHLYvnHLDqIKeQLsRtPrpa.PtLqlnpdk....hdIfdF........dmkDikLlnYdp.ypaIKapVAV
Tn4003        IASYaLLThLVAkecgLevGEFIhtfGdaHIYsnHMDaIHtQLsRdsylp.PqLkintdk.....sIfdi........nyeDleLinYes.hpaIKapIAV
Phage T4      IASYatLvhIVAkmcnLipGDLIfsgGntHIYmnHVEqcKeiLrRePkel.ceLvisglpykfrylstkeqlkyvlklrpkDfvLnnYvs.hpxIKgkMAV
Phage phi3T   VfqYnVLqrMIAqvtgyelGEYIfniGdcHVYtrHIDnLKiQMeReqfea.PeLwinpev....kdfynF........tvdDfkLinYkh.gdkLlfeVAV
```

Fig. 1.6. **TS sequence alignment.**
Alignment of the 21 TS sequences reported to date. **BOLD UPPERCASE** letters indicate completely conserved residues in all 21 TS sequences. UPPERCASE letters indicate a residue in one of the groups (I/L/M/V) (D/E) (S/T) (F/Y/W) (F/L) (K/R) in all 21 TS sequences. The numbering scheme refers to VZV TS. The secondary structure elements (^^^^^ = β-sheet; ----- = α-helix) are based on TS from *L.casei* (see Hardy *et al.*, 1987; Perry *et al.*, 1990). TS sequences shown include those for VZV (Thompson *et al.*, 1987); HVS (Honess *et al.*, 1986); HVA (Richter *et al.*, 1988); human (Takeishi *et al.*, 1985); mouse (Perryman *et al.*, 1986); *Candida albicans* (Singer *et al.*, 1989); *Pneumocystis carinii* (Edman *et al.*, 1989); *Saccharomyces cerevisiae* (Taylor *et al.*, 1987); *Crithidia fasiculata* (Hughes *et al.*, 1989); *Leishmania amazonensis* (Nelson *et al.*, 1990); *Leishmania major* (Beverley *et al.*, 1986); *Leishmania tropica* (Grumont *et al.*, 1986); *Plasmodium chabaudi* (Cowman & Lew, 1989); *Escherichia coli* (Belfort *et al.*, 1983b); *Bacillus subtilis* (Iwakura *et al.*, 1988); *Lactococcus lactis* (Ross *et al.*, 1990a); *Lactobacillus casei* (Maley *et al.*, 1979b); *Staphylococcus aureus* transposon *Tn*4003 (Rouch *et al.*, 1989); coliphage T4 (Chu *et al.*, 1984); *B.subtilis* phage phi3T (Kenny *et al.*, 1985). Only the TS portion of the DHFR/TS sequences of *Leishmania* spp., *C.fasiculata* and *Plasmodium* spp. are shown (see Perry *et al.*, 1990). An additional, larger scale, loose copy of this figure is also provided for easy reference.

the same group has shown that the two enzymes are co-ordinately regulated (Lazar *et al.*, 1989). These studies have been performed in overproducer cell lines and the elevated levels of specific mRNAs should facilitate cloning and sequencing of the TS and DHFR from carrot cells, which may resolve some of the discrepancies reported by these two groups.

### 1.5.3 Amino acid sequence homology

The amino acid sequence of TS from *Lactobacillus casei* was determined in 1979 by Maley and co-workers (Bellisario *et al.*, 1979; Maley *et al.*, 1979a,b). Since then, the complete nucleotide sequence of this gene has been determined (Andrews *et al.*, 1985; Pinter *et al.*, 1988) and shown to be identical to that found by amino acid sequencing techniques (Maley *et al.*, 1979b). Moreover, the TS gene from twenty other organisms has been cloned, sequenced and the amino acid sequence predicted (see legend from Fig. 1.6 for references). A striking feature of the TS sequences is the very high degree of amino acid conservation, approximately 11% (27 out of a core of ~250 residues) of the residues are absolutely conserved in the 21 known complete sequences (see Fig. 1.6). If the two most divergent TS sequences, those of bacteriophage ø3T and *L. lactis* are excluded, the degree of conservation rises to approximately 24% (61 out of a core of ~250 residues). This is equivalent of one in every four residues being completely conserved in 19 TS sequences (but see Appendix). Such a degree of amino acid sequence conservation places TS amongst the most highly conserved enzymes, a group that includes proteins such as the cytochromes (Dickerson 1971). The cytochromes however, are much smaller than TS (less than half the number of amino acids) and much of their surface is involved in interactions with oxidase and reductase. It has been suggested that some reasons why TS is so highly conserved are related to the chemical reaction catalysed and the conformational dynamics that occur during catalysis (Finer-Moore *et al.*, 1990). Conservation may

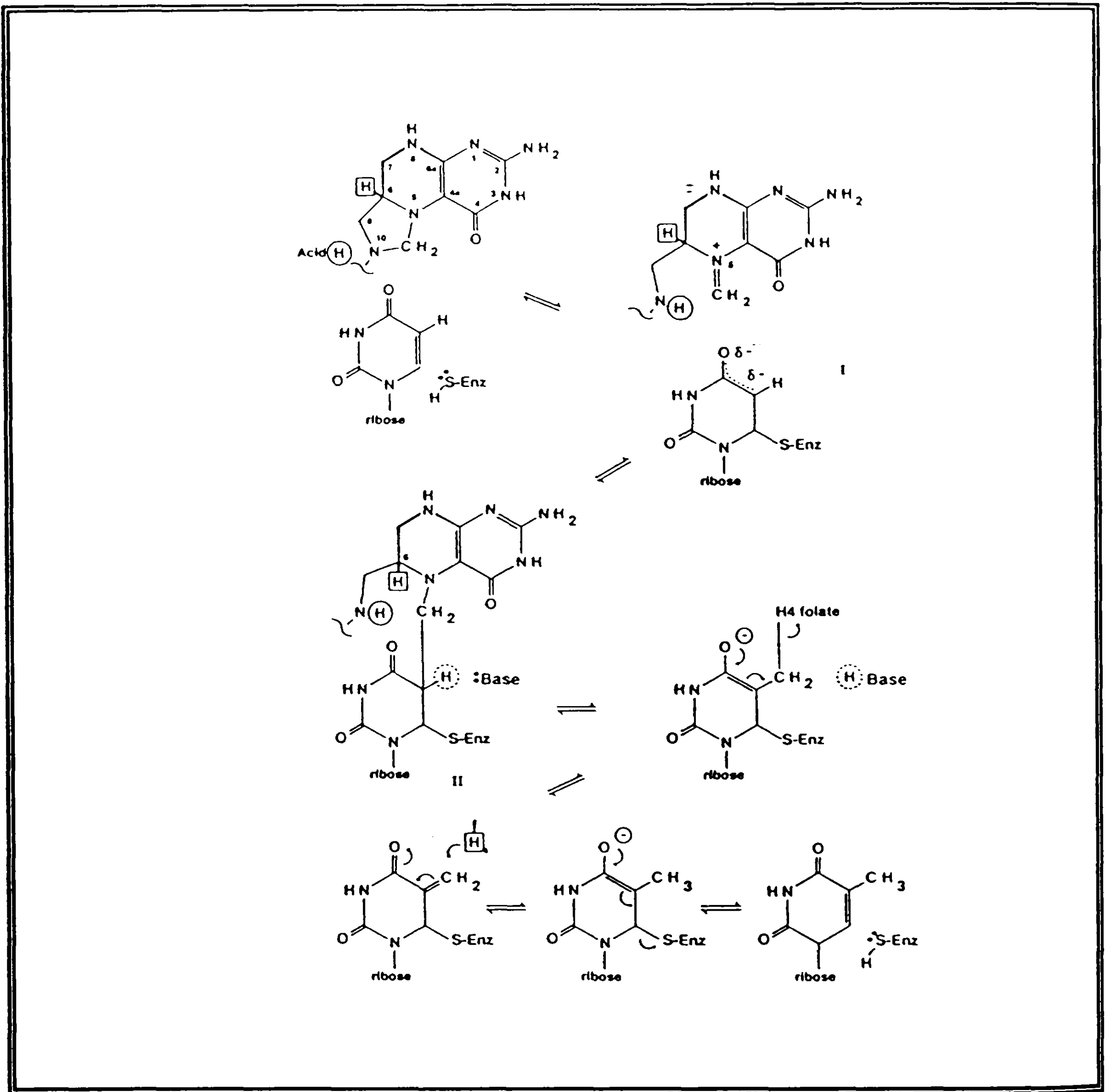



Fig. 1.7. **Catalytic mechanism of TS.**
The proposed mechanism for methyl transfer to dUMP by TS, showing the numbering scheme for $CH_2FH_4$. The formation of a covalent bond between the catalytic thiol of C198 of VZV TS and C-6 of dUMP activates C-5 of dUMP for condensation with the one-carbon unit of $CH_2FH_4$ (intermediate I). Elimination and hydride transfer from C-6 of the cofactor result in the generation of dTMP and $FH_2$ (taken from Finer-Moore *et al.*, 1990).

also be due to the need to maintain interactions with other enzymes such as DHFR.

The probability of structural homology at such levels of primary sequence conservation is very high. The value for the minimum amino acid sequence homology between a pair of proteins likely to reflect significant structural similarity has been estimated to be 26% (Sander & Schneider, 1991). The likelihood of structural homology, based on the high level of amino acid identity between *E.coli* and *L.casei* TS (60%), has facilitated the interpretation of the X-ray data obtained from *E.coli* TS by molecular replacement, by using the co-ordinates of the *L.casei* TS structure (Hardy *et al.*, 1987; Perry *et al.*, 1990). In addition, the high degree of amino acid sequence identity between VZV and *E.coli* TS (48%), has allowed models of the VZV TS to be constructed with a high degree of confidence, from the known structures of unliganded, and ligand bound forms of *E.coli* TS (see Section 1.10; Celia Schiffer and Robert Stroud, personal communication).

## 1.6 Catalytic mechanism of TS

It has been recognized for over 30 years that $N^5,N^{10}$-methylene tetrahydrofolate serves the dual function of both one carbon donor and reductant in the reaction catalysed by TS (Humpreys & Greenberg, 1958; Friedkin, 1959). Friedkin first proposed in 1959 that the reaction involved the condensation of $N^5,N^{10}$-$CH_2FH_4$ with dUMP to produce a 5-thymidylyl-tetrahydrofolate intermediate. This would then undergo a 1,3-hydride shift to produce dTMP and $FH_2$. The first step represents an electrophilic substitution reaction in which the one-carbon unit of $N^5,N^{10}$-$CH_2FH_4$ replaces the hydrogen at the 5-position of dUMP without a change in the oxidation level. The second step involves hydride transfer to the exocyclic methylene group of dUMP, by the hydrogen originating from the 6-position of the co-factor, to produce free dTMP and $FH_2$ (see Fig. 1.7).

Since then, numerous groups have conducted exhaustive

investigations into the reaction mechanism of TS. This includes a fully detailed stereospecific reaction mechanism that was described after the determination of the crystal structure of the covalent inhibitory ternary complex of TS from *E.coli* (Finer-Moore *et al.*, 1990; Matthews *et al.*, 1990b). However, a brief summary of the early data, is necessary for a complete understanding of the reaction mechanism.

### 1.6.1 Identification of catalytic residues

Studies of the inhibitory ternary complex, comprised of FdUMP and $N^5,N^{10}-CH_2FH_4$ covalently bound to TS, have identified cysteine 198 (C183 in VZV) of *L.casei* TS as forming a covalent bond to FdUMP (Bellisario *et al.*, 1979). Further studies proved the existence of a covalent bond linking C-5 of FdUMP to the methylene group and N-5 of the co-factor (Pellino & Danenberg, 1985). Moore and co-workers (1986) proved the existence of a stable binary complex, consisting of dUMP (or dTMP) covalently bound to TS. Both findings are as predicted by the proposed reaction mechanism (Friedkin, 1959). Moore and co-workers (1986) also revealed that FdUMP and dUMP (and also dTMP) were linked to the same residue, cysteine 198, thus augmenting the use of the FdUMP based covalent ternary complex for studies of the catalytic mechanism of TS.

In additon to the active-site cysteine residue, arginine and tyrosine residues have been proposed to play a role in catalysis, on the basis of chemical modification experiments (Lewis & Dunlap, 1981; Rosson *et al.*, 1980). Structural studies have implicated four highly conserved arginine residues in the coordination of the 5'-phosphate of dUMP and two conserved tyrosine residues are also predicted to form part of the active site (Finer-Moore *et al.*, 1990; Matthews *et al.*, 1990b).

### 1.6.2 Ordered binding of substrate and cofactor

Kinetic investigations by Daron and Aull (1978) suggested that the process of substrate addition and

product release is an ordered sequential mechanism. dUMP binds first, followed by cofactor to form the covalent ternary complex. Catalysis takes place, and the oxidised cofactor is released followed by the product, dTMP. However, polyglutamate derivatives of $N^5,N^{10}-CH_2FH_4$ serve as more efficient cofactors, and indeed are the normal form of folates found *in vivo* (Kisliuk & Gaumont, 1974; Galivan *et al.*, 1976a; Dolnick & Cheng, 1978; Eto & Krumdieck, 1981; Foo & Shane, 1982). When Galivan and co-workers (1976a) demonstrated that the pentaglutamate form of the cofactor could form a binary complex with TS in the absence of dUMP, this suggested that the ordered sequential reaction seen *in vitro* with the monoglutamate form of the co-factor (Daron & Aull, 1978), may not follow *in vivo*. Indeed, Lu and co-workers (1984) showed that for the tetraglutamate form of $N^5,N^{10}-CH_2FH_4$, the order of substrate binding and product release was reversed *in vivo*.

### 1.6.3 Conformational changes associated with catalysis: ternary complex formation

A series of conformational changes have been associated with the catalytic process performed by TS. The binding of dUMP or FdUMP causes a change in the fluorescence and ultra-violet (UV) spectra, indicating changes in the environment of tyrosine and tryptophan residues (Donato *et al.*, 1976; Galivan *et al.*, 1975; Leary *et al.*, 1975). However, no change was seen in the circular dichromic (CD) spectrum upon the binding of FdUMP or cofactor independently to *L.casei* TS (Manavalan *et al.*, 1986). It has been suggested that, because of the ordered sequential mechanism of binding seen *in vitro*, these spectral changes are associated with the creation of a binding site for $N^5,N^{10}-CH_2FH_4$ (Santi & Danenberg, 1984).

The UV and CD spectra alter considerably upon the addition of $N^5,N^{10}-CH_2FH_4$ to the binary complex, due to subsequent ternary complex formation (Donato *et al.*, 1976; Danenberg *et al.*, 1974; Santi *et al.*, 1974; Manavalan *et al.*, 1986). Formation of the ternary complex results in a

considerably more compact structure for the protein, corresponding to an apparent decrease of 3.5% in the Stokes radius (Lockshin & Danenberg, 1980) and correlating well with the fluorescence quenching observed upon ternary complex formation (Bellisario *et al.*, 1976). These factors provide evidence that changes in the protein conformation, as opposed to changes in the environment of the ligand, contribute to the majority of the observed variation in the fluorescence spectrum (Santi & Danenberg, 1984).

TS from all sources is thought to be an obligate homodimer with residues from both subunits forming two active sites (Hardy *et al.*, 1987). In the presence of co-factor, FdUMP binds with a stoichiometry of approximately 2 mols FdUMP per mol of enzyme forming a FdUMP.$N^5$,$N^{10}$-$CH_2FH_4$.TS ternary complex (the 2:2:1 complex; Danenberg *et al.*, 1974; Danenberg & Danenberg, 1979; Plese & Dunlap, 1977; Santi & Danenberg, 1984). However, the 2:2:1 ternary complex decays within 24h to a significantly more stable complex with only 1 mol FdUMP bound per mol of enzyme, (Aull *et al.*, 1974; Santi & Danenberg, 1984), suggesting that the two binding sites are not equivalent. Furthermore, in the presence of triglutamated forms of $N^5$,$N^{10}$-$CH_2FH_4$, the two binding sites show a twenty-fold difference in affinity for dUMP (Lockshin and Danenberg, 1979).

In the absence of co-factor, only a single molecule of dUMP or FdUMP is bound per mol of enzyme (Galivan *et al.*, 1976a; Beaudette & Kisliuk, 1980; Langenbach, 1976), although Plese and Dunlap (1978) report that approximately 1.5 nucleotides were bound per enzyme molecule.

These observations have been proposed to reflect a sequential action of the catalytic sites, where the binding of one molecule of dUMP at either site results in minor conformational changes, and renders the second site inaccessible. This corresponds to the observed changes in the fluorescence and UV spectra (Donato *et al.*, 1976; Galivan *et al.*, 1975; Leary *et al.*, 1975). Binding of co-factor causes a much greater conformational change, thus

providing access to the second site (Donato *et al.*, 1976; Danenberg *et al.*, 1974; Santi *et al.*, 1974; Manavalan *et al.*, 1986). A second dUMP molecule is then bound and the interaction of a second cofactor molecule completes the formation of the ternary complex (Santi & Danenberg, 1984).

### 1.6.4 Stereospecific reaction mechanism

The current, widely accepted chemical basis of the reaction catalysed by TS was described by Santi and Danenberg (1984) and is shown in Fig. 1.7 (Pogolotti & Santi, 1977; Heidelberger *et al.*, 1983; Santi & Danenberg, 1984; Finer-Moore *et al.*, 1990). The covalent bond formed between the active site cysteine residue and C-6 of dUMP serves to activate C-5 of dUMP for condensation with the one-carbon unit of $N^5,N^{10}$-$CH_2FH_4$. An activated form of the cofactor, the iminium ion 5-$CH_2$=$FH_4$, has been postulated to act as an electophilic species in the condensation reaction (Kallen & Jencks, 1966; Benkovic & Bullard, 1973). An acidic group on the enzyme has been suggested to catalyse opening of the imidazoline ring of $N^5,N^{10}$-$CH_2FH_4$ by protonation at N-10. A basic group of the enzyme was postulated to abstract a proton from C-5 of dUMP (Benkovic, 1980; Fife & Pellino, 1981). Subsequent ß-elimination and hydride transfer from C-6 of the cofactor would then generate dTMP and dihydrofolate.

In addition, the TS catalysed reaction proceeds under tight stereochemical control. A detailed experimental analysis led to the proposal that dUMP would bind in the syn-conformation, to account for the observed stereochemical nature of the reaction (Slieker & Benkovic, 1984). CD and $^{19}$FNMR studies suggested that dUMP is bound in a syn-conformation in the non-covalent TS.dUMP binary complex (Leary *et al.*, 1975; Lewis & Dunlap, 1981). In contrast, molecular modelling of dUMP, based upon the crystal structure of *L.casei* TS with bound phosphate suggests that dUMP would bind with O-2 anti to the ribose ring (Hardy *et al.*, 1987). In studies of the covalent ternary complex, dUMP has indeed been shown to be bound in the anti-

conformation by $^{19}$FNMR studies of the *L.casei* TS.FdUMP.$N^5,N^{10}$-$CH_2FH_4$ complex (Byrd *et al.*, 1978) and by X-ray crystallographic studies of the *E.coli* TS.dUMP.CB3717 complex (Montfort *et al.*, 1990). FdUMP, a close analogue of the substrate dUMP, adopts the same conformation in both the *E.coli* TS.FdUMP.CB3717 (Matthews *et al.*, 1990a) and TS.FdUMP.$N^5,N^{10}$-$CH_2FH_4$ complexes (Matthews *et al.*, 1990b).

### 1.6.5 Reaction mechanism based upon X-ray crystallographic data

Recently, two groups have independantly published detailed schemes for a stereospecific mechanism for the TS catalysed reaction. Both groups have reported essentially identical structures for the ternary complexes, but have described different mechanisms for the reaction. For this reason, the reports by the two groups are discussed separately, followed by a brief summary of the differences.

#### 1.6.5.1 Model of Matthews and co-workers

Matthews and co-workers (1990b) determined the structure of the TS.FdUMP.$N^5,N^{10}$-$CH_2FH_4$ covalent ternary complex from *E.coli*, which was thought to accurately represent the catalytic intermediate of the TS catalysed reaction (Danenberg, 1977). They have shown that FdUMP and $CH_2FH_4$ are precisely bound by a number of highly conserved residues. FdUMP is found in the anti-conformation. The bond between Cys146 and C6 of FdUMP, and the C5-C11 bond are in a trans-diaxial arrangement. The tetrahydrozine ring of the co-factor has its C6-C9 bond axial, C6-H equatorial. The methylene bridge of the nucleotide orients the C5-H such that it is also equatorial. These observations concur with $^{19}$FNMR studies of the *L.casei* TS.FdUMP.$N^5,N^{10}$-$CH_2FH_4$ ternary complex (Byrd *et al.*, 1978).

To assess the validity of the covalent ternary complex as a model for the catalytic intermediate, Matthews and co-workers (1990b) attempted to predict the stereochemical course of the reaction, from the *E.coli* ternary complex structure, and compare this with experimental results of

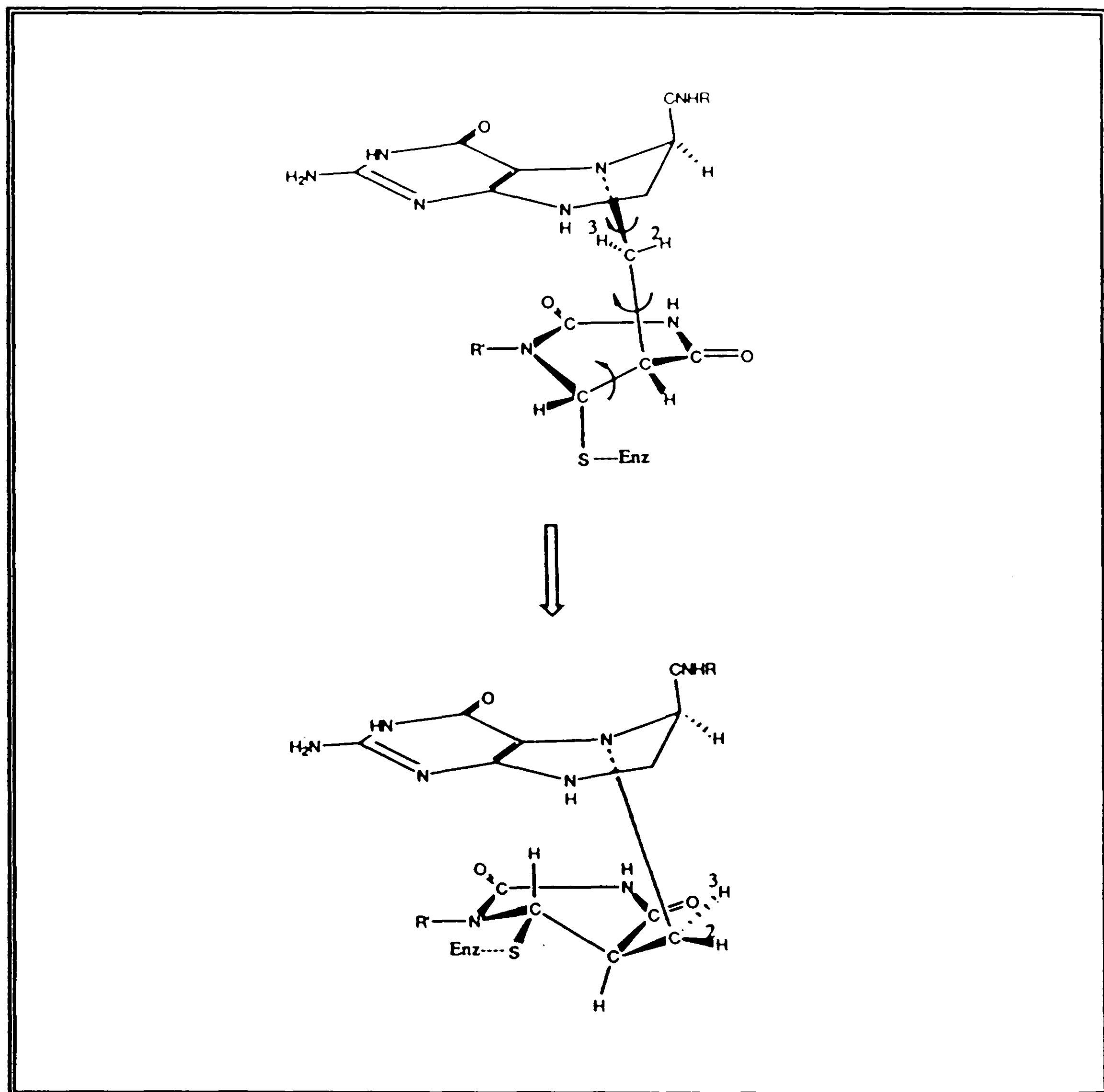



Fig. 1.8. **Catalytic mechanism of TS.**
The proposed conformational isomerization of the catalytic intermediate that preceeds β-elimination (see text for details). R' represents deoxyribose, CNHR represents the remainder of the cofactor molecule, and Enz-S indicates the catalytic thiol, C198 in VZV TS (taken from Matthews *et al.*, 1990b).

Sleiker and Benkovic (1984). They argued that if this complex was an accurate representation of the catalytic intermediate, then the ensuing elimination reaction should result in the formation of a methyl group at C5 of dUMP with the S configuration. However, the first step in the reaction is the elimination across the C5-C6 bond, which is in a syn-conformation (this is henceforth referred to as syn-elimination). This would result in the formation of an exocyclic methylene derivative of dUMP having a Z configuration. Since the C6 hydrogen of $N^5,N^{10}$-$CH_2FH_4$ approaches from above the *si* face of $CH_2$=dUMP the resultant methyl group would have the configuration R.

An alternative is to consider the possibility of anti-elimination across the C5-C6 bond. Indeed, anti-eliminations are more favoured than syn-eliminations, and the fact that thiols are better leaving groups than tertiary amines, adds weight to this notion. However, anti-elimination would result in the bizarre situation where the substrate and cofactor are regenerated. Clearly these contradictory observations suggest that the structure of the TS.FdUMP.$N^5,N^{10}$-$CH_2FH_4$ ternary complex does not accurately represent the catalytic intermediate (Matthews *et al.*, 1990b).

To rationalize these observations, Matthews and co-workers (1990b) proposed that if the pyrimidine ring adopted the opposite pucker (C5 below the plane, with C5-C11 bond equatorial, and the C6 bond to Cys 146 lying above the plane), then a reaction mechanism that agrees with the observed stereochemistry (Sleiker & Benkovic, 1984), could be readily formulated (see Fig. 1.8). The pyrimidine ring of dihydrouridine, a nucleoside which has a similarly substituted pyrimidine ring to that of deoxyuridine (the unphosphorylated precursor of dUMP), can adopt both the above forms of puckering with equal energies (Suck *et al.*, 1972). This premis lead Matthews and co-workers (1990b) to suggest that the uracil ring of dUMP may also be capable of adopting either of the puckering conformations.

Fig. 1.9. **Catalytic mechanism of TS.**
The proposed structure for the TS.dUMP.$CH_2FH_4$ complex (intermediate II in Fig. 1.7) which is consistent with the dUMP conformation seen in the crystal structure of the TS.dUMP.CB3717 complex. Rotation about the N1-ribose bond must occur prior to hydride transfer (see text for details; taken from Finer-Moore *et al.*, 1990).

Molecular modelling has shown that a pyrimidine ring of opposite pucker to that observed in the TS.FdUMP.$N^5$,$N^{10}$-$CH_2FH_4$ ternary complex could readily be accommodated at the TS active site, with only minor perturbations in the relative juxtapositions of the folate and nucleotide to each other, and the overall relationship to the protein (Matthews *et al.*, 1990b). Such an intermediate could undergo anti-periplanar elimination across C5-C11 to give an exocyclic methylene, having the configuration required to yield a methyl group with the correct chirality, following hydride transfer from $N^5$,$N^{10}$-$CH_2FH_4$.

1.6.5.2 **Model of Stroud and co-workers**

Workers from Robert Stroud's laboratory have recently determined the structure of the *E.coli* TS with bound phosphate and the TS.dUMP.CB3717 complex (Perry *et al.*, 1990; Montfort *et al.*, 1990). They have shown that dUMP is bound in the anti-conformation, as is also seen in the TS.FdUMP.$CH_2FH_4$ and TS.FdUMP.CB3717 structures of Matthews and co-workers (1990a,b). Workers in Stroud's group have used molecular modelling techniques (Singh *et al.*, 1986; Weiner *et al.*, 1984, 1986) to generate a description of the catalytic intermediate which is consistent with the stereo-specific reaction mechanism. They described several alternatives for the catalytic mechanism, but their favoured model is described here. The Cys146 sulphydryl group and methylene bridge are covalently bound in a trans-diaxial manner to C6 and C5 of dUMP respectively. The para-aminobenzoic acid ring of $N^5$,$N^{10}$-$CH_2FH_4$ and the phosphate moiety of dUMP are close to the positions of the same groups in the crystal structure of the TS.dUMP.CB3717 ternary complex and no major conformational changes are made to the protein (see Fig. 1.9; Finer-Moore *et al.*, 1990). dUMP is bound with its pyrimidine ring in the anti-conformation, with respect to its ribose moiety, and has S absolute configuration at C6. This leaves the pyrimidine ring of dUMP and the pteridine ring of the cofactor rotated only slightly from their positions observed in the crystal

structure.

Stroud and co-workers propose that syn-elimination occurs across the C5-C6 bond to form the exocyclic methylene derivative of dUMP. However, if this intermediate were then to undergo hydride **transfer**, it would result in the formation of dTMP with R absolute configuration for the methyl group. As mentioned previously, the R configuration is incompatible with the experimental data (Sleiker & Benkovic, 1984).

To overcome this, Stroud and co-workers have proposed that the pyrimidine ring of dUMP undergoes a 180° rotation about the N1-ribose ring bond, after the process of syn-elimination. This now leaves the C6-hydrogen of the co-factor orientated for hydride transfer to the *re* face of the exocyclic methylene group (see Fig. 1.9), which would result in the formation of a methyl group with the absolute configuration S, consistent with the experimental data (Sleiker & Benkovic, 1984).

1.6.5.3 **Summary**

These studies have confirmed the conformation of the bound ligands in the TS ternary complex first studied by $^{19}$FNMR (Byrd *et al.*, 1978). In addition, they have demonstrated that this complex does not represent the true catalytic intermediate.

The reaction mechanisms postulated by both groups, although based partly on crystal structures of different ternary complexes, share many common features, suggesting that the schemes are essentially correct. The major discrepancy is that workers from Stroud's group (Finer-Moore *et al.*, 1990) infer that reorientation of the pyrimidine ring is necessary for the proposed stereo-specificity of the TS catalysed reaction (Slieker & Benkovic, 1984), whereas Matthew's group (1990b) propose that alternate ring puckering of the pyrimidine ring occurs. Further work will be required to rationalize these alternate observations and reveal the precise nature of the reaction mechanism.

## 1.7 Structural features of TS

Prior to X-ray crystallographic studies of TS, Manavalan and co-workers (1986) used the techniques of circular dichroism (CD) and four different empirical methods to predict elements of secondary structure in TS. The CD spectrum was analysed for the various secondary structure elements and suggested 33% α-helix, 25% ß-strand, 20% turns and 16% other structures. Four algorithms which were used to predict the elements of secondary structure in TS (Burgess *et al.*, 1974; Chou & Fasman, 1974; Garnier *et al.*, 1978; Lim, 1974). The average values for the α-helical (32%) and ß-sheet (23%) content were comparable to the CD data. In addition, the value for α-helical content compares favourably with the observed value of 36.4% seen in the crystal structure of the native *L.casei* TS. However, their predicted locations in the primary sequence showed little relationship to the sites of helices in the native enzyme structure and the values for the ß-sheets show little correlation with the observed figure of 15.5% (Hardy *et al.*, 1987).

### 1.7.1 X-Ray crystallography

The understanding of many facets of TS function and dynamic behaviour during catalysis has been significantly enhanced by X-ray crystallographic studies. In the past four years, the 3-D structures of TS from various sources in several forms have been solved. These include the structures of the native unliganded enzymes of both *L.casei* (Hardy *et al.*, 1987) and *E.coli* (Perry *et al.*, 1990; Matthews *et al.*, 1990b). These structures reveal that TS, although not bound to any ligands, does have a phosphate ion bound at the dUMP phosphate site. The structure of the ternary complex of the *E.coli* enzyme with different combinations of substrate and cofactor analogues has also been reported (Montfort *et al.*, 1990; Matthews *et al.*, 1990a,b). Human TS has also been crystallized (Schiffer *et al.*, 1991), the first example of TS crystallization from phosphate and/or substrate free medium, thus representing a

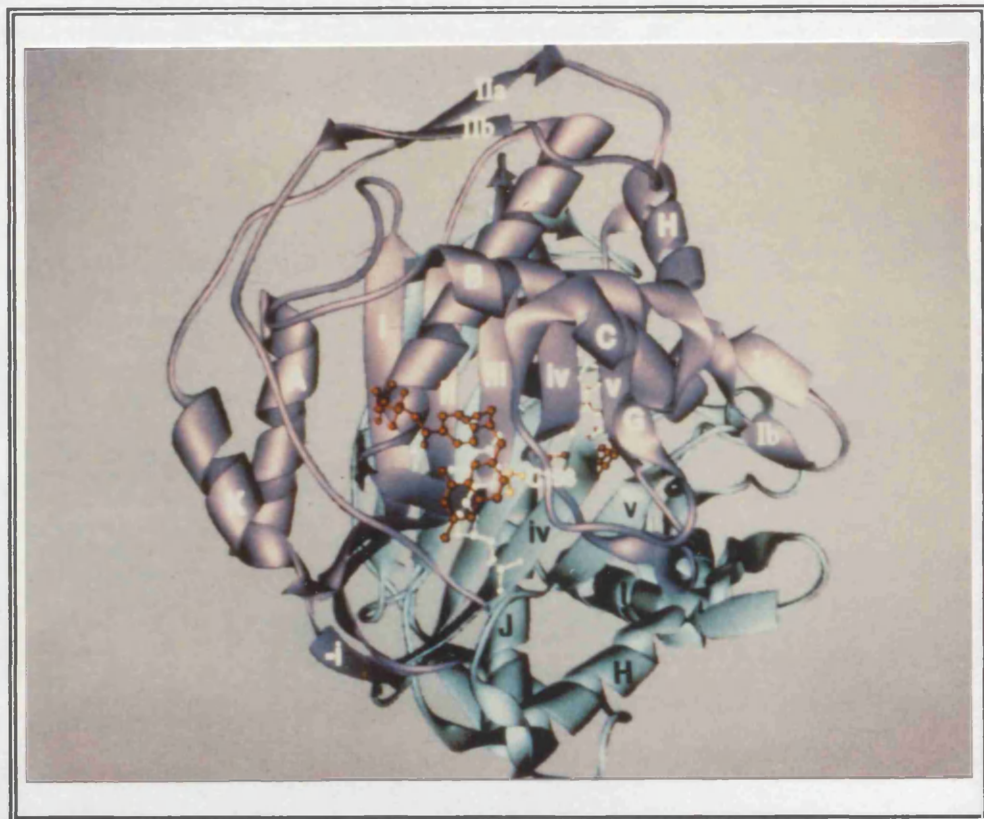



Fig. 1.10. ***E.coli* TS**.
The complete TS.dUMP.CB3717 ternary complex of *E.coli* TS is shown here, with one monomer in purple and the second coloured in blue. The five β-strands of each monomer are number i through v, and the α-helices from A through C, and G through K. Note that helix designations are from the *L.casei* TS and that helices D, E and F are absent in many other TSs. The active site residue of *E.coli* TS, C198 (C183 in VZV TS), is located at the N-terminal end of β-strand iv and is coloured in yellow. dUMP is shown in white, with the C-5 position juxtaposed to C198. The cofactor analogue, CB3717, is coloured in red, with the quinazoline ring lying on top of the pyrimidine ring of dUMP. The ligands are illustrated for both active sites (taken from Finer-Moore *et al*., 1990).

true native enzyme. Indeed, the addition of phosphate to the suspension cracks the crystals, suggesting that a possible phosphate induced conformational change takes place (Schiffer *et al.*, 1991). The crystal structures of human TS and bacteriophage T4 TS have been solved (Robert Stroud, personal communication).

### 1.7.2 General structural features

Comparison of the *L.casei* and *E.coli* primary structures reveals only one major difference, namely a 50 amino acid insertion around the middle of the *L.casei* enzyme (Hardy *et al.*, 1987; see Fig. 1.6). Comparison of the native enzyme structures reveals that, apart from the aforementioned insertion, the structures are very similar. Use of a difference distance matrix has identified a common core of 95 α-carbon atoms, from a total of 528 common to TS of both species, per dimer, as bearing a constant relationship of ±0.5 Ångstöms (Å) to each other (Perry *et al.*, 1990). A difference of this magnitude is comparable to the difference observed between two crystal structures of the same protein solved in different space groups (Chothia & Lesk, 1986).

The insertion in *L.casei* TS represents a small domain of 74 amino acids arranged as three short α-helices. The larger domain, common to all species of TS, is comprised of three layers of secondary structure. The first layer has five ß-strands that form the dimer interface. The central layer is made up exclusively of helices, a feature not normally seen in α/ß domains (Richardson, 1981), whilst the third layer consists of several other helices and the carboxy-terminal region (see Fig. 1.10).

### 1.7.3 Specific features of the large domain

#### 1.7.3.1 ß-kink of strands i, ii and iii

Hardy and co-workers (1987) described an aligned discontinuity in ß-strands i through iii of the *L.casei* native enzyme at residues G33, A258 and Q217 twisting the bottom left corner of the ß-sheet outwards. Matthews and

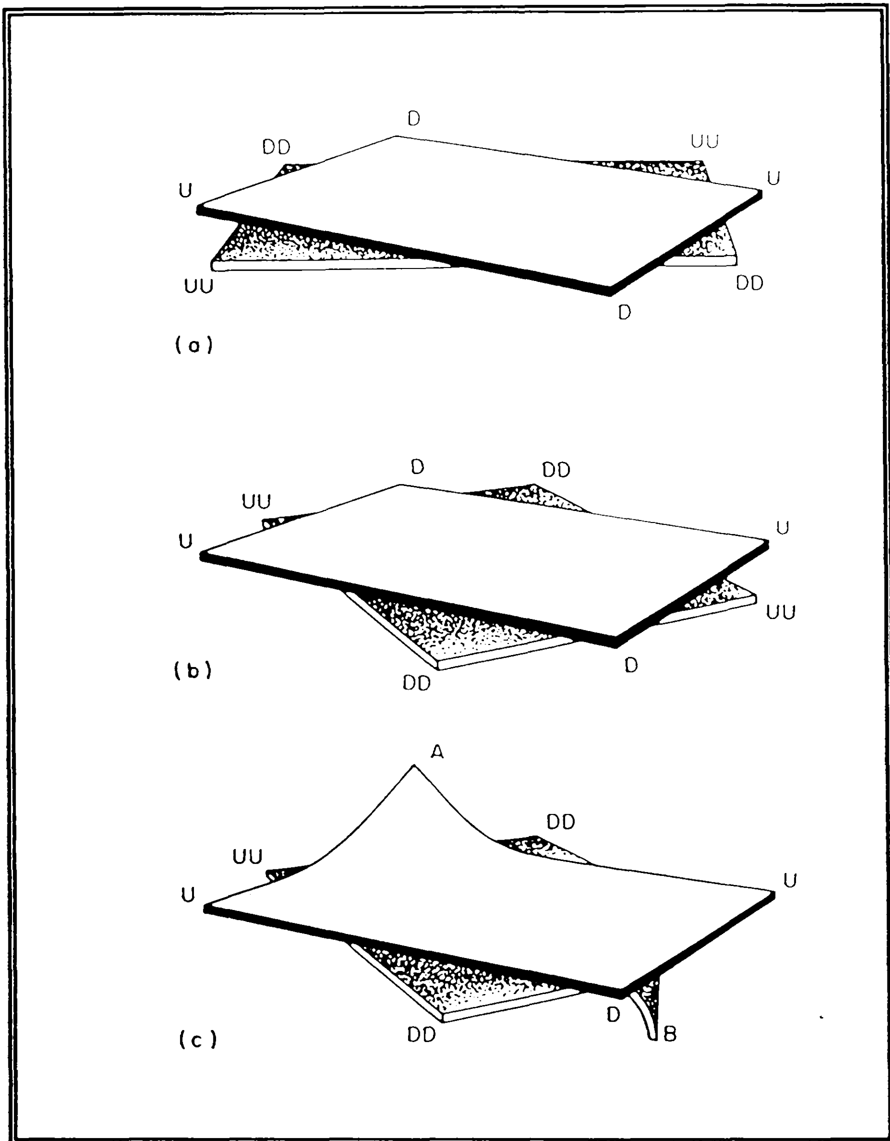



Fig. 1.11. **β-interface of TS.**
Face to face packing geometry of two aligned β-sheets. For the top sheet, U and D indicate up-pointing and down-pointing regions respectively. Corresponding portions of the bottom sheet are labelled UU and DD. (a) represents the anti-clockwise rotation of the bottom sheet relative to the top. (b) represents the clockwise rotation of the bottom sheet relative to the top, causing D to closely approach UU which would produce an unfavourable cavity (see text for details). (c) shows the stacked β-kinks in TS which "turn" the corner A of the top sheet upward and the corner B of the bottom sheet downward, thereby facilitating favourable packing of side chains across the entire dimer interface (adapted from Matthews *et al.*, 1989a).

co-workers (1989) looked at the corresponding structure from the ternary complex of *E.coli*, and noted that the discontinuity corresponds to the amino acids G31, G204 and R166. These residues correspond to G33, G256 and R218 in the *L.casei* TS. However, it is immediately apparent that these residues are not those involved in the discontinuity of the ß-strands in native *L.casei* TS. Assuming that both structures are correct, and that the *L.casei* amino acids involved adopt the same secondary structural element as those corresponding to the *E.coli* residues in the ternary complex, then this suggests that significant movement of ß-strand ii relative to ß-strands i and iii occurs upon ternary complex formation (see Sections 1.6.3 and 1.8).

1.7.3.2 **Dimer-interface**

The dimer interface of TS comprises two ß-sheets of five or six strands that stack face to face, and are related by a unique right-handed twist (Hardy *et al.*, 1987; Matthews *et al.*, 1989). This is in marked contrast to other proteins where the packing of ß-sheet interfaces has previously been observed only with a left-handed twist (Chothia *et al.*, 1977; Cohen *et al.*, 1981; Chothia & Janin, 1981). These studies suggest that the commonly observed, left-handed rotation in two such aligned sheets, is the result of the natural twist between individual strands in a single sheet. This left-handed rotation creates a deformed rectangular surface such that diagonally opposed corners both point either up or down. When brought together, two such surfaces would adopt a left-handed rotation relative to each other. This conformation is favoured by the down-pointing corners of the top sheet folding over the edge of the lower sheet, and the up-pointing corners of the lower sheet folding over the edge of the top sheet. A right-handed rotation would force interior regions of the sheet to separate, creating an unfavourable cavity in the interface (Chothia *et al.*, 1977; Cohen *et al.*, 1981; see Fig. 1.11).

As the two ß-sheets of TS are related by a right-handed

twist, such a cavity could be expected to exist between the sheets. However, at the point where a down-pointing corner of the top sheet would be predicted to interfere with an up-pointing corner of the bottom face, **the ß-kink**, which is comprised of the stacked ß-bulges, acts to reorientate a portion of the bottom sheet so that it bends in the same direction as the top sheet. This favours a close association between the top and bottom sheets,and also maintains optimal side-chain packing in the central region of the interface. The two-fold symmetry of TS means that a similar conformation occurs on the top sheet (Matthews *et al.*, 1989). This data suggests that efficient side-chain packing of residues from opposing ß-sheets can occur at the centre of a ß-sheet interface, either as a result of left-handed or right-handed rotation between the sheets, provided that in the latter case, the unfavourable edge interactions can be relieved by bulging (Matthews *et al.*, 1989).

1.7.3.3 **The J-helix**

A core structural element around which the large domain is folded, is the J-helix. This six-turn 3-10 helix is unusually hydrophobic, more so than some membrane spanning helices, and is almost completely buried from solvent. It sits on top of the ß-sheet, is flanked by the amphipathic helices A and H, and is capped by helix B (Hardy *et al.*, 1987).

1.7.3.4 **Active site cavity**

The active site is a shallow cavity in the large domain. The left hand wall is formed by residues in the ß-kink of strands ii and iii, and the loop connecting helix A and ß-strand i. The ceiling is composed of the solvent exposed accessible residues of the N-terminus of the J-helix, and the right hand wall by the N-terminus of strand iv, including the catalytic cysteine residue. The rear of the cavity is comprised of four residues from the ß-sheet of the second subunit (Hardy *et al.*, 1987). Of the 25

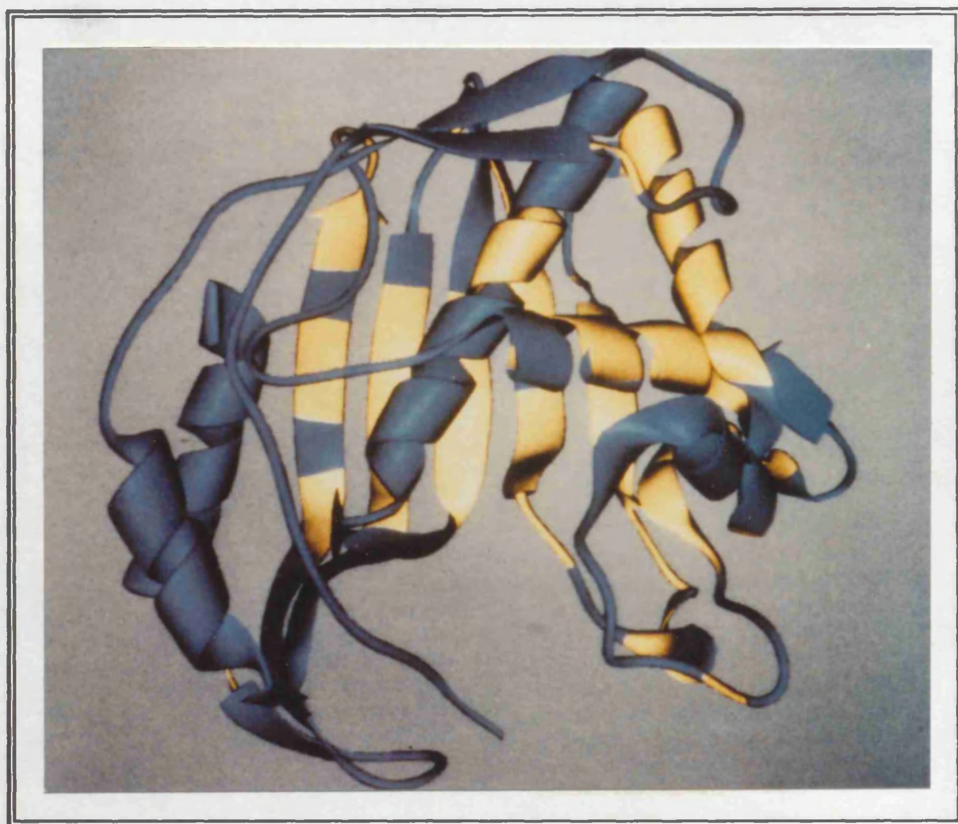


Fig. 1.12. **TS monomer**.
A single monomer of *E.coli* TS, presented in the same orientation as in Fig. 1.10, is shown to illustrate the residues involved in the conformational changes associated with ligand binding. Residues shown in yellow deviate from their original position by a RMS value of less than 0.5Å upon ligand binding, whereas residues shown in blue undergo substantial movement (see text for details). (Taken from Finer-Moore *et al.*, 1990).

residues that line the active site cavity, 12 are completely conserved in the 21 TS sequences reported to date. These include R163' and R164 (VZV numbering; amino acids from the second subunit are distinguished by a prime (') following the residue number of on monomer).

## 1.8 Conformational changes during ligand binding: segmental accomodation

In an attempt to understand the conformational changes that occur during the binding of substrate and co-factor, two groups have analysed the differences between the structures of the unliganded and ligand bound forms of *E.coli* TS (Matthews *et al.*, 1990a,b; Monfort *et al.*, 1990; Finer-Moore *et al.*, 1990; Perry *et al.*, 1990). To quantify the movement, a core of 155 alpha-carbon ($C_\alpha$) atoms whose mutual disposition is unchanged, +/- 0.5 Ångströms, between the liganded and unliganded forms, was identified (Monfort *et al.*, 1990). Most non-core residues had shifted between 0.5 to 1.25 Ångströms, the largest shift being associated with the C-terminus (see Fig. 1.12). In contrast to the domain shifts, seen in proteins such as hexokinase (Anderson *et al.*, 1978) or citrate synthase (Remington *et al.*, 1982), the changes seen in TS are due to segmental accommodation, where α-helices, ß-strands and connecting loops undergo concerted movement towards the active centre (Matthews *et al.*, 1990b; Monfort *et al.*, 1990).

The central J-helix pivots around a central point such that the N-terminal end moves ~0.5 Ångströms towards the active centre, causing the region of ß-strands i, ii and iii below the ß-kink to move towards the same location (see Section 1.6.5). Indeed, the ß-kink acts as one of several hinge points in the structure, about which other segments of secondary structure move. The region containing the active site residue, C146 (C183 in VZV numbering), moves inward, whilst the loop containing the two arginine residues 126' (164'; primed numbers indicate residues from the other subunit of the dimer) and 127' (163') also moves

inwards. The two largest shifts involve the loop containing R21 (38) and T22 (39) that moves up to interact with the C-terminus, and the movement of the four C-terminal residues, P261 (E298) to I264 (L301). In the unliganded structure, these four residues point out into the solvent, suggesting they are disordered (Matthews *et al.*, 1990a). However, on the binding of ligands, this region moves ~4.0 Ångströms, sequestering several water molecules, to anchor the co-factor in place. This directs the exact alignment of substrate and co-factor to allow the formation of the observed H-bonds between the carboxylate group of I264 and residues R21, T22 and W83 (via a water molecule), and between R21, A263, a water molecule and the cofactor (see Fig. 1.10; Matthews *et al.*, 1990a,b; Monfort *et al.*, 1990; Finer-Moore *et al.*, 1990).

## 1.9 Structural plasticity and covariant accomodation

The introduction of amino acid substitutions by site-directed mutagenesis affects protein function in an, as yet, poorly understood manner. The variations are often tolerated by adjustment of neighbouring residues and can result in unexpected properties. The accomodation of substitutions is due to the structural plasticity of the protein, the ability of a protein to adapt to changes in its primary structure, whilst maintaining the integrity of its tertiary fold. To quantitatively describe the capacity for structural plasticity, in response to substitutions, two approaches could be envisaged. The first involves the comparison of the structures of many site-directed variants of a single protein. The second would be a comparison of two structures that differ at many positions, with an attempt to generalize the effects of substitutions.

A detailed comparison of the structures of unliganded TS from *L.casei* and *E.coli* has shown that they are 89% conserved in terms of atom sites (sharing 3824 common atoms out of 4300 from *E.coli*). This comparison has provided an ideal opportunity to investigate the plastic adaptations

seen in proteins by the second method (Perry *et al.*, 1990).

A detailed study of the buried region of TS from both structures revealed three different mechanisms by which substitutions could be accomodated (Perry *et al.*, 1990). The first is the occurrence of local changes, whereby large effects observed close to the site of mutation dissipate as a function of distance. This is usually achieved by a mechanism that involves local rearrangements, often the repacking of adjacent residues, with only minimal effeots on the overall architecture of the protein (Perry *et al.*, 1990).

A second mechanism is covariant accomodation where multiple sequence changes combine to minimize the distortion of the main chain atoms and simultaneously preserve the volume of the packed side-chains. A representative example of this is found in the ß-sheet, structurally the most conserved region of the TS protein. F255 at the centre of ß-strand ii of *L.casei* is substituted by glycine in *E.coli* (see Fig. 1.6). The space 'vacated' by the glycine residue is 'filled' by the side-chains of W253 and T258 (*L.casei* numbering), whereas the residues H253 and A258 surround F255 in *L.casei*. This corresponds to a net difference of only a single atom in these three residues (Perry *et al.*, 1990). An example of a covariant pair across the dimer interface is seen where T200 and L201 of *L.casei* TS are replaced by A200 and F201 (*L.casei* numbering) in *E.coli* TS, thus preserving a similar side chain volume (Perry *et al.*, 1990). In both examples, the positions of the main chain $C_\alpha$ atoms are essentially unchanged. In general, it has been found that when a substitution introduces three or more heavy atoms (carbon, oxygen, sulphur or nitrogen), this nearly always results in a second site, or covariant change nearby (Perry *et al.*, 1990).

These observations correlate with experimental data obtained from mutagenesis studies of the hydrophobic core of the bacteriophage lambda repressor protein, where a net change of >+2 or <-3 methylene groups has been shown to

inactivate the protein (Lim & Sauer, 1989). Moreover, where a mutation increases the volume of the core in excess of the permitted levels, second site mutations that reduce the net volume, can restore activity (Lim & Sauer, 1991).

A final mechanism by which plastic adaptation can occur is by the concerted movement of individual elements of secondary structure. Several of the helical regions of TS, which contain the largest structural differences between the enzyme from *L.casei* and *E.coli*, have been accomodated in this fashion. The 'ensemble' movement of helices B, C and G is the most prominent example of this process, where these structures have shifted together as a unit in the *L.casei* structure, relative to *E.coli*. This movement has occurred to accomodate the interactions between the 3-helix unit and the helix D of *L.casei*, a helix that is absent from *E.coli* TS (Perry *et al.*, 1990). It is interesting to note the preservation of the side-chain packing interactions between these three helices during the course of evolution.

## 1.10 Models of VZV TS

Two models of the VZV TS, one of the ternary complex and the other of the native enzyme with bound phosphate, have been constructed by Celia Schiffer and Robert Stroud (personal communication). The models are based upon the coordinates of the ternary complex of *E.coli* TS with CB3717 and dUMP (Montfort *et al.*, 1990) and the coordinates of the *E.coli* TS structure with bound phosphate respectively (Perry *et al.*, 1990). All residues from the *E.coli* sequence that differed from the equivalent residue of the VZV sequence were substituted by the latter using the graphics program INSIGHT (Dayringer *et al.*, 1986). The substituted residues were positioned in a conformation which overlapped as much as possible with that of the residue existing in the *E.coli* structure. Where a VZV residue was larger than the one it replaced, it was placed in a conformation which corresponds to that most frequently

**Bacteriophage T4** (LaPat-Polasko *et al.*, 1990; Frasca *et al.*, 1988).

| Mutant | TS activity | Mutant | TS activity |
|---|---|---|---|
| K48R | 6.5% | P155A | 87.0% |
| K48M | <1.0% | C156S | 0.07% |
| K49R | 56.0% | C156R | NDA |
| R137G | <1.0% | H157V | 23.0% |
| R137K | 32.0% | | |
| R176K | <1.0% | | |

K48 (K65); K49 (K66); P155 (P182); C156 (C183); H157 (H184).

***E.coli*** (Dev *et al.*, 1988, 1989).

| Mutant | TS activity | Mutant | TS activity |
|---|---|---|---|
| C146S | 0.02% | H147V/L/Q/N/G | 6.5-10.4% |
| C146G | NDA | H147E/D | 0.047-0.052% |
| C146T | NDA | H147R/K | 0.007-0.008% |

C146 (C183); H147 (H184).

***L.casei*** (Climie & Santi, 1990; Climie *et al.*, 1990)

| Mutant | TS activity | Mutant | TS activity |
|---|---|---|---|
| R179A | 31% | H199L/K/R/M | non-complementing |
| R179E | 24% | H199/N/W/Y/E | non-complementing |
| R179T | 24% | H199P/S/G | complementing |
| R179K | 26% | H199A/T/V | complementing |

R179 (R164); H199 (184).

**Mouse** (Zhang *et al.*, 1990).

| Mutant | TS activity |
|---|---|
| R44V | <1.0% |

R44 (R38).

Table 1.3. **Site-directed mutagenesis of TS.**
Summary of all the site-directed mutagenesis data of TS from different sources. TS activity was measured by the tritium release assay, except for the H199 mutants of *L.casei* which were assayed for complementation. NDA indicates no detectable activity. / indicates alternative substitutions. Numbers in parentheses refer to equivalent amino acid from VZV TS. The convention K48R indicates that the wild type amino acid lysine at position 48 has been replaced by arginine.

found for that amino acid in protein structures. The model lacks the first 19 amino acids from the VZV sequence, and the two short insertion sequences common to all eukaryotic TSs (residues 105-116 and 134-141; VZV numbering), that are not found in the *E.coli* enzyme. At the time the model was generated, these regions of the human TS being studied by this group (Schiffer *et al.*, 1991), were unresolved. The model of the VZV native TS was refined by 200 cycles of conjugate gradient minimization with the refinement package of XPLOR (Brünger *et al.*, 1987), clearing up the bad contacts generated by the amino acid substitutions. The minimization resulted in a root mean square (RMS) shift of 0.38 Ångströms of the alpha-carbons ($C\alpha$). The model of the VZV ternary complex was refined with 150 cycles of XPLOR. Due to difficulties in the initial stages of energy minimization, the $C\alpha$ atoms were constrained to stay fixed, thus allowing only the movement of side chains to remove any energetically unfavourable contacts.

## 1.11 Mutational analysis of TS

In recent years, several groups have studied the role that individual amino acids play in the structure and function of TS from various sources. In particular, two groups have used different methods of saturation site-directed mutagenesis on TS from *L.casei* and *E.coli*. Some of their findings are summarised in this section.

### 1.11.1 Site-directed mutagenesis

Several groups have analysed a small subset of residues, thought to play key roles in the catalytic mechanism of TS, by site-directed mutagenesis. The results are summarised in Table 1.3.

These studies have served to highlight several interesting features of some of the residues predicted to play important roles in catalysis and substrate recognition (Finer-Moore *et al.*, 1990; Matthews *et al.*, 1990b). The substitution of the totally conserved active site cysteine

in *E.coli* TS with Gly or Thr inactivated the enzyme. Substitution with Ser yielded an enzyme with very low but detectable activity. The serine hydroxyl group to some extent replacing the thiol of cysteine. This confirms the crucial role of this amino acid in catalysis (Dev *et al.*, 1988). More interestingly, several other highly conserved residues thought to play essential roles in activity were found to be dispensable, in that certain substitutions had little effect on enzyme activity. When proline 155 of bacteriopage T4 (P182 in VZV), which is totally conserved in all 21 TS sequences reported, was replaced by alanine, only a small drop in $k_{cat}$ was observed compared with that of the WT enzyme (LaPat-Polasko *et al.*, 1990).

The mutability of three of the four highly conserved arginine residues thought to play a role in the co-ordination of the phosphate group of dUMP was also studied. Replacement of R44 (R38) by valine, in mouse TS, resulted in a greater than 100-fold decrease in enzyme activity (Zhang *et al.*, 1990). Substitution of lysine for the invariant R137 (R164) in T4 or R179 (R164) in *L.casei* only reduced $k_{cat}$ by 3- to 4-fold (LaPat-Polasko *et al.*, 1990; Santi *et al.*, 1990). Several other amino acid substitutions, including a glutamine, were tolerated at position 179 in *L.casei* TS (Santi *et al.*, 1990). Substitution of arginine by glycine inactivated T4 TS (LaPat-Polasko *et al.*, 1990). In contrast to R137 of T4 TS, the completely conserved R176 (R203) cannot be replaced by lysine (LaPat-Polasko *et al.*, 1990).

Analysis of the folate-binding region of T4 TS revealed that substitution of arginine for the highly conserved residue K48 (K65), significantly diminished TS activity, whilst replacement of the neighbouring highly conserved amino acid K49 (K66) decreased activity by less than 2-fold (LaPat-Polasko *et al.*, 1990).

The role of the histidine residue, adjacent to the active site cysteine, has been studied by several groups. This histidine is conserved in all TS sequences except bacteriophage phi3T (replaced by valine; Kenny *et al.*,

| VZV | *L.casei* | | *E.coli* | | Phage T4 | | Mouse | |
|---|---|---|---|---|---|---|---|---|
| Original amino acid | Original amino acid | Substituted amino acids | Original amino acid | Substituted amino acids | Original amino acid | Substituted amino acids | Original amino acid | Substituted amino acids |
| R38 | R23 | LHV | R21 | LH | | | R44 | V |
| R163 | R178 | SPLFEK | R126 | SPLFEK | | | | |
| R164 | R179 | AEK | R127 | GAEK | R137 | G | | |
| P182 | P197 | A | | | P155 | A | | |
| C183 | C198 | SGAPLFYER | C146 | SGAPLFYER | C156 | S | | |
| H184 | H199 | GASPLYERKV | H147 | GASPLYERK | H157 | V | | |
| R203 | R218 | GHKAYFLSHEK | R166 | GHKAYFLS | R176 | K | | |
| N214 | N229 | ACQLY | N177 | ACQLY | | | | |

Table 1.4. **Random mutagenesis of TS**.
This table displays all the random mutagenesis data reported to date where applicable to TS from more than one source. Amino acids shown in **bold** complement the growth of *E.coli thyA* strains. Amino acids shown in outline fail to complement the growth of *E.coli thyA* strains. Note that amino acid on the same line occupy equivalent positions in the various TS species.

1985) and *L.lactis* (replaced by alanine; Ross *et al.*, 1990). Substitution of histidine by valine in TS from *L.casei* (H199), *E.coli* (H147) and bacteriophage T4 (H157) produced active enzyme (Climie & Santi, 1990; Dev *et al.*, 1989; LaPat-Polasko *et al.*, 1990). Interestingly, substitution of alanine in place of histidine 199 in the *L.casei* TS also produced active enzyme (Climie & Santi, 1990).

In summary, these initial experiments have shown that whilst some invariant residues cannot tolerate amino acid substitutions, others can be mutated freely with little deleterious effect on activity. Moreover, the effect of the substitution of one amino acid for another may, or may not, have an analogous effect on TS activity when the same substitution is made to TS from an alternate source.

### 1.11.2 Saturation site-directed mutagenesis

Two groups have recently described different methods for the large scale random mutagenesis of targeted residues of TS from *L.casei* and *E.coli* (Climie *et al.*, 1990; Michaels *et al.*, 1990). Some of the results from these experiments are shown in Table 1.4.

Workers in Dan Santi's group constructed a synthetic *L.casei* TS gene that possessed over 30 unique restriction sites in the TS coding sequence and was shown to express TS to a level of 10% of the total soluble protein in *E.coli* (Climie & Santi, 1990). A technique of mutagenesis (Wells *et al.*, 1985), using synthetic DNA cassettes containing a mixture of 32 possible codons (represented by NNG/C) that encode all 20 amino acids and the amber stop codon (TAG), in conjunction with a growth complementation assay was developed. Climie and Santi (1990) demonstrated the feasibility of this method for creating the majority of the single amino acid substitutions of a targeted residue. Using this technique, they went on to identify 125 mutants, which they grouped into three different phenotypic classes (active, sick and dead/inactive) at 12 different amino acid positions (Climie *et al.*, 1990). These residues had all

previously been predicted to play roles in the direct binding of, or lie in close proximity to dUMP (Hardy *et al.*, 1987). Eight of the 12 residues are invariant, whilst the other four are conserved in 15 to 19 of the 21 TS sequences reported to date.

All the sequenced substitutions of R23, C198, and R218 resulted in enzyme inactivation, whilst a fourth residue, Y261, only retained activity when substituted by methionine. The remaining eight residues, five of which are invariant, could be readily substituted by at least two, and in some cases as many as twelve, different amino acids without inactivating the enzyme. Although none of the subtitutions of R23 retained activity, the mutant R23G, which would mimic the only naturally occurring substitution mutant of this residue was not isolated. To assess the level of activity necessary to facilitate complementation, the $k_{cat}$ value of several mutants was determined. Mutants with greater than 1% activity gave full complementation, although some had levels comparable to WT *L.casei* TS, whereas mutants with less than 0.1% activity could not complement the growth of an *E.coli thyA* strain (Climie *et al.*, 1990).

Workers from Jeffrey Miller's group, in collaboration with David Matthews, adopted a different approach for the saturation mutagenesis of *E.coli* TS (Michaels *et al.*, 1990). They replaced a normal codon with the amber nonsense stop codon (TAG) at 20 different sites in the *E.coli* TS gene on a low-copy number plasmid. By transforming these 20 *thyA* plasmids into 13 different *E.coli* strains, each of which harboured a different amber suppressor tRNA, nearly 250 mutants were generated. These were analysed by growth complementation and enzyme assays were performed on a selected set of mutants (Michaels *et al.*, 1990). This technique has been used to generate over 1,600 amino acid substitutions of the *E.coli lac* repressor (Miller *et al.*, 1979; Kleina & Miller, 1990).

Seven of the 20 residues studied were totally conserved, eight were highly conserved (invariant in >15 TS sequences)

and the remaining five were non-conserved (residue found in <7 TS sequences). This latter group (E14, R35, D105, N121, E223) complemented the growth of all 13 strains showing that all the amino acid substitutions gave at least some level of activity (Michaels *et al.*, 1990). The highly conserved residues show varied tolerance to substitution. Some could only tolerate conserved changes, such as F30 (L or Y) or G204 (A or S), whilst others such as Q33 could be replaced by any of the amino acids provided by the suppressor strains (Michaels *et al.*, 1990). Surprisingly, even the invariant residues could tolerate substitutions. C146S retained some activity, in agreement with a previous report (Dev *et al.*, 1988), as did D169C, although no other substitutions of these residues facilitated growth complementation. Growth complementaton was seen in the remaining conserved residues by substitutions of two to seven residues (Michaels *et al.*, 1990). Growth complementation was shown to correlate with an enzyme activity of less than 1.0%, whereas mutants that did not complement growth were shown to have specific activities of <0.32% the wild type level (Michaels *et al.*, 1990).

The efficiency of suppression can vary from 5% to 100% depending upon the sequence context of the amber codon (Kleina *et al.*, 1990). To investigate the effects of varying levels of suppression a comparison of the phenotypes of four suppression mutants and four missense mutants was made (Michaels *et al.*, 1990). Two of the missense mutants had the same phenotype as the corresponding suppressor mutant (H147A and D169H). However, the other two missense mutants complemented the growth of *E.coli thyA*, whilst their corresponding suppressor mutants did not (H147E and D169E). On this basis, it was proposed that overproduction of mutant forms of TS can sometimes lead to growth complementation (Michaels *et al.*, 1990). The growth complementation ability of several of the suppressor mutants from this study (Michaels *et al.*, 1990) can also be compared to the enzyme activities of the purified missense mutant enzymes of *E.coli* TS generated by Dev and co-workers

(1989). H147G complements the growth of *E.coli thyA* and has an enzyme activity 10% of the wt *E.coli* TS, whilst H147K has only 0.008% the activity of the wt enzyme and fails to complement growth. However, the mutant H147R, has only 0.007% activity of the wild type TS, almost identical to that of H147K, but can successfully complement the growth of *E.coli thyA*. Interestingly, the missense mutant H147E, which has 0.05% the activity of the wild type enzyme can complement the growth of *E.coli thyA*, although its corresponding suppressor mutant cannot (Dev *et al.*, 1989; Michaels *et al.*, 1990). This comparison shows that complementation of *E.coli thyA* strains does not always correlate with enzyme activity.

Where the same substitution has been made to equivalent residues in TS from both *E.coli* and *L.casei*, the ability to complement the growth of *E.coli thyA* strains is not always the same. As seen in Table 1.4, of 45 such mutants, 25 (56%) are inactive in both enzymes, nine (20%) are active in both, nine (20%) are active in *E.coli* TS but not *L.casei* TS, and the remaining two mutants (4%) are active in *L.casei* TS but not *E.coli* TS. In a comparison of the limited data available for equivalent residues from phage T4, of five mutants, one is active in all three TS species, two are active in T4 and *L.casei*, and one each are inactive in T4 but active in *E.coli* or *L.casei* (LaPat-Polasko *et al.*, 1990). The only mutant reported for the murine TS (Zhang *et al.*, 1990) has similar levels of enzyme activity to the equivalent mutant in TS of *L.casei*.

When the results of these groups are compared, they show that many of the invariant residues of TS studied, may be substituted by at least one amino acid and still retain at least some catalytic activity. This poses the paradoxical question; why are so many of the residues of TS completely conserved in all the TS sequences reported to date, yet they can be substituted by other amino acids and still retain activity?

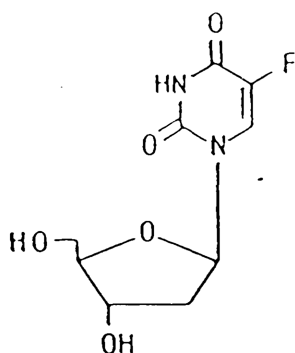
## 1.12 TS inhibitors

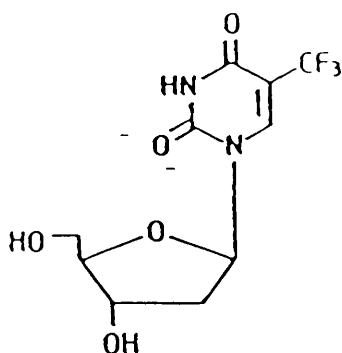
Since the discovery that 5-fluorouracil (FU) is a potent TS inhibitor by Heidelberger and colleagues in 1957, the growth in the number and diversity of TS inhibitors has been explosive. There are currently too many compounds to list individually, and their mechanisms of action are also too diverse to cover completely. However, a number of representative compounds, that fall broadly into seven different groups, will be discussed and, where known, their mechanisms of action described.

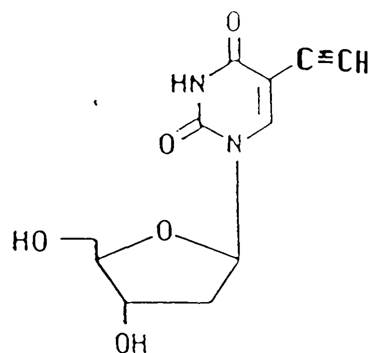
### 1.12.1 Nucleoside analogues requiring activation

All of the nucleoside analogues discussed in this section, predominantly 5-substituted-2'-deoxyuridine analogues, exert their inhibitory effects at the level of the corresponding 5'-monophosphate. For these analogues to be active *in vivo*, they must first be enzymically activated (see Sections 3.1 and 3.4). For studies of purified enzyme, the 5'-monophosphate derivative must be used. The necessity for the phosphate group was shown by the fact that deoxyuridine (UdR) is not a substrate for the purified enzyme (Reyes & Heidelberger, 1965), with an estimated $K_i$ of 0.2mM for human TS (Lockshin & Danenberg, 1981), which corresponds to a 500-fold drop in binding activity compared to dUMP. In addition, inorganic phosphate is a competitive inhibitor of nucleotide binding, with a $K_i$ of 10mM (Lockshin & Danenberg, 1981). In *in vivo* studies, the necessity for the 5'-monophosphate group has been demonstrated by the fact that mutations in the activation step lead to resistance to these analogues (see Section 1.14).

A study of the quantitative structure-activity relationships of different TS inhibitors of *L.casei* TS has shown that analogues with a 5-substituent that increases electron withdrawal from the heterocyclic ring of uracil, increase the affinity of the analogue for the enzyme, whilst bulky groups at the C-5 position are usually

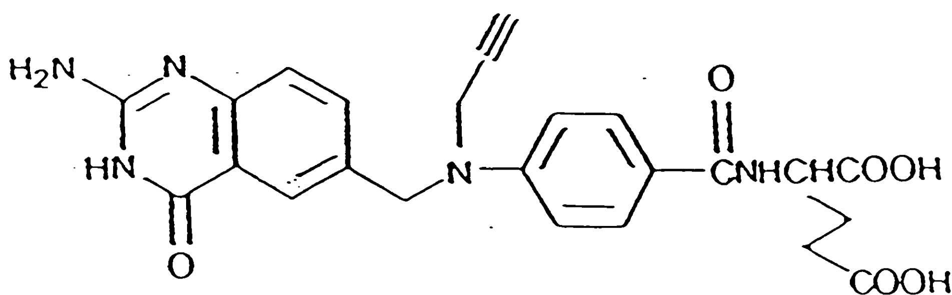



| **FUdR** | **$CF_3$UdR** | **EYUdR** |
|---|---|---|
| 5-fluoro-2'-deoxyuridine | 5-(trifluoro-methyl)-2'-deoxyuridine | 5-ethynyl-2'-deoxyuridine |

Fig. 1.13a. **TS inhibitors: 5-substituted deoxyuridine analogues.**




**CB3717** $N^{10}$-propargyl-5,8-dideazafolate

Fig. 1.13b. **TS inhibitors: folate analogues.**

detrimental to binding for steric reasons (Wayata & Santi, 1977). Consistent with these findings is the observation that the 5'-monophosphate derivatives of 5-fluoro-2'-deoxyuridine (FUdR), 5-(trifluoro-methyl)-2'-deoxyuridine ($CF_3$UdR; trifluorothymidine TFT), 5-nitro-2'-deoxyuridine ($NO_2$-UdR) and 5-ethynyl-2'-deoxyuridine (EYUdR) have consitently emerged as the most potent TS inhibitors in both *in vitro* and *in vivo* experiments with nucleoside analogues (see Fig. 1.13a; Balzarini *et al.*, 1982; Lewis & Dunlap, 1981).

The mode of action of these compounds is now reasonably well understood. FdUMP, the active form of FUdR, forms a covalent ternary complex with TS in the presence of the cofactor, $N^5,N^{10}$-$CH_2FH_4$, which is similar to the covalent catalytic intermediate (see Section 1.6; Lewis & Dunlap, 1981). This complex is stable and terminates the reaction because the fluorine atom cannot be abstracted, as occurs with the C-5 hydrogen in the course of the normal TS reaction, presumably due to the strength of the C-5-F bond (Santi & Danenberg, 1984).

In contrast to FdUMP, $CF_3$dUMP and $NO_2$dUMP can inhibit TS in the absence of $N^5,N^{10}$-$CH_2FH_4$, but still interact with the active site cysteine residue forming a 5,6-dihydro-pyrimidine derivative. Their ability to interact with TS in the absence of the cofactor is thought to be the basis of their tight-binding to, and inhibition of, TS (Danenberg & Heidelberger, 1974; Wayata *et al.*, 1980).

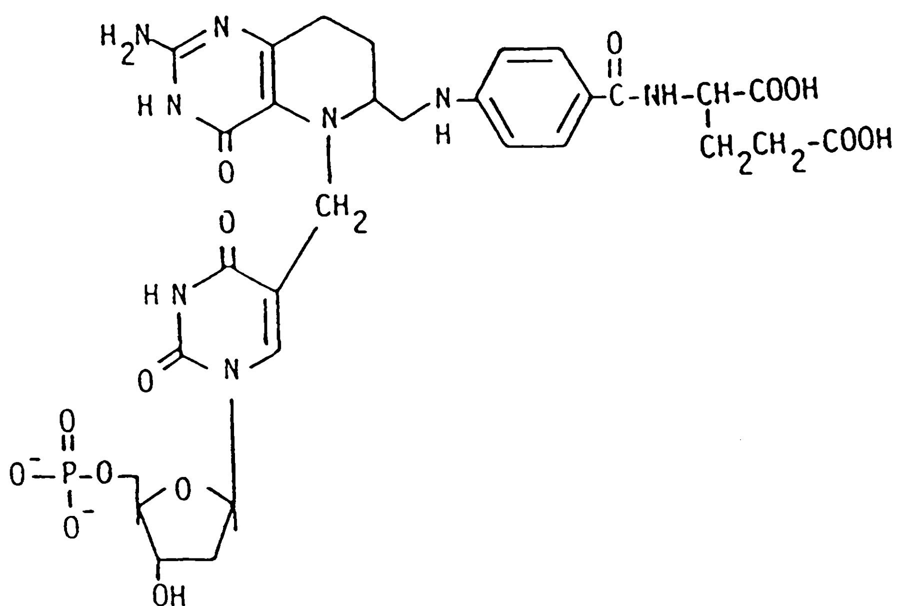
EYUdR is a potent inhibitor of S-49 murine lymphoma cells with an $IC_{50}$ value of 12µM, and inhibits TS only in the presence of the cofactor, by forming an inhibitory ternary complex (Barr *et al.*, 1981).

Danenberg and co-workers (1981) confirmed that a ternary complex between EYdUMP, $N^5,N^{10}$-$CH_2FH_4$ and TS was formed but it was much weaker than that formed by FdUMP. Furthermore, the EYdUMP-based complex was disrupted upon denaturation of the enzyme, in contrast to the FdUMP-based complex, which remains stable upon denaturation (Danenberg *et al.*, 1981).

### 1.12.2 **Cofactor analogues**

2-amino-4-hydroquinazolines (Bird *et al.*, 1970) and folylpolyglutamates (Kisliuk *et al.*, 1974) were the first folate analogues shown to inhibit TS. A major discovery that stimulated the rapid expansion of research into the use of anti-folates directed at TS, was the demonstration that pteroyl-hexaglutamate could inhibit bacteriophage T2 TS at a concentration 100-fold lower than that required to inhibit *E.coli* TS (Maley *et al.*, 1979c). This was a feature that none of the potent substrate based TS inhibitors had exhibited. Folate analogues have several other advantages over nucleoside analogues for chemotherapeutic use, including a decreased susceptiblity to metabolic degradation, and they pose no significant risk of mutagenicity or carcinogenicty (Keyomarsi & Moran, 1990). Moreover, the high levels of dUMP found in many tumours (that would reduce the effects of nucleotide analogues such as FdUMP) would potentiate binding of folate analogues (Danenberg & Lockshin, 1982; Myers *et al.*, 1974; Keyomari & Moran, 1988, 1990; Pogolotti *et al.*, 1986). Conversely, the existance of low levels of $N^5N^{10}$-$CH_2FH_4$ in many tumours (which reduces the cytotoxicity of FdUMP by preventing maximal formation of the TS covalent ternary complex) would result in little competition for exogenous anti-folate analogues binding to TS (Houghton *et al.*, 1981; Evans *et al.*, 1981; Jackson *et al.*, 1983; Sikora *et al.*, 1988).

Several other folate analogues have been shown to inhibit TS. One of the first, and subsequently most studied, compounds to emerge from these investigations was $N^{10}$-propargyl-5,8-dideazafolate (CB3717; see Fig. 1.13b) which has a $K_i$ value of $1 \times 10^{-9}$M for TS isolated from murine leukemia L1210 cells (Jones *et al.*, 1981). Phase I and II clinical trials have shown that CB3717 has demonstrable clinical activity against ovarian, breast and liver cancers, but the occurrence of renal toxicity in a sporadic, unpredictable and life-threatening manner, has precluded its further development as an anti-cancer drug




**N⁵dUMP-8-deazatetrahydrofolate**

Fig. 1.14. **TS inhibitors: Bisubstrate analogues.**

(Harrap *et al.*, 1989). However, workers in this group have used this compound as a parent compound to identify equally effective, but less toxic derivatives (Harrap *et al.*, 1989; Jones *et al.*, 1989; Hughes *et al.*, 1990), which have met with significant success (Jackman *et al.*, 1990; Marsham *et al.*, 1990).

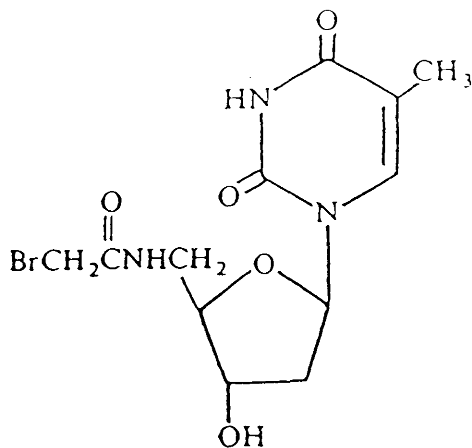
CB3717 inhibits TS from murine L1210 cells by the formation of a complex with dUMP, that still allows FdUMP to bind. This suggests that the covalent inhibitory complex formed is similar to the 1:1:1 ternary complex of TS.dUMP.$N^5N^{10}$-$CH_2FH_4$ (see Section 1.6; Keyomarsi & Moran, 1990). TS inhibition mediated by the binding of CB3717 also results in increased levels of the enzyme, presumably by protecting the enzyme against proteolysis (Keyomarsi & Moran, 1990).

In addition to its role in the treatment of disease, CB3717 has been used to probe the mechanism of action of TS by virtue of its use in structural studies of the *E.coli* TS (Matthews *et al.*, 1989, 1990a; Finer-Moore *et al.*, 1990; Montfort *et al.*, 1990).

Dihydrofolate, which accumulates at high levels when DHFR is inhibited, can directly inhibit TS (Bunni *et al.*, 1988; Seither *et al.*, 1991). Chu and co-workers (1990) point out that whilst the direct inhibition of TS due to a build up of $FH_2$ is indeed a major cause of growth inhibition, depletion of $FH_4$ also plays a role. This lead them to suggest that inhibition of *de novo* thymidylate synthesis is a multi-factorial process, a proposal backed by the comprehensive mathematical model of folate cycle kinetics (Morrison & Allegra, 1989).

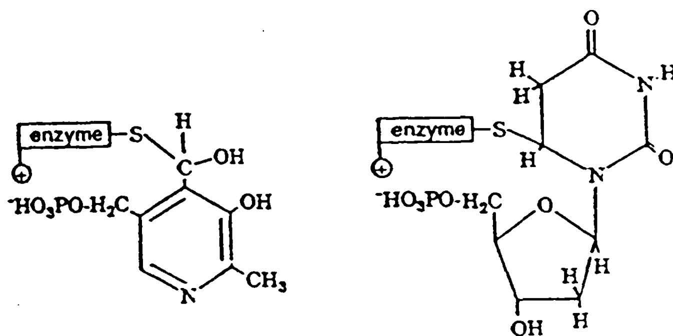
### 1.12.3 **Bisubstrate analogues**

Several co-called bisubstrate analogues, comprised of a folate molecule, covalently linked to a nucleotide analogue and designed to mimic the non-enzymic portion of the catalytic intermediate of the TS reaction, have been shown to inhibit TS (Park *et al.*, 1979; Amarnath *et al.*, 1984). Two such compounds, $N^5$-dUMP-$N^{10}$-propargyl-8-deazatetra-




**BAT**
5'-bromoacetamido-2'-deoxythymidine

Fig. 1.15a. **TS inhibitors: 5'-haloacetamido-2'-deoxythymidine analogues.**




**PLP** pyridoxal phosphate

Fig. 1.15b. **TS inhibitors: *in vivo* regulators.**
This figure illustrates the proposed interaction of PLP (right) with TS in comparison to the interaction between TS and thymidylate (taken from Chen *et al.*, 1989).

hydrofolate and $N^5$-dUMP-8-deazatetrahydrofolate (see Fig. 1.14) inhibit purified TS from murine leukemic L1210 cells in a competitive manner, with respect to dUMP, with a $K_i$ value of 0.09µM for both compounds (Veda *et al.*, 1986). This study has also shown that rather than entering into the active site of the enzyme to produce an inhibitory ternary complex, these compounds inhibit TS by blocking the nucleotide and folate binding sites. They should therefore, be referred to as 'substrate-product' rather than 'bi-substrate' analogues, as the folate analogue is already attached to the 5-position of dUMP.

### 1.12.4 Nucleoside analogues not requiring activation

A novel group of compounds, the 5'-haloacetamido-2'-deoxythymidines, do not require phosphorylation or the co-factor to inhibit TS (see Fig. 1.15a; Sani *et al.*, 1986). The three compounds 5'-bromo-, 5'-chloro- and 5'-iodo-acetamido-2'-deoxythymidines (BAT, CAT and IAT respectively) inhibit TS from L1210 cells with $IC_{50}$ values revealing their relative potencies as BAT (1.2µM) > IAT (6.0µM) > CAT (62.0µM) (Sani *et al.*, 1986). More detailed kinetic studies with BAT suggest that it acted as a competitive inhibitor with respect to dUMP. It had a $K_i$ value of 5.4µM and complete and irreversible inactivation can be achieved by pre-incubation of TS in the presence of 50µM BAT for 30min (Sani *et al.*, 1986). Replacement of the 5'-methyl group of BAT with H, F, Br, I or an ethyl group reduced the $IC_{50}$ by approximately 50% (Elliott *et al.*, 1987). However, none of these compounds has proved more inhibitory to the growth of murine L1210 or P388 cells than BAT (Brockman *et al.*, 1984; Elliott *et al.*, 1987).

### 1.12.5 *In vivo* regulators

Recent studies have revealed that pyridoxal phosphate (PLP) reversibly inhibits TS from *L.casei* with a $K_i$ value of 0.6-0.9µM. They also showed that inhibition is competitive with respect to dUMP (Chen *et al.*, 1988, 1989). The observed changes in the absorption spectra of PLP

resulting from the addition of TS (Chen *et al.*, 1989) were consistent with the formation of a thiohemiacetal by addition of a sulphydryl group to the carbonyl group of PLP (Matsuo, 1957; Buell & Hansen, 1960). This suggests an interaction between the catalytic cysteine residue and the carbonyl group of PLP, similar to that formed between TS and dUMP during catalysis (Moore *et al.*, 1986; Chen *et al.*, 1989; see Fig. 1.15b).

In addition, the common metabolite, glyceraldehyde-3-phosphate (G-3-P), was shown to act as a competitve inhibitor with respect to dUMP, with a $K_i$ value of 12 to 13μM (Bures *et al.*, 1991). Although no evidence for the formation of a thiohemiacetal between the sulphydryl group of TS and the carbonyl group of G-3-P was given by Bures and co-workers (1991), such an intermediate is known to exist in the reaction catalysed by G-3-P dehydrogenase (Harris & Waters, 1976). They proposed that the mechanism of interaction with both enzymes is similar (Bures *et al.*, 1991).

The most intriguing feature of these studies is the possible role of PLP and G-3-P in regulating DNA synthesis by modulating TS activity in both eukaryotic (Rode *et al.*, 1980; Chen *et al.*, 1989; Matherly *et al.*, 1989) and prokaryotic cells (Tolleson *et al.*, 1990). The observation that low levels of PLP are found in various rapidly proliferating cancer cells, and that the administration of pyridoxine or pyridoxal (precursors of PLP) inhibits the proliferation of these cells both *in vitro* and *in vivo* (Thanassi *et al.*, 1985, and references therein), augments this idea.

### 1.12.6 **Anti-sense oligonucleotides**

An alternative approach to the inhibition of enzyme function other than the use of drugs is to specifically inhibit enzyme synthesis. The use of anti-sense oligodeoxynucleotides (oligos) to arrest translation of specific mRNA molecules (Hélène & Toulmé, 1990) has recently been directed at the DHFR-TS mRNA of *Plasmodium*

*falciparum*. A 49-mer targeted at the conserved region of TS gave 50% inhibition at 6µM and 90% inhibition at 45µM (Sartorius & Franklin, 1991a). Such an oligo is very specific to the DHFR-TS mRNA of *P.falciparum*. It also differs considerably from the human TS mRNA, sharing 29/49 (59%) nucleotide sequence identity (Ayusawa *et al.*, 1986). However, the fact that certain useful modifications such as the use of methylphosphonates or poly(L-lysine) conjugates cannot be used for long oligos, may delay their development as anti-malarial drugs (Sartorius & Franklin, 1991b).

### 1.12.7 Oligopeptide inhibition of TS

Recently, oligopeptides have been shown to inhibit VZV TS activity. A linear pentadecapeptide corresponding to the amino acid sequence of ß-strand i (LPPCHTLCQFYVANG) inhibits TS *in vitro* ($IC_{50}$ = 50µM). N- and C-terminal deletions of this oligopeptide have defined the active core as CHTLCQFY, which corresponds precisely to those amino acids at the interface of the two monomers (H. S. Marsden and R. Thompson, personal communication). These oligopeptides are thought to disrupt activity by preventing the correct formation of the dimer.

## 1.13 Nucleoside and nucleoside analogue metabolism

To understand the mechanisms controlling the activation of nucleoside analogues, and thus potential modes of resistance, a thorough knowledge of nucleoside metabolism is necessary. The aim of this section is to describe the different pathways by which analogues are activated and to provide evidence that different pathways are followed in different tissues, and in different organisms. The pathways described below are for FU and its derivatives. FU metabolism resembles that of uracil, but some differences are apparent. The metabolism of other nucleoside analogues is largely unexplored, but thought to resemble that of FU and uracil.

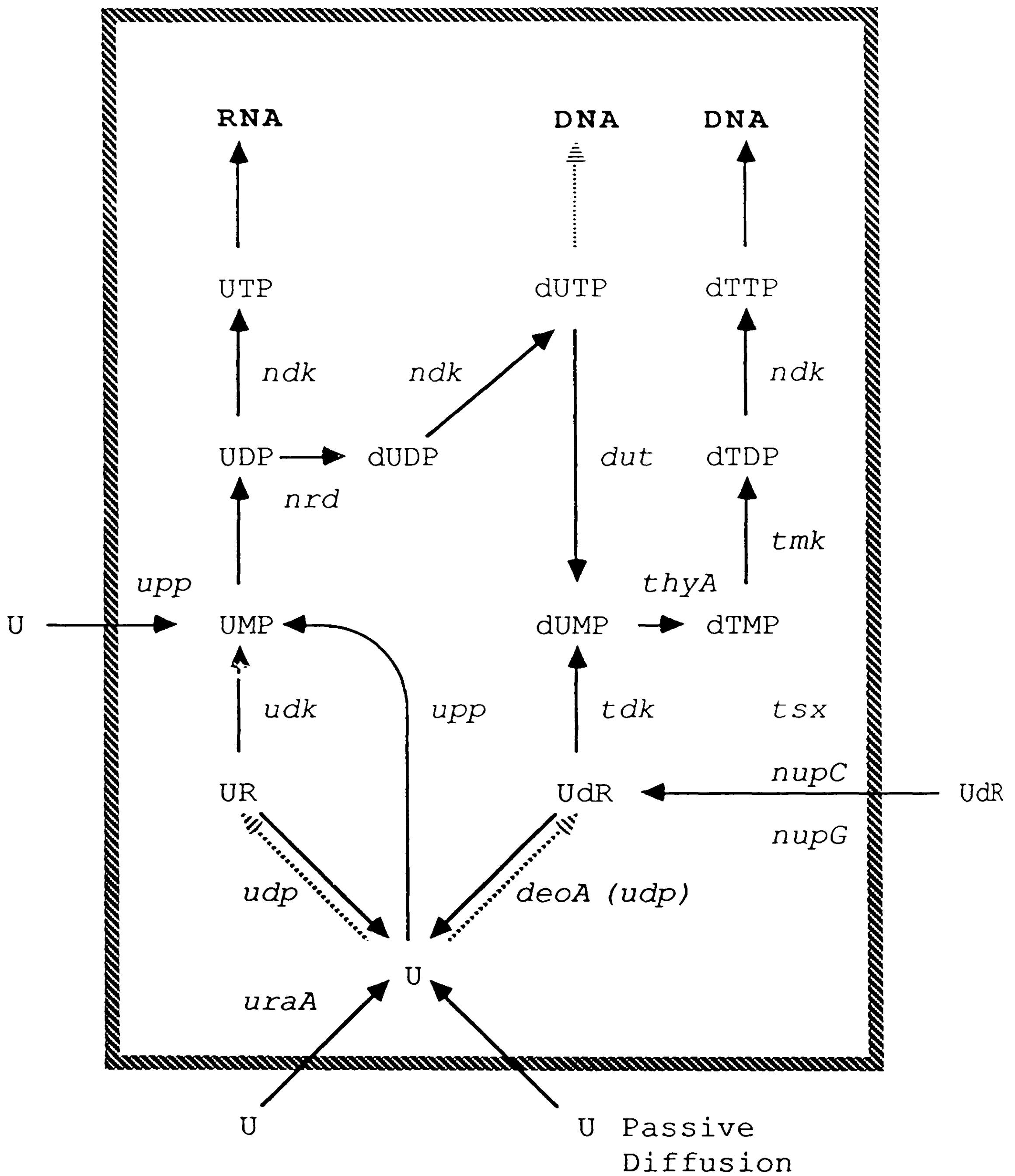



Fig. 1.16. **Nucleotide Metabolism.**
The pathways of nucleotide metabolism in *E.coli*. The structural and biochemical similarity of uracil and FU means that most pathways for their metabolism are the same (see text for details). Dotted lines indicate minor pathways
The following abbreviations are used for the various nucleotides described in this scheme: U - uracil; UR, UMP, UDP and UTP - uridine and its mono-, di- and triphosphate derivatives. UdR, dUMP, dUDP and dUTP - deoxyuridine and its mono-, di- and triphosphate derivatives. dTMP, dTDP and dTTP - the mono-, di- and triphosphate derivatives of thymidine. The following abbreviations are used for the genes of various proteins: *uraA*, *nupC*, *nupG* and *tsx* - genes for proteins involved in nucleobase and nucleoside uptake. *udp* - uridine phosphorylase; *deoA* - thymidine phosphorylase; *upp* - uracil phosphoribosyl transferase; *udk* - uridine kinase; *tdk* - thymidine kinase; *tmk* - thymidylate kinase; *thyA* - thymidylate synthase; *nrd* - ribonucleotide reductase; *dut* - dUTPase; *ndk* - nucleoside diphosphate kinase.

### 1.13.1 **Mammalian Systems**

There are at least two methods by which FU can inhibit the growth of cells. One is the inhibition of TS by FdUMP, a metabolite of FU, whilst the other involves the incorporation of FU into RNA via the formation of FUTP.

When examining the cytotoxic effects of FU mediated via TS, the two routes for conversion of FU to FdUMP must be considered (Cory *et al.*, 1977). The most direct route requires the formation of FUdR from FU, catalysed by thymidine phosphorylase (TdRPase) (and to a minor extent by uridine phosphorylase (URPase), followed by the phosphorylation of FUdR to FdUMP, a reaction catalysed by TK (Cory *et al.*, 1977).

The second route commences with the formation of FUMP, which may be synthesized from FU by two separate pathways. The first pathway involves the conversion of FU to FUR, catalysed by URPase, and the subsequent phosphorylation of FUR by uridine kinase (Cory *et al.*, 1977). The alternate pathway involves the direct formation of FUMP from FU in a reaction catalysed by uracil phosphoribosyl transferase (Cory *et al.*, 1977). To form FdUMP, FUMP must then be phosphorylated to FUDP by nucleotide kinases, reduced to FdUDP by ribonucleotide reductase, phosphorylated again to FdUTP and then converted to FdUMP by dUTPase (see Fig. 1.16; Cory *et al.*, 1977).

The formation of FdUMP alone is not necessarily sufficient for the inhibition of TS, as suitable levels of cofactor are also required for formation of the inhibitory ternary complex. In a human colon adenocarcinoma xenograft, the concentrations of $CH_2FH_4(Glu)_n$ have been shown to be too low to form stable inhibitory complexes with TS and FdUMP which may explain, in part, why certain tumours are insensitive to FU (Radparvar *et al.*, 1989a,b).

FU is also ineffective in the treatment of certain tumours where there is little or no dependance on *de novo* thymidylate synthesis. Mutants of a human carcinoma cell-line, selected for a deficiency in TS activity, were shown

to have a 32- and 750-fold increase in resistance to FU and FUdR respectively, compared to wild type parental cells. They also formed drug insensitive tumours in athymic nude mice (Houghton *et al.*, 1989), suggesting that such tumours can scavenge thymine *in vivo* to facilitate normal tumour growth, in the absence of *de novo* thymidylate synthesis.

Studies have shown that levels of enzymes involved in the activation of FU such as TdRPase vary tremendously in human tissues: high levels are seen in T- and B-lymphocytic malignant lymphomas and in normal liver and lung tissues, but are low in lymphoblastic and myelogenous leukaemic cells. This again could explain why some tumours respond to FU whilst others are insensitive (Yoshimura *et al.*, 1990). Moreover, the finding that TdRPase levels are high in human spleen yet remain undetectable in rat spleen, could explain the observed differences in response to FU between tumours from the same tissue of different organisms (Zimmerman & Seidenberg, 1964).

A block in the supply of thymidylate (thymidylate stress) is known to induce nucleotide pool imbalances and lead to rapid thymineless death of mammalian cells (Koyama *et al.*, 1982; Ayusawa *et al.*, 1988). Double-stranded DNA breaks (Ayuasawa *et al.*, 1988; Seno *et al.*, 1985), chromosomal damage, especially in regions of active DNA replication (Hori *et al.*, 1984a,b), and increased recombination (Ayusawa *et al.*, 1986; Barclay *et al.*, 1982; Kunz & Haynes, 1982), have all been shown to be a consequence of thymidylate stress.

More recently, an almost identical form of cell death has been observed in FUdR treated mouse cells, direct evidence that FUdR can induce thymineless death (Yoshioka *et al.*, 1987; Ayusawa *et al.*, 1988). In contrast, the use of cytotoxic drugs such as FUdR to induce thymidylate stress in androgen-independent rat prostatic adenocarcinoma AT-3 cells, results in the activation of programmed cell death. A characteristic feature of this process is the fragmentation of genomic DNA into discrete multiples of a nucleosomal unit of 180bp (Searle *et al.*, 1982; Kyprianou &

Isaacs, 1989).

FU can be toxic to cells due to the incorporation of FU, via FUTP, into RNA and can lead to cell death solely by the disruption of RNA processing (Akazawa *et al.*, 1986; Greenhalgh & Parish, 1990). These aberrant processing events include a dose-dependent increase in the mammalian DHFR mRNA intron V levels (Will & Dolnick, 1986), erroneous processing of rRNA precursors (Greenhalgh & Parish, 1989) and maturation of tRNA molecules (Santi & Hardy, 1987).

### 1.13.2 Bacterial systems

Although FU was originally developed as an anti-cancer drug (Duschinsky *et al.*, 1957), it was also shown to inhibit the growth of several species of micro-organisms, including *E.coli* (Heidelberger *et al.*, 1957). Studies investigating the mode of FU toxicity have suggested three different mechanisms of action; inhibition of TS, incorporation of FU into RNA and the disruption of cell-wall synthesis. To understand how FU kills the bacteria in these different ways, a description of the metabolism of FU is necessary. Comprehensive studies of these pathways have given an insight into the mechanisms of resistance (Neuhard & Nygaard, 1987; O'Donovan & Neuhard, 1970).

Before FU can be metabolised, it must enter the cell by one of three mechanisms, the first being by a passive diffusion process (Munch-Peterson & Mygind, 1983; Neuhard & Nygaard, 1987). The gene product of *uraA* is responsible for a specific transport system that brings FU into the cell (Burton, 1977, 1983), whilst phosphoribosyltransferase, encoded by *upp*, catalyses the uptake and concomitant conversion of FU to the respective monophosphate derivative by the addition of a ribose-5'-phosphate group (Neuhard & Nygaard, 1987).

FU is predominantly converted to FUMP (if it was not already as part of the uptake process) by uracil phosphoribosyltransferase (Neuhard & Nygaard, 1987). Mutant strains of *E.coli* that lack this enzyme are resistant to FU. FU may also be converted to FUMP by the concerted

action of URPase and UK, but this route only operates in the presence of high intracellular levels of FU or ribose-5'-phosphate donors. In contrast to *E.coli* mutants in *upp* that confer resistance to FU, an *E.coli udp udk* double mutant is sensitive to FU (Neuhard & Nygaard, 1987). The formation of FUdR from FU, catalysed by TdRPase is minimal, as this enzyme functions in a predominantly catabolic fashion (Neuhard & Nygaard, 1987; see Fig. 1.16).

FUMP is readily converted to FUDP whereupon it has two fates. When further phoshorylated to form FUTP it can become incorporated into RNA (Heidelberger, 1957; Horowitz & Chargaff, 1959; Brockman *et al.*, 1960; Brockman & Anderson, 1963). When FUTP is incorporated into mRNA, the result is the synthesis of enzymes with altered activities (Horowitz & Chargaff, 1959; Horowitz *et al.*, 1960; Bussard *et al.*, 1960). This was shown to occur due to aberrant amino acid incorporation into proteins (Gros, 1960; Naona & Gros, 1960a,b), although a second group failed to find evidence for such a phenomenon (Aronson, 1961).

The alternate fate for FUDP is conversion to FdUMP which follows the same pathway as the conversion of UDP to dUMP, via ribonucleotide reductase (Neuhard & Nygaard, 1987). FdUMP can then inhibit TS as described (see Section 1.12.i). FUdR can also be used to inhibit the growth of *E.coli*. It enters the cell via one of three specific nucleoside uptake systems, encoded by the *nupC*, *nupG* or *tsx* loci (Munch-Peterson & Mygind, 1983; Hantke, 1976; McKeown *et al.*, 1976). Once inside the cell, FUdR is phosphorylated by TK, to form FdUMP (Neuhard & Nygaard, 1987). A strain of *E.coli* with mutations in the *udp* (URPase) and *deoA* (TdRPase) genes cannot degrade FUdR and is more sensitive to FUdR than wild type. However, if such a strain is simultaneously deficient in TK activity, it will be resistant to FUdR (Neuhard & Nygaard, 1987). The inhibition of TS activity results in the so-called 'thymineless death' of *E.coli*.

In bacteria, FU can exert its toxicity by an additional method (Tomasz & Borek, 1959), inhibiting the incorporation

of αε-diaminopimelic acid into the cell walls, leading to the formation of osmotically sensitive 'spheroplast'-like bodies (Rogers & Perkins, 1960; Tomasz & Borek, 1962).

### 1.14 **Resistance to TS inhibitors**

As discussed above, the arrest of cell or tumour growth by TS inhibitors, is only accomplished under certain conditions. However, in some circumstances tumours and cells can develop or induce mechanisms that allow both to overcome the toxic effects of these drugs, thus conferring resistance. Examples of drug resistance, and the molecular basis of this phenomenon, will be described for both mammalian and bacterial systems.

#### 1.14.1 **Mammalian systems**

Shortly after the introduction of FU reports of drug resistance began to emerge (Reichard *et al.*, 1959; Heidelberger *et al.*, 1960a,b). Three possible mechanisms have been identified; the prevention of activation, TS overproduction and structural alterations to TS.

Characterization of a variety of FU and FUdR resistant cell lines has shown that a significant decrease in any of a number of enzymes that prevent the formation of FdUMP or FUTP, thereby preventing inhibiton of TS (Reichard *et al.*, 1959; Anderson & Law, 1960; Umeda & Heidelberger, 1968; Slack *et al.*, 1976; Kessel & Wodinsky, 1970; Mulkins & Heidelberger, 1982) or disruption of RNA processsing (Brockman & Law, 1960), can be responsible for the drug resistant phenotype.

The cause of resistance to FUdR in at least two murine cell lines has been shown to be due to a 6- to 10-fold overproduction of TS (Baskin & Rosenberg, 1975; Priest *et al.*, 1980). This form of resistance has also been observed for mouse FM3A mammary adenocarcinoma cells resistant to the folate analogue CB3717 (Danenberg & Danenberg, 1989). TS was overproduced by up to 200-fold in the presence of CB3717, but the levels of TS decayed, with a half-life of

about 4 weeks, upon removal of the drug. The cause of overproduction was specific amplification of the TS gene (Danenberg & Danenberg, 1989).

TS analysed from crude extracts of a fluorouracil resistant cell line (Heidelberger *et al.*, 1960b), was insensitive to FdUMP inhibition (Heidelberger *et al.*, 1960a). This suggests that the structure of the active site is altered in the mutant enzyme in such a way that it can descriminate between dUMP and FdUMP, in a manner that the wt enzyme cannot. Unfortunately, the resistant cell line was lost before the enzyme could be purified and studied in more detail (Hartmann & Heidelberger, 1961).

Heidelberger and colleagues proposed that the enzyme is unlikely to distinguish between the dUMP and FdUMP on the basis of the size of the fluorine and hydrogen atoms as their Van der Waals radii are very similar (Pauling, 1944; Heidelberger *et al.*, 1960a). Instead, they proposed that the fluorine substituent, which would increase the acidity of the pyrimidine carbonyls, which may permit the enzyme to distinguish between dUMP and FdUMP (Heidelberger *et al.*, 1960a). The determination of the structure of the *E.coli* ternary complex has recently shown that the pyrimidine ring C4 carbonyl group is directly involved in hydrogen-bonding, with a completely conserved residue in the active site, thus giving some support for this mechanism (Matthews *et al.*, 1990b; Montfort *et al.*, 1990; Finer-Moore *et al.*, 1990).

Jastreboff and co-workers (1983) isolated a different cell line resistant to FUdR. A comparison of TS purified from both the FUdR resistant cell line, and the wild type cell line from which it was derived, revealed that TS from the FUdR resistant cell line had a higher turnover number (91 $min^{-1}$ versus 31 $min^{-1}$) and increased $K_i$ value for FdUMP (19nM versus 1.9nM) compared to the enzyme from the wild type cell line (Jastreboff *et al.*, 1983). Although this altered form of TS might be expected to account for the drug resistant phenotype of these cells, it may not be the sole determinant. This is because the levels the TS are

twice those seen in the wild type cell line, and a deficiency in TK could not be ruled out (Jastreboff *et al.*, 1983). No further reports concerning the further characterization of TS from this cell line have been published.

A cell line derived from a human colorectal tumour that expressed two forms of TS differing in electrophoretic mobility, showed resistance to FUdR (Berger & Berger, 1988). When cell lines expressing either the variant or wt form of TS were isolated, drug resistance was shown to be linked to expression of the variant form of TS (Berger *et al.*, 1988). The variant form of TS has a more basic pI and exhibits a three-fold reduction in affinity for both FdUMP and $N^5,N^{10}$-$CH_2FH_4$ (Berger *et al.*, 1988). The isolation and sequence determination of a cDNA clone, specific to this form of TS, has shown it differs from wild type TS only by substitution of histidine for the highly conserved tyrosine at position 33 (Y21 in VZV TS, see Fig. 1.6; Barbour *et al.*, 1990). It is possible that mutations at other loci such as TK may contribute to the drug resistant phenotype of cells expressing the TS variant (Berger *et al.*, 1988; Barbour *et al.*, 1990), and a detailed study of the kinetic parameters of the purified enzyme would resolve the situation. This is the only recorded instance of a variant of TS where the amino acid changes, possibly responsible for resistance to FdUMP, have been determined.

### 1.14.2 **Bacterial systems**

In bacteria resistance to fluoropyrimidines has been investigated during the course of experiments designed to elucidate the pathways of nucleotide metabolism (Neuhard & Nygaard, 1987; O'Donovan & Neuhard, 1970). In contrast to mammalian systems, only one mechanism of drug resistance has been described, namely the prevention of activation.

Mutant strains of *E.coli* that lack uracilphosphoribosyl-transferase activity are resitant to FU. Strains with an additional deficiency in uridine kinase, are also resistant

to FUR, but remain sensitive to FUdR (Neuhard & Nygaard, 1987). Strains that lack TK are not resistant to FUdR (O'Donovan & Neuhard, 1970), but mutants that lack TdRPase and URPase activity in addition to a deficiency in TK activity, are resistant to FUdR (Neuhard & Nygaard, 1987). Some strains of *E.coli* and *Streptococcus faecalis* resistant to FU have been shown to lack URPase activity and show decreased incorporation of FU into RNA respectively (Brockman *et al.*, 1960; Brockman & Law, 1960).

The overproduction of TS leading to resistance to FU has not been studied in bacteria. However, at least two strains of *L.casei* have been isolated that are resistant to the DHFR inhibitors MTX and dichloromethotrexate. These overproduce both DHFR and TS 100-fold, by the process of gene amplification (Crusberg *et al.*, 1970). Extracts of *S.faecalis* resistant to MTX also show increased levels of DHFR and TS (Suolinna *et al.*, 1967). Such strains would be expected to show resistance to FU when compared with the parental strains from which they were derived.

More importantly, the *Lactobacillus* strains were instrumental in many of the studies of TS, because they provided high levels of enzyme at a time preceeding the modern techniques that allow the high-level expression of proteins (see Sections 1.6 and 1.7).

## 1.15 **Aims of the project**

The aim of the project was to develop a system, employing random mutagenesis techniques, to allow the isolation of variants of VZV TS that were resistant to one or more of a range of nucleoside analogues. The identification of the amino acid changes in these variants would yield information as to which residues were involved in substrate recognition and binding.

As a complementary approach, we chose site-directed mutagenesis to explore the roles of secondary structure elements which early investigations had suggested could be functionally important.

# 2. MATERIALS AND METHODS

## 2.1 Bacteria and bacteriophage

### 2.1.1 Bacterial strains

| Strain | Genotype, source and reference |
|---|---|
| BW313 | *relA1, spoT1, dut-1, ung-1, thi-1*, Hfr KL16(PO45)<br>R.Hayward, University of Edinburgh; Künkel (1985) |
| C600 (*tdk*) | $F^-$, *thi1, thr1, leuB6, lacY1, tonA21, supE44*,<br>Laboratory Stock; Appleyard, 1954<br>*tdk* variant; Kaehler (1984). |
| CB158 | *mutD*<br>Laboratory stock; Degnen & Cox (1974). |
| CJ236 | *dut1, ung1, thi1, relA1* / pCJ105<br>P.C.Huang, John Hopkins University; Kunkel (1984); Joyce & Grindley (1984) |
| DH5α | *supE44*, Δ(*argF-lacZYA*)*U169, hsdR17, recA1, endA1, gyrA96, thi-1, relA1*, (ø80lacZ M15)<br>BRL; Hanahan (1983). |
| JM101 | *supE44, thi*, Δ*(lac-proAB)* / *F'*[*traD36*, *proAB*$^+$, *lacI*$^q$,*lacZ M15*]<br>Laboratory stock; Messing (1983) |
| JM109 | *recA1, supE44, endA1, hsdR17, gyrA96, relA1, thi*, (*lac-proAB*) / *F'*[*traD36*, *proAB*$^+$, *lacI*$^q$, *lacZ M15*]<br>Laboratory stock |
| MC1061* | *hsdR, mcrAB, araD139*, Δ(*araABC-leu*)*7679, lac X74, galU, galK, rpsL, thi1*<br>Laboratory stock; Meissner *et al.* (1987) |
| RZ1032 | *relA1, spoT1, dut1, ung1, thi1, supE44, zbd-279::Tn10*, Hfr KL16(PO45)<br>R.Hayward, University of Edinburgh; Kunkel (1985) |
| SØ928* | Δ*deo*-11 (= Δ*deoA,B,C,D*), *lac, thi1, upp1, udp1, tonA*,<br>B.Mygind, University of Kobenhavn; Valentin-Hansen *et al.* (1978) |

| | |
|---|---|
| TG1 | *supE44, hsd 5, thi1,* Δ*(lac-proAB) / F'[traD36, proAB⁺, lacI^q, lacZ M15]* Amersham; Gibson, (1984) |
| W3110 | *lacI^qL8* Laboratory stock; Brent & Ptashne (1981) |
| XL-1* | *supE44, hsdR17, recA1, endA1, gyrA96, thi1, relA1, lac⁻ / F'[proAB⁺, lacI^q, lacZ M15 Tn10]* Stratagene; Bullock *et al*.. (1987) |
| χ2913 | *thyA 752* *E.coli* Genetic Stock Centre; R. Curtiss (personal communication) |

Note: *thyA* derivatives were made of strains marked * as described (see Section 2.7.2).

2.1.2 **Bacteriophage**

Bacteriophage M13R408 (Russel *et al*., 1986) was provided by J. Mottram, University of Glasgow. Bacteriophage M13K07 (Vieira & Messing, 1987) was a laboratory stock.

2.2 **Plasmids**

2.2.1 **Plasmids provided**

| Plasmid | Source & Reference |
|---|---|
| pBS+/pBS- | Stratagene<br>Fernandez *et al*. 1987 |
| pCB105 | D. Sherratt, University of Glasgow;<br>Boyd & Sherratt (1986) |
| pGEM3Zf+/- | Promega<br>Dotto *et al*. (1984) |
| pGL271 | *Nco*I-*Hin*dIII fragment from VZV *Eco*RI *m* cloned into pKK240-11;<br>R. Thompson, University of Glasgow;<br>Thompson *et al*. (1987) |
| pTMORF36.1 | 1043bp fragment containing VZV orf36 (TK) cloned into pKK233-2;<br>T. McKee, University of Glasgow;<br>McKee (1990) |

| | |
|---|---|
| pUC19-HTS | Human TS cloned under the control of an *L.casei* promoter; P. J. Greene and D. V. Santi, University of California, San Francisco (personal communication) |

### 2.2.2 Plasmid constructs

| Plasmid | Details |
|---|---|
| pAD113 | TS containing *Eco*RI fragment of pGL271 cloned into pBS+ (contains a C787T nucleotide substitution in the pBS+ backbone (numbering of pBS+) |
| pAD203 | TS containing *Eco*RI/*Xba*I fragment of pAD113 cloned into pGEM-3Zf(+) previously digested with *Eco*RI/*Xba*I |
| pAD301 | *rrnB* containing *Hinc*II/*Pst*I fragment of pGL271 cloned into pAD203 previously digested with *Hin*dIII and *Pst*I, *the Hin*dIII overhanging end converted to a blunt end |
| pAD403 | TS containing *Bam*HI/*Hin*dIII fragment of pAD113 cloned into *Bam*HI/*Hin*dIII digested pBS- (contains the mutations A-48G, A-47T; see pAD768) |
| pAD507 | TK containing *Eco*RI/*Hin*dIII fragment of pTMORF36.1 was cloned into pBS+ previously digested with *Eco*RI/*Hin*dIII |
| pAD601 | Constructed exactly as pAD507 except that pBS- was used |
| pAD768 | Oligo 68 was used to change the promoter of pAD403 by site-directed mutagenesis |
| pAD876 | The human TS containing *Sac*I/*Hin*dIII fragment of pHTS-pUC19 was cloned into pBS- previously digested with *Sac*I/*Hin*dIII |
| pAD901 (pCBTK3) | The *Bam*HI/*Hin*dIII fragment of pVZVTK+5 (pAD507/52 - see Section 3.4.i) was cloned into pCB105 previously digested with *Bam*HI/*Hin*dIII |

## 2.3 Bacterial growth media

### 2.3.1 Liquid media

Luria-Bertani Medium (LB)

Per litre:
10g NaCl
10g Bactopeptone
5g yeast extract
Adjusted to pH7.0
prior to sterilization

2xYT Medium (2YT)

Per litre:
16g Bactotryptone
10g yeast extract
5g NaCl
Adjusted to pH7.0
prior to sterilization

Minimal Medium (M9)

Per 400ml:
300ml $H_2O$
80ml M9[x5] salts (see below)
20ml 15% casamino acids
4ml 20% glucose
500µl 1M $MgCl_2$
100µl 0.8mg/ml thiamine

M9[x5] salts

Per litre:
64g $Na_2HPO_4.7H_2O$
15g $KH_2PO_4$
5g $NH_4Cl$
2.5g NaCl

### 2.3.2 Solid media

LB

As above except supplemented with 15g Bacto-agar

M9

As above except 300ml $H_2O$ replaced by 300ml
2% Bacto-agar

Top Agar (TA)

Per Litre:
7g Bacto-agar

## 2.4 Chemicals, drugs, radiochemicals and enzymes

### 2.4.1 Chemicals

All chemicals were purchased from BDH or Sigma Chemical Company except where stated below. Dimethyl-sulphoxide (DMSO) (spectophotometric grade) was from Fluka and was stored at -70 C in 1ml aliquots. *N,N,N',N'*-tetramethyl-ethylenediamine (TEMED) and ammonium persulphate (APS) were from Bio-Rad. Bactopeptone, casamino acids and yeast

extract were from Difco. SeaPlaque and NuSeive agarose were from FMC BioProducts. Organic solvents and acids were from May and Baker. Ecoscint A **was from** National Diagnostics. α-thio-dNTPs, dNTPs and **ddNTPs** were from Pharmacia. Glycogen was from **Boehringer-Mannheim**. Ethyl methanesulphonate (EMS) was from Koch-Light laboratories. Ampicillin was from Beecham research. Wacker silane GF38 was purchased from Wacker Chemicals.

2.4.2 **Drugs**

All drugs not purchased from BDH or Sigma Chemical Company were the kind gift of Wellcome PLC. These included PYaraU, PYUdR, EYaraU, EYUdR, BVdU, FIAC and ACV, and were supplied as 40mM solutions in DMSO.

2.4.3 **Radiochemicals**

Radiochemicals were purchased from Amersham and New England Nuclear/DuPont as follows:

| | | |
|---|---|---|
| [$\alpha^{32}$P]dATP | ~3,000mCi/mmol | (PB10204) |
| [$\alpha^{35}$S]dATP | >1,000mCi/mmol | (SJ1304) |
| [$\gamma^{32}$P]dATP | >5,000mCi/mmol | (PB10218) |
| 5-[$^{3}$H]dUMP | 20mCi/mmol | (TRK287) |
| [methyl-$^{3}$H]TdR | 25mCi/mmol | (TRK120) |
| [$^{35}$S]Sequetide™ | 400μCi | (NEG-034N) |

2.4.4 **Enzymes**

T4 DNA ligase and restriction enzymes were from Boehringer-Mannheim and the manufacturers SuRE/cut buffer system was employed. T7 DNA polymerase and T4 gene 32 protein were from Pharmacia. RNase A and lysozyme were from Sigma. T4 DNA polymerase and T4 polynucleotide kinase were from New England Biolabs. Reverse **transcriptase was** from Life Sciences Incorporated. **Sequenase™** from United States Biochemicals (including **sequencing solutions**) and the large fragment of DNA polymerase I **(Klenow) was from** Northumbria Biologicals Limited.

## 2.5 Commonly used buffers

| | |
|---|---|
| TBE | 90mM Tris, 90mM boric acid, 1mM EDTA |
| TE | 10mM Tris-Cl pH8.0, 1mM EDTA |
| Ø/$CHCl_3$ | phenol/chloroform 1:1 ($^v/_v$) |
| Ø/TE | phenol equilibrated with TE |
| 5xRTB | 250mM Tris-Cl pH8.3, 30mM $MgCl_2$, 300mM KCl, 5mM DTT, 450µg/mL BSA |
| 10xKB | 1M Tris-Cl pH8.0, 100mM $MgCl_2$, 70mM DTT |
| 10x ligase buffer | 660mM Tris-Cl pH7.5, 100mM $MgCl_2$, 10mM ATP, 10mM DTT |
| 10xNTB | 500mM Tris-Cl pH7.2, 100mM $MgSO_4$, 1mM DTT, 500mg/ml BSA |
| 20xSSC | 175.3g/l NaCl, 88.2g/l Trisodium citrate, pH7.0 |
| 50xDHB | 10g/l Ficoll, 10g/l polyvinylpyrrolidone, 10g/l BSA (filtered through Whatman paper). |

## 2.6 Growth and maintenance of bacteria and bacteriophage

### 2.6.1 Bacteria

#### 2.6.1.1 Laboratory stocks of *E.coli*

25µl of a glycerol stock of each laboratory strain of strain of *E.coli* was used to inoculate 5ml LB, supplemented with antibiotics as appropriate, and incubated overnight at 37°C. A 10µl loopful of the culture was streaked for single colonies on selective plates and incubated overnight at 37°C. A well isolated single colony was used to make a fresh glycerol stock.

#### 2.6.1.2 New strains of *E.coli*

New strains provided by the *E.coli* Genetic Stock Centre on filter discs were placed on LB plates and moistened with a few drops of sterile $H_2O$ and streaked for single colonies and incubated overnight at 37°C. Isolated single colonies were streaked onto selective media, incubated overnight at 37°C and well isolated single colonies used to make a glycerol stock. New strains, provided as glycerol stocks,

were treated as above.

2.6.1.3 **Overnight cultures**

A well isolated single colony, or 25µl of a glycerol stock was used to inoculate 5 or 10ml LB, supplemented with the appropriate antibiotics. The culture was incubated, with vigorous shaking, at 37°C overnight.

2.6.1.4 **Glycerol stocks**

A 10ml overnight culture of the appropriate clone was set up in a 15ml Falcon tube. Cells were pelleted by centrifugation at 3000rpm for 10min in a Beckman GPR benchtop centrifuge. The cell pellet was resuspended in 5ml 2% bactopeptone. 5ml 80% glycerol and 50µl 100µg/ml TdR were added, mixed by repeated inversion and stored at -20°C.

2.6.1.5 **Antibiotics**

Antibiotics were used at the following final concentrations, unless otherwise stated. Ampicillin (Amp) 100µg/ml; Chloramphenicol (Cm) 10µg/ml; Tetracyclin (Tet) 10µg/ml).

2.6.2 **Bacteriophage**

2.6.2.1 **Titration of helper phage**

10ml 2YT was inoculated with 1ml of an overnight culture (grown in 2YT) of *E.coli* TG1 and incubated, with vigorous shaking, at 37°C for 1h. 3mL aliquots of molten top agar were equilibrated to 45°C in a waterbath. A series of 1:10 dilutions of the helper phage were made in 2YT. 200µl *E.coli* TG1 and 100µl of diluted helper phage (10-fold dilutions in the range $10^6$ to $10^9$) were mixed together, added to the molten TA and immediately poured onto LB plates. Once the top agar layer solidified the plates were incubated at 37°C overnight. Plaques were counted the following day and the phage titre calculated.

2.6.2.2 **Growth of helper phage**

100ml 2YT was inoculated with a single plaque of helper phage and incubated at 37°C, with vigorous shaking, overnight. The culture was cleared by centrifugation at 9,000rpm at 4°C for 10min. The supernatant was decanted to a fresh tube and recleared by centrifugation. This supernatant was transferred to a 300ml bottle, heated at 55°C for 30min, allowed to cool to room temperature and stored at 4°C. The titre of the virus was determined as described above.

## 2.7 Bacterial manipulations

### 2.7.1 Transformation

Two methods were routinely used for the preparation of competent cells and transformation.

2.7.1.1 **Calcium chloride** (Humphreys *et al.*, 1979).

100ml LB was inoculated with 3ml of an appropriate standing overnight culture and incubated at 37°C, with vigorous shaking, until $A_{600}$ reached ~0.3. Cells were pelleted by centrifugation at 9000rpm in a Sorvall SS34 for 5min at 4°C prior to resuspension in 50ml ice-cold 10mM $CaCl_2$ (supplemented with 50µg/ml TdR for *thyA* strains). The cells were re-pelleted as before and resuspended in 50ml ice-cold 75mM $CaCl_2$. Cells were incubated for 30min on ice, repelleted and resuspended in 3ml 75mM $CaCl_2$.

10 to 100ng DNA, in a small volume of $H_2O$ (<5µL), was added to 200µl competent cells and incubated on ice for 30min prior to heat shock at 42°C for 2min. Cells were used to inoculate 3ml LB and incubated at 37°C, with vigorous shaking for 1h. 100µl was spread onto LB plates containing the appropriate antibiotic, and incubated at 37°C overnight.

2.7.1.2 **DMSO-PEG4000** (Chung and Miller, 1988).

20ml LB was inoculated with 1ml of an appropriate, standing overnight culture and incubated at 37°C, with vigorous shaking, until $A_{600}$ reached ~0.3. Cells were

pelleted as above and resuspended in freshly prepared, ice-cold transformation and storage buffer (TSB). TSB consists of 7.25ml LB, 2.0ml 50% PEG-4000, 500µl DMSO, 100µl 1M $MgCl_2$, 100µl 1M $MgSO_4$ and 50µl TdR. Cells were incubated on ice for 15 to 30min prior to use.

10 to 100ng DNA, in a small volume of $H_2O$ (<5µl), was added to 50µl competent cells and incubated on ice for 5 to 30min. 450µl of TSB was added to the cells and incubated at 37°C for 30min to 2h. 100 to 200µl was spread onto LB plates supplemented with the appropriate antibiotics and incubated at 37°C overnight.

For both methods a positive control (~10ng supercoiled plasmid DNA) and a negative control (no DNA) were included.

### 2.7.2 Generation of *thyA* strains of *E.coli*

2ml minimal media supplemented with TdR (100µg/ml) and trimethoprim (20µg/ml) was inoculated with 40µl of an appropriate overnight culture and incubated, with vigorous shaking, at 37°C overnight. 40µl of this culture was used to inoculate 2ml of fresh media and incubated as before. A single loopful of this culture was used to streak for single colonies on minimal media plates supplemented with TdR (100µg/ml) and incubated at 37°C overnight. Individual single colonies were patched onto minimal media plates and minimal media plates supplemented with TdR (100µg/ml) and incubated at 37°C overnight. Colonies growing only on minimal media plate supplemented with TdR were classified as $Thy^-$ strains (Stacey & Simpson, 1965).

## 2.8 DNA isolation

### 2.8.1 Mini-prep plasmid DNA preparation

Mini-prep DNA was prepared from *E.coli* using the method of Maniatis *et al.*, (1982). 1ml of an overnight culture was pelleted by centrifugation at 11,000rpm for 15s in a bench-top mini-centrifuge. The supernatant was discarded and the pellet resuspended in 100µl fresh lysozyme solution (4mg/ml lysozyme, 50mM glucose, 25mM Tris-Cl pH8.0, 10mM

EDTA pH8.0) and incubated at room temperature for 5min. 200µl 0.2M NaOH, 1% SDS was added and incubated on ice for 5min. 150µl potassium acetate solution (5.0M KOAc 60ml, glacial acetic acid 11.5ml, $H_2O$ 28.5ml) was added, vortexed for 10s and incubated on ice for 5min. After centrifugation for 5min, the supernatant was removed and phenol extracted prior to ethanol precipitation. The lyophilised pellet was resuspended in 50µl $H_2O$. 1 to 5µl was used for transformation or restriction enzyme analysis.

### 2.8.2 Midi-prep plasmid DNA preparation

This method was routinely used for the isolation of up to 100µg DNA of purity comparable to caesium chloride banded DNA. The method utilizes the Qiagen™ pack-100 gravity flow DNA purification columns (Diagen). 100ml of LB, supplemented with the appropriate antibiotic, was inoculated with 1ml of an overnight culture and incubated at 37°C, with vigorous shaking, until $A_{600}$ reached ~0.6. Cells were harvested by centrifugation at 7,000rpm at 4°C for 10min in a Sorvall GS3 rotor. The supernatant was discarded. The bacterial pellet was resuspended in 4ml buffer P1 (25mM Tris-Cl pH8.0, 10mM EDTA pH8.0, 100µg/ml RNase A) and transferred into a disposable 30ml tube suitable for the Sorvall SS34 rotor. 4ml buffer P2 (200mM NaOH, 1% SDS) was added, mixed gently and incubated at RT for 5min. 4ml buffer P3 (2.55M KOAc pH4.8) was added, mixed gently and the sample immediately centrifuged at 9,000rpm in a Sorvall SS34 at 4°C for 30min. A Qiagen™ pack-100 column was equilibrated with 2ml buffer QBT (750mM NaCl, 50mM 3-(*N*-morpholino)-propanesulphonic acid (MOPS), 15% EtOH, 0.15% Triton X-100 pH7.0). The supernatant was applied directly to the pre-equilibrated column and left to flow through. The column was washed twice with 4mL aliquots of buffer QCT (1,500mM, 50mM MOPS, 15% EtOH, 0.15% Triton X-100 pH7.0). DNA was eluted with 2ml of buffer QFT (1500mM NaCl, 50mM MOPS, 15% EtOH, 0.15% Triton X-100 pH8.0). DNA was precipitated with 0.8vol isopropanol, by centrifugation at 9,000rpm at 4°C for 30min in a Sorvall

SS34 rotor. The lyophilised DNA pellet was resuspended in $H_2O$, and the concentration determined spectrophotometrically (1 $A_{260}$ unit = 50µg/mL plasmid DNA).

### 2.8.3 Maxi-prep DNA plasmid preparation

350ml LB, supplemented with the appropriate antibiotic, was inoculated with 10ml of overnight culture and incubated at 37°C, with vigorous shaking, until $A_{600}$ reached ~0.6. 175µl chloramphenicol (200mg/ml in EtOH) was added and the incubation continued overnight. Cells were harvested by centrifugation at 9,000rpm at 4°C for 10min in a Sorvall GS3 rotor and the supernatant discarded.

Plasmid DNA was extracted by an adaptation of the 'maxi-boiling technique' of Holmes and Quigley (1981). The bacterial cell pellet was resuspended in 10ml STET (0.1M NaCl, 10mM Tris-Cl pH8.0, 1mM EDTA pH8.0, 5.0% Triton X-100). This was transferred to a 100ml flask containing 15ml STET. 3ml of freshly prepared lysozyme solution (10mg/ml) was added. The flask was heated over the flame of a Bunsen burner until *just* boiling and transferred to a boiling water bath for 40s. The contents were transferred to tubes suitable for the Sorvall SS34 rotor and centrifuged at 18,000rpm at 4°C for 20min. The supernatant was decanted into a 50ml Falcon tube. An equal volume of isopropanol was mixed with the supernatant and incubated at RT for 15min. After centrifugation, at 3,000rpm for 15min, in a Beckman GPR tabletop centrifuge, the supernatant was discarded and the pellet dried well prior to being resuspended in 10.5ml $H_2O$. CsCl was added to give a final density of 1.55g/mL. Ethidium bromide (EtBr) was added to a final concentration of 0.5mg/ml. The mixture was incubated on ice for 10min and then cleared by centrifugation at 3,000rpm at 4°C for 10min. The cleared supernatant was decanted to a quick-seal tube and the DNA was banded by ultra-centrifugation at 40,000rpm at 15°C for 48h in an angled Ti50 rotor. DNA was visualized by long-wave UV illumination and the lower, supercoiled plasmid DNA band, was removed with a large gauge needle and

syringe. DNA was extracted twice with water saturated butan-1-ol and dialysed against water for 5h at RT. After ethanol precipitation, the pellet was washed, dried and resuspeded in $H_2O$. The DNA concentration was determined spectrophotometrically.

### 2.8.4 ssDNA preparation from phagemid vectors

#### 2.8.4.1 Midi-scale ssDNA preparation

ssDNA was prepared essentially as described (Dente *et al.*, 1985). 20ml 2YT in a 50ml Falcon tube, was inoculated with 660µl overnight culture (grown in 2YT) and incubated with vigorous shaking at 37°C, until $A_{600}$ reached ~0.28-0.30. The helper phage M13R408 was added at a multiplicity of infection (moi) of 20:1 (assuming a bacterial cell density of ~$2.5x10^8$ cells/ml at $A_{600}$ ~0.28-0.30) and the culture was incubated for a further 8h. The culture was cleared by two centrifugation steps at 9,000rpm at 4°C for 10min in a Sorvall SS34 rotor. The bacterial cell pellet was discarded. The cleared supernatant was stored at 4°C or processed immediately.

A quarter volume of 3.75M $NH_4OAc$/20%PEG-6000 was added to the supernatant, vortexed briefly and the solution incubated on ice for 30min. The ssDNA and phage were pelleted by centrifugation at 9,000rpm at 4°C for 20min. The supernatant was discarded and the walls of the tube dried thoroughly with a tissue. The pellet was resuspended in 200 to 400µl TE, and $CHCl_3$ extracted twice. The DNA was repeatedly extracted with O/$CHCl_3$ until there was no material visible at the interface. A final $CHCl_3$ extraction was performed prior to $NH_4OAc$/EtOH precipitation. The lyophilised DNA pellet was resuspended in $H_2O$, and the concentration determined spectrophotometrically (1 $A_{260}$ unit = 40µg/ml ssDNA).

#### 2.8.4.2 Maxi-scale ssDNA preparation

This method is essentially the same as the midi-scale method except as follows. 200ml of 2YT, in a 2l baffled flask, was inoculated with 6.6ml overnight culture and

incubated and infected with helper phage as before. The supernatant was cleared by two centrifugation steps using a Sorvall GS3 rotor. The ssDNA was purified as described in above.

### 2.8.5 Plasmid DNA sequencing preparations

The following method of DNA preparation is that of Hattori and Sakaki (1986). 1ml of overnight culture was treated exactly the same as in 2.8.1, except that the final DNA pellet was resuspended in 50µl TE. RNaseA was added to a final concentration of 10µg/ml and incubated at 37°C for 30min. 30µl 20% PEG-6000/2.5M NaCl was added, mixed thoroughly by vortexing and incubated on ice for 1h. After a 10min spin, the DNA pellet was washed, lyophilised and resuspended in 18µl TE. To denature the DNA, 2µl 2M NaOH was added and incubated for 5min at RT. 100µl 100% EtOH and 8µl 5M $NH_4OAc$ were added and incubated in a dry ice bath for 10min, followed by 10min centrifugation. The DNA pellet was washed with 75% EtOH, lyophilized and resuspended in 20µl $H_2O$. 2 to 4µl was used directly for sequencing or stored at -20°C.

## 2.9 DNA manipulations

### 2.9.1 Restriction enzyme digests

Restriction enzyme digests were usually performed in 10 or 20µl reaction volumes. Restriction enzymes were added at a concentration of 1 to 10 units (U)/µg DNA and incubations performed at 37°C for 1 to 4h or overnight. Commercial restriction enzyme buffers were usually provided as 10x solutions and were diluted appropriately. When mini-prep DNA was digested, RNase A was added at a final concentration of 50µg/ml. For cloning experiments, DNA was always Ø/$CHCl_3$ extracted and ethanol precipitated prior to use.

### 2.9.2 Removal of 5'-phosphate groups

1U of calf intestinal phosphatase (CIP) was added to the

restriction digest reactions when digestion was complete. After incubation at 37°C for 30min, CIP was inactivated by incubation at 70°C for 15min. DNA was extracted with Ø/$CHCl_3$ and ethanol precipitation.

### 2.9.3 Isolation of DNA fragments from agarose gels

#### 2.9.3.1 Electroelution

The DNA fragment to be isolated was visualised by long-wave UV illumination, excised using a scalpel and transferred to a piece of dialysis tubing containing approximately 1ml of 0.5xTBE. The dialysis tubing was placed in a BRL gel electrophoresis kit perpendicular to the current flow. Electrophoresis was performed at 200V in 0.5xTBE for 1h. Successful elution was determined visually by long-wave UV illumination. Electrophoresis at 200V for 20 to 30s with reverse polarity was performed to release DNA from the side of the dialysis tubing. DNA was removed from the dialysis tubing and mixed with 0.05 vol 1.0M MOPS (pH7.0) and 0.1 vol 5.0M NaCl. DNA was purified using a QIAGEN-tip 5 (Diagen). The QIAGEN-tip 5 was equilibrated by repeatedly pipetting 300µl buffer A (400mM NaCl, 50mM MOPS, 15% EtOH pH7.0) in and out. A syringe was used to force the DNA sample through at a flow rate of ~250µl/min the QIAGEN-tip 5 was washed by pipetting 750µl buffer B (750mM NaCl, 50mM MOPS, 15% EtOH pH7.0) in and out five times using fresh buffer each time. The DNA was finally eluted by pipetting 200µl buffer F (1500mM NaCl, 50mM MOPS, 15% EtOH pH7.5) in and out three times. This process was repeated twice and the eluates pooled. DNA was precipitated using 0.8 vols isopropanol.

#### 2.9.3.2 Low melting point agarose

DNA was excised as described above from a 1% agarose gel made from SeaPlaque low melting point agarose. The gel piece was transferred to a 1.5ml eppendorf tube, covered with TE and incubated at 65°C for 10min. Ø/TE, prewarmed to 65°C, was used to extract the DNA twice. A third extraction was performed using Ø/TE equilibrated to RT.

After Ø/$CHCl_3$ extraction, DNA was ethanol precipitated using glycogen (final concentration ~20µg/ml) as a carrier.

### 2.9.3.3 Use of NA45 ion-exchange membrane

One half of an NA45 ion-exchange membrane (Schleicher and Schuell, N° 417006) was pre-treated in accordance with the method of Dretzen *et al.* (1981). After a brief wash in $H_2O$, the filter was washed once in TEN (TE with NaCl at a final concentration of 1M), then four times in TE. A small slit was made in the gel with a single edged razor blade just in front of the band to be eluted. The pretreated filter was inserted into the slit and electrophoresis continued for 10 to 20min. The filter was removed and the presence of DNA bound to the filter was confirmed by brief long-wave UV illumination. DNA was eluted by incubation at 37°C for 1h in 250µl TEN. Buffer was removed to a fresh tube and a further 150µl TEN added to the membrane for incubation at 37°C for 1h. The two aliquots of TEN containing the DNA were pooled and successful elution confirmed by visualization with long-wave UV. The pooled DNA sample was passed through glass wool packed in a Gilson P1000 tip. To maximise recovery, the tip was placed into a 15ml disposable tube and centrifuged at 1,000rpm for 2min in a Beckman GPR tabletop centrifuge. The DNA was ethanol precipitated using glycogen as a carrier.

### 2.9.4 Conversion of 5'-overhangs to blunt ends

Protruding 5' overhangs, generated by restriction endonuclease treatment of plasmid DNA, were converted to blunt ends using the method described by Maniatis *et al.* (1982). 500 to 1000ng of completely digested DNA was Ø/$CHCl_3$ extracted, ethanol precipitated and resuspended in 10µl $H_2O$. To this was added 1µl 2mM dNTPs, 2.5µl 10xNTB, 2U of the Klenow fragment of DNA polymerase I and the volume adjusted to 25µl with $H_2O$. After incubation at 22°C for 30min, the enzyme was inactivated by incubation at 70°C for 5min. DNA was Ø/$CHCl_3$ extracted and excess dNTPs removed by two ethanol precipitations.

### 2.9.5 DNA ligation

Ligations were usually performed in a total volume of 10µl. 100ng vector and 100ng DNA fragment to be insert were mixed with 1µl 10xLB (Boehringer-Mannheim) or 2µl 5xLB (BRL), 1.0µL (1U) T4 DNA ligase and the volume adjusted to 10µl with $H_2O$. The ligation reaction was incubated at 16°C overnight. 1 to 3µl was used for transformation of *E.coli.*

### 2.9.6 Gapped duplex DNA (gdDNA) synthesis

gdDNA was synthesized essentially as described (Kramer & Fritz, 1987). pAD113 was digested with *Hin*dIII and *Nco*I and the largest DNA fragment isolated from a 1% agarose gel (see Section 2.9.3). Uracil enriched, single-stranded (U-ss) pAD113 was separated from M13R408 helper phage in the same way. The concentration of the purified fragments was estimated by running them against samples of known concentration on a 1% agarose gel. gdDNA was formed by mixing 500ng U-ssDNA with 1,250ng *Hin*dIII/*Nco*I fragment in 30µl buffer containing 187.5mM KCl, 12.5mM Tris-Cl pH7.4 The mixture was boiled at 100°C for 3min, incubated at 65°C for 8min and allowed to cool to room temperature over a period of 20 to 30min.

## 2.10 Oligonucleotides

### 2.10.1 Synthesis and purification

Oligonucleotides synthesized on the Biosearch 8600 DNA synthesizer were provided in 1ml of ammoniacal solution. This was heated at 55°C for 5h and subsequently lyophilized overnight. The pellet was resuspended in 50µl $H_2O$ and 30µl was stored at -20°C. The remaining 20µl was mixed with 20µl loading buffer without dyes (1xTBE, 90% formamide). The sample was boiled for 2 to 3min prior to gel electrophoresis through a 15% polyacrylamide gel (acrylamide: bisacrylamide 29:1) containing 8M urea. 4µl stop solution (see Section 2.13.1.1) was loaded in a separate well to provide tracking dyes. Electrophoresis was performed at 250V. For oligonucleoties of length 15 to 32 bases in

length, the BPB was allowed to migrate $^2/_3$ of the way through the gel. For oligonucleotides 42-57 bases in length the BPB was allowed to migrate to the end of the gel. To isolate the oligonucleotides, the gel was removed from the glass plates and wrapped in cling film. DNA was visualized by illumination of the gel, whilst on a thin layer chromatography (TLC) plate, using a shortwave UV lamp. The DNA appears as a dark band against a uniform fluorescent background on the TLC plate. The band was excised from the gel and eluted in 1mL $H_2O$ by overnight incubation at 37°C. The DNA solution was butan-2-ol extracted twice to reduce the volume prior to EtOH/NaOAc precipitation. The DNA pellet was resuspended in 50µl $H_2O$ and the concentration determined spectrophotometrically (1 $A_{260}$ unit = 20µg/ml oligonucleotide).

Oligonucleotides synthesized on the Applied Biosystems 394/4 4 column DNA synthesizer were provided in 200µl of ammoniacal solution. This was incubated at 55°C overnight. After 15s centrifugation the solution was decanted to a fresh tube. 0.1vol 3M NaOAc and 2.5vol ethanol were added and incubated at -70°C for 10min. After a 5min spin in a microcentrifuge, the supernatant discarded and the pellet was rinsed with 70% ethanol before lyophilization. The pellet was resuspended in 200µl TE, 160µl of which was stored as a crude extract at -20°C. The remaining solution was ethanol precipitated as before. The final pellet was resuspended in 50µl TE and the concentration determined spectrophotometrically.

As many of the oligonucleotides were subsequently phosphorlylated, ethanol precipitation was performed using NaOAc rather than $NH_4OAc$, as $NH_4^+$ ions can inhibit T4 polynucleotide kinase (Sambrook *et al.*, 1989).

### 2.10.2 **Synthesis of spiked oligonucleotides**

These oligos were synthesized on the Biosearch 8600 DNA synthesizer using phosphoramidite mixes containing 96% of the normal nucleotide and 1.33% of the other three nucleotides (see Section 3.1).

### 2.10.3 Radioactive 5'end labelling of oligonucleotides

100 to 200ng oligonucleotide was labelled with 5µl [gamma-$^{32}$P] dATP using 1U T4 polynucleotide kinase in 1xKB in a final volume of 25µl at 37°C for 30min. The volume was adjusted to 100µl with $H_2O$ and the oligonucleotide was ethanol precipitated twice, using glycogen as a carrier. The final pellet was resuspended in 20µl $H_2O$ and was immediately used as a probe for hybridization.

### 2.10.4 Oligonucleotide sequences

A complete list and sequences of the oligonucleotides used in this project are shown here.

#### 2.10.4.1 Site-directed mutagenesis oligos

| Number | Sequence | Mutation |
|---|---|---|
| 52 | TCAAACTGCT**T**ACGAAGCATG | S288L(VZV TK) |
| 57 | GAGCTTGCATT**G**CAAATAAAG | G48A |
| 58 | GTATCGAGCTTG**GGGGGG**T**G**CAAATAAAG | G48A/M49P/49P50 |
| 68 | CCACACA**TT**ATACGAGCCG | promoter |
| 69 | AATCCATTA**TTACT**CCATTTCC | L301Δ/A300E |
| 70 | CTTGTTTTAAGT**G**CTGAAGTTCGC | Y21H |
| 72 | GAGCTGCAG**C**ACCTGGGGCAG | Y33H (Human TS) |
| AR1 | AATCCATTA**TTA**AGCCATTTCC | L301Δ |
| AR2 | AGCCATT**GG**CATTTTTAGGGG | E298P |
| 67 | GTTAAAAGAGGAAACGTTTCATTTCGC | 58T59 |

Nucleotides in **bold** type-face represent substitutions, whilst nucleotides in **bold underlined** type-face represent insertions. The mutations refer to the amino acid changes the oligo was designed to introduced and apply to the VZV TS gene, unless otherwise stated. S288L represents the substitution of Ser at position 288 by Leu. 48P49 represents the insertion of Pro between amino acids 48 and 49.

2.10.4.2 **Sequencing oligos**

| Number | Sequence | |
|---|---|---|
| 3 | TATTCTGAAATGAGC | |
| 4 | GGAATCGGAACGTTA | |
| 5 | ATAGGAATGGCTTCC | |
| 6 | GGATATCCCCTTAAT | |
| 7 | GGGGATGCACATATT | |
| 53 | TGGGATTGAAGACAC | (VZV TK) |
| 59 | GTCTGTGTATATCAT | |
| 60 | CCATTGTATGAATTA | |
| 61 | TTTAATTGTATCTAT | |
| 62 | CCACGACGGCCCTCC | |
| 74 | CATCTCTCAGGCTGTAGCGCG | (Human TS) |
| 56 | CAATGTTGAACGGTAC | |

These oligos were used as sequencing oligos of the VZV TS gene unless otherwise indicated.

2.10.4.3 **Spiked oligos**

| Number | Sequence Targeted | Size |
|---|---|---|
| R01 | 109 to 165 | (57) |
| R02 | 160 to 213 | (54) |
| R03 | 208 to 249 | (42) |
| R04 | 241 to 297 | (57) |
| R05 | 466 to 513 | (48) |
| R06 | 499 to 546 | (48) |
| R07 | 535 to 579 | (45) |
| R08 | 574 to 621 | (48) |
| R09 | 613 to 666 | (54) |
| R10 | 655 to 708 | (54) |
| R11 | 700 to 753 | (54) |

The numbering system refers to the VZV TS gene, with nucleotide +1 referring to the A in the ATG codon specifying the first methionine residue (see also Fig. 3.4).

## 2.11 Mutagenesis

### 2.11.1 Chemical mutagenesis

#### 2.11.1.1 Ethyl methanesulphonate

30ml LB was inoculated with 1ml *E.coli* GL271 ($\chi$2913 transformed with pGL271) overnight culture and incubated at 37°C, with vigorous shaking until $A_{600}$ reached ~0.6 ($4x10^8$ cells/ml). 10ml was pelleted by centrifugation at 9,000rpm for 5min at 4°C in a Sorvall SS34 rotor and resuspended in 5ml M9 medium. The cells were repelleted and then resuspended in a fresh 5ml aliquot of M9 medium. 50µl EMS (100%) was added, mixed by vortexing and incubated at 37°C for 15min. The cells were repelleted as before and resuspended in 5ml LB. Twelve 100µl aliquots were used to inoculate twelve 10ml LB overnight cultures. Mini-prep DNA was prepared from each of the overnight cultures. The mutagenized DNA was used to transform *E.coli* $\chi$2913.

#### 2.11.1.2 Hydroxylamine

3 to 5µg plasmid DNA, in 100µL buffer containing 0.4M hydroxylamine, 50mM sodium phosphate (pH6.0), and 1mM EDTA, was incubated at 75°C for 30 or 60 min. A spin-X tube (Costar) was loaded with 1 to 2ml of Sephadex G25 slurry in TE containing 0.02% $NaN_3$ to form a desalting column. The column was washed four times with TE. DNA was loaded onto the column and forced through by centrifugation for 20 to 30s in a microcentrifuge. DNA was ethanol precipitated, dried and resuspended in 10 to 20µl $H_2O$. 1 to 2µl was used to transform *E.coli* $\chi$2913 or MC1061*thyA*.

#### 2.11.1.3 Sodium bisulphite

The following method for mutagenesis was based on those of Pine and Huang (1987). 2.5 to 25µg ssDNA or 2 to 5µg gdDNA were mutagenized in 1ml buffer containing 3M sodium bisulphite (mixture of $Na_2SO_3$/$NaHSO_3$), 0.5mM hydroquinone, 3.75M NaCl, 0.375mM tri-sodium citrate by incubation at 37°C, in the dark, for 1 to 4h. The reaction was terminated by the following sequence of dialysis steps:

1 1l 5mM potassium phosphate (pH6.0), 0.5mM hydroquinone at 4°C for 2h.
2 As step 1.
3 1l 5mM potassium phosphate (pH6.0) at 4°C for 2h.
4 1l 0.2M Tris-Cl(pH9.2), 50mM NaCl, 2mM EDTA at 37°C for 16 to 24h.
5 1l 2mM Tris-Cl(pH8.0), 2mM NaCl, 2mM EDTA at 4°C for 6 to 12 h.

DNA was recovered from the dialysis tubing and ethanol precipitated using glycogen as a carrier. 1 to 2µg gdDNA was made completely double stranded in 10µL 1xLB (BRL) containing 2U Klenow fragment of DNA polymerase I, 1U T4 DNA ligase, 1 to 2µg T4 gene 32 protein and 0.5mM dNTPs incubated at room temperature for 2h. 1 to 5µl was used to transform *E.coli*. ssDNA was made completely double stranded using phosphorylated oligo 3 as a primer for second strand synthesis (see section 2.11.3). Alternatively, ssDNA was made completely double-stranded *in vivo* by being used to transform *E.coli* BW313, and mini-prep dsDNA extracted from these bacteria. 1 to 5µl dsDNA was used to transform *E.coli* XL-1*thyA* or SØ928*thyA*.

### 2.11.2 **Growth in a mutator strain**

50µl of an overnight culture of *E.coli* CB158 containing pGL271 was used to inoculate either 10ml LB supplemented with ampicillin, 10ml M9 media supplemented with ampicllin or 10ml M9 media supplemented with ampicillin and FUdR (5µM), and incubated overnight at 37°C with vigorous shaking. 50µl of each overnight culture was used to inoculate 10ml of the fresh corresponding media and were incubated overnight at 37°C with vigorous shaking. The process was continued until 5 growth 'cycles' had been completed. DNA was extracted from each of the three final cultures and used to transform *E.coli*.

### 2.11.3 **Oligonucleotide directed mutagenesis**

Oligonucleotides were used to introduce defined mutations at specific sites or mutations at random across a

short targeted region essentially as described by Mead *et al.* (1986), with slight modifications (Nick Davis-Poynter, personal communication). The strand selection scheme of Künkel was employed in which a uracil-containing template strand is degraded in an *ung*$^{+}$ host strain (Künkel, 1985; Künkel *et al.*, 1987). 1nmol oligonucleotide was phosphorylated in 20µl of 1xKB containing 1U T4 polynucleotide kinase at 37°C for 30min. T4 polynucleotide kinase was inactivated by incubation at 70°C for 15min. The phosphorylated oligonucleotide was diluted in $H_2O$ to a final concentration of 5pmol/µl and stored at -20°C. 1µl phosphorylated oligonucleotide was annealed to 2µl (0.1pmol) U-ssDNA (see below) in 10µl buffer containing 10mM Tris-Cl pH8.0 and 10mM $MgCl_2$ at 37°C for 30min, or at 55°C for 5min and cooled slowly to room temperature. To generate dsDNA, 1µl 5mM dNTP mix (a 5mM solution of all four dNTPs), 2µL 10xLB, 5U Klenow fragment of DNA polymerase I, 1U T4 DNA ligase, 2.0µg T4 gene 32 protein and $H_2O$ to a final volume of 20µL. After 2h incubation at RT, a further 5U Klenow fragment of DNA polymerase I and 1U T4 DNA ligase were added and incubated at RT for another 2h. The DNA was stored at -20°C or 1 to 5µl used to transform *E.coli*.

U-ssDNA was made in exactly the same manner as ssDNA (see Section 2.8.4) except DNA was propagated in either *E.coli* BW313, RZ1032 or CJ236.

### 2.11.4 Misincorporation mutagenesis

The following method was based upon that of Lehtovaara *et al.* (1988) with modifications based upon the protocol of Shiraishi and Shimura (1988). 5pmol phosphorylated oligonucleotide was annealed to 2.5pmol U-ssDNA in 25µl 15mM Tris-Cl pH8.0, 7.5mM $MgCl_2$ by incubation at 37°C for 30 min. 155µl 15mM Tris-Cl pH8.0, 7.5mM $MgCl_2$ was added and the sample divided into twelve 15µl aliquots. These were sub-divided into four groups of three and referred to as $A^-$, $C^-$, $G^-$ and $T^-$ ($N^-$ collectively). 3.0µl of the appropriate $N^-$ reaction mix (see below) was added to each aliquot. 2.0µl of one of three concentrations of the

appropriate limiting nucleotide was added (0.75, 2.0 and 3.25µM). Each reaction was initiated by the addition of 2U Klenow fragment of DNA polymerase I, incubated at room temperature for 5min, and terminated by the addition of 1.5µl 0.25M EDTA. The three tubes from each group were pooled, Ø/$CHCl_3$ extracted and ethanol precipitated three times (to remove free nucleotides) in the presence of 2.5M $NH_4OAc$ using glycogen as a (20µg/ml) carrier. The final pellet was washed with 75% ethanol and lyophilized.

Two methods were used for the misincorporation of nucleotides and subsequent complete second strand synthesis and ligation. The first method was based on that of Lehetovaara (1988). Each of the four DNA samples ($A^-$, $C^-$, $G^-$ and $T^-$), was resuspended in 10µL $dH_2O$ and 6.3µL 5xRTB, 12.4µl of the appropriate misincorporation mix (misinc N mix) and 10U RT were added. Samples were incubated at 37°C for 1.5h. A further 10U reverse RT were added and incubation continued for 1.5h. 3.5µl 5mM dNTPs were added to all four tubes and incubated at 37°C for 30min. 15µl 5x RTB, 4µl (10U) Klenow fragment of DNA polymerase I, 14µl (14U) T4 DNA ligase, 2µl (7µg) T4 gene 32 protein and 7µl 5mM ATP were added to each tube. After mixing, the contents were divided between two tubes and incubated at 14°C overnight. DNA was stored at -20°C. 1 to 5µl DNA was used to transform bacteria.

The $N^-$ reaction mixes used were as follows:

$A^-$ reaction mix 3.3mM dNTPs except dATP

$C^-$ reaction mix 3.3mM dNTPs except dCTP

$G^-$ reaction mix 3.3mM dNTPs except dGTP

$T^-$ reaction mix 3.3mM dNTPs except dTTP

The misincorporation mixes (misinc-N-mix) used were as follows:

| | | | |
|---|---|---|---|
| Misinc-A-mix | 75µM dCTP, | 0.4µM dGTP, | 200µM dTTP |
| Misinc-C-mix | 200µM dATP, | 50µM dGTP, | 1µM dTTP |
| Misinc-G-mix | 50µM dATP, | 250µM dCTP, | 100µM dTTP |
| Misinc-T-mix | 50µM dATP, | 100µM dCTP, | 200µM dGTP |

The second method was based upon that of Shiraishi and Shimura (1988). Each of the four samples was resuspended

in 20µl $H_2O$ and divided into four 5µl aliquots, one each for the misincorporation of the four different α-thio dNTPs. Misincorporation was performed in 20µl of 130mM *N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulphonic acid (HEPES) pH7.6, 0.2mM $MnCl_2$, 2mM ß-mercaptoethanol (ß-ME), 0.1mg/ml BSA, 0.1mM α-thio dNTP and Klenow fragment of DNA polymerase I (1U) at 25°C overnight. 6.0µl 5xLB, 2.5µl (5U) Klenow, 1.0µl T4 DNA ligase, 0.5µl (2µg) T4 gene 32 protein and 1µl 5mM dNTPs were added and incubated for 2h at RT. A further 5U Klenow fragment of DNA polymerase I and 1U T4 DNA ligase were added, and the incubation continued for a further 2h. DNA was stored at -20°C. 1 to 5µl DNA was used to transform bacteria.

## 2.12 Mutant screening

### 2.12.1 Growth complementation assays

#### 2.12.1.1 Thymidylate synthase

1 to 5µl mutagenized DNA was used to transform a suitable *E.coli thyA* strain. Transformants were picked, using sterile tooth picks, and patched onto M9 plates, and M9 plates supplemented with thymine (50µg/ml) or TdR (100µg/ml) and incubated at 37°C overnight. Appropriate clones were used to set up overnight cultures. Mini-prep DNA was prepared and used to retransform a suitable *E.coli thyA* strain. The phenotype was confirmed by streaking for single colonies on minimal plates and minimal plates supplemented with TdR and incubating them overnight at 37°C. $Thy^+$ clones were sequenced directly. $Thy^-$ clones were analysed by restriction enzyme analysis prior to DNA sequencing. In both cases the entire TS coding region was sequenced for each clone.

#### 2.12.1.2 Thymidine kinase

1 to 5µl mutagenised pAD507 DNA was used to transform *E.coli* C600*tdk* cells. Transformants were patched onto M9 plates supplemented with trimethoprim (10µg/ml), TdR (100µg/ml), and AdR (100µg/ml) (MTTA plates) and M9 plates

supplemented with Amp. Clones growing on both types of plates have TK activity whilst clones that grew only on minimal plates were predicted to lack TK activity. Mini-prep DNA from both types of clone was prepared and used to transform fresh *E.coli* C600*tdk* cells. The phenotypes were reaffirmed and the relevant portion of the TK gene sequenced.

### 2.12.2 Drug resistance screening protocols

#### 2.12.2.1 Method 1

A 100µL aliquot of *E.coli* χ2913, transformed with mutagenized DNA, was used to inoculate 3ml LB-Amp and incubated overnight at 37°C. A second 100µl aliquot was spread onto LB-Amp plates, and incubated overnight at 37°C. Transformants were replica plated onto M9 plates supplemented with 7.5µM FU (FU7.5 plates) and incubated for 16 to 40h at 37°C. 100µl of the overnight culture was spread onto FU7.5 plates and incubated for a further 16 to 40h. Single colonies growing on either plate were streaked onto fresh FU7.5 plates, with *E.coli* GL271 streaked as a negative control, and incubated as before. The TS gene of clones growing on these plates was sequenced.

#### 2.12.2.2 Method 2

*E.coli* MC1061*thyA* was transformed with mutagenized and wt DNA. Colonies were replica plated from LB-Amp plates onto M9 plates supplemented with 4.0µM FUdR (FUdR4 plates) and incubated as described above. Single colonies were streaked onto fresh FUdR4 plates and incubated for a further 16 to 40h. Mini-prep DNA was extracted from clones growing on these plates and used to transform fresh *E.coli* MC1061*thyA*. Transformants were replica plated as above. Clones that did not grow were discarded. Clones that did grow (i.e. showing plasmid linked drug resistance) were streaked on fresh FUdR4 plates and incubated for 16 to 40 h. The TS gene of clones growing on these plates was sequenced.

2.12.2.3 **Method 3**

Mutagenized DNA was used to transform the *E.coli* strain SØ928*thyA*. 3ml LB-Amp, was inoculated with 100µl transformed cells and incubated, with vigorous shaking, at 37°C overnight. 100 to 200µl overnight culture was spread onto FUdR5, 10 or 20 plates and incubated at 37°C for 16 to 40h. Single colonies were streaked onto FUdR1, 2 or 5 plates and incubated as before. Mini-prep DNA was made from clones that yielded single colonies on FUdR plates, and used to transform fresh *E.coli* SØ928*thyA* cells. Transformants were streaked onto FUdR1, 2 or 5 plates and incubated as before. The coding region of the TS gene of clones growing on these plates was completely sequenced. By way of variation, bacteria were grown overnight in minimal media (liquid or solid) at all stages immediately prior to growth in the presence of FUdR.

2.12.2.4 **Method 4**

Mutagenized DNA was transformed into *E.coli* SØ928*thyA* harbouring the VZV TK expressing plasmid, pCBTK3, and treated essentially as described in method 3 except that the overnight growth steps prior to growth in the presence of drug were always performed. 5-ethynyl-2'-deoxyuridine (EYUdR) was used at 1, 5 and 10µM for initial screening and 0.5, 1 and 2µM when colonies were streaked for single colonies on fresh plates. FUdR was used as in method 3.

2.12.2.5 **Method 5**

Screening was performed as in method 4, but after the initial overnight growth step in M9 media, the bacteria were grown in minimal media, supplemented with either FUdR or EYUdR, through several growth cycles prior to the initial plating out onto minimal media plates containing the drugs. Screening was continued as for method 4.

2.12.3 **Slot blot hybridization**

Up to 72 overnight cultures, including control cultures, were set up. Slot blot hybridization was performed using a

micro-sample filtration manifold kit (SRC 072/0 Minifold II; Schleicher and Schuell). Two gel blotting papers (Schleicher and Schuell) and a piece of GeneScreen *Plus*™ membrane (DuPont) were soaked in hybridization buffer (20mM Tris-Cl pH7.5, 3xSSC, 5xDHB, 0.5% SDS, 20µg/ml denatured salmon testes DNA; Maniatis *et al.*, 1982) for 5 to 10min. The kit was completely assembled and connected to a vacuum pump. 100µl overnight culture was loaded into each of the 72 wells and the vacuum continued for 5mins after the lást well was loaded. The membrane was removed from the kit and washed in 0.5M NaOH for 5 min. After a second wash in 1.0M Tris-Cl pH7.5 for 2 min, a final wash in 0.5M Tris-Cl, 1.5M NaCl for 5 to 10min was performed. The membrane was placed in a polythene bag filled with 20 to 30ml hybridization buffer and heat sealed. Prehybridization was performed at 65°C for 2h. After prehybridization, the membrane was transferred to a 50ml Falcon tube containing 20ml hyridization buffer. 20µl $^{32}P$ labelled oligonucleotide (see Section 2.10.3) was added and hybridization performed at room temperature for 2h with continuous rotation. The membrane was washed with 6xSSC at room temperature, placed in a polythene bag and heat sealed before overnight exposure to Kodak S film 100. If the desired clone could not be differentiated from control clones, a series of washes at increasing temperatures (usually in increments of 5°C) were performed until the desired clones were identified.

## 2.13 **Mutant Characterisation**

### 2.13. **DNA sequencing**

#### 2.13.1.1 **Sequencing solutions**

USB sequencing solutions:

Labelling mix: 1.5µM dCTP, 1.5µM dGTP, 1.5µM dTTP

Termination mix: A contained 8µM ddATP
C contained 8µM ddCTP
G contained 8µM ddGTP
T contained 8µM ddTTP

All four termination mixes also included 80µMdATP, 80µMdCTP, 80µM dGTP, 80µM dTTP, 50mM NaCl.

Pharmacia sequencing solutions:
Labelling mix: 2.0µM dCTP, 2.0µM dGTP, 2.0µM dTTP
Termination mix: A contained 15µM ddATP
C contained 15µM ddCTP
G contained 15µM ddGTP
T contained 15µM ddTTP

All four termination mixes also included 150µMdATP, 150µMdCTP, 150µM dGTP, 150µM dTTP, 50mM NaCl, 40mM Tris-Cl pH7.5, 10mM $MgCl_2$.

Common solutions:
Annealing buffer (5xTM): 0.1M Tris-Cl pH8.0, 0.1M $MgCl_2$.
Stop solution: 95% deionized formamide, 20mM EDTA pH7.5, 0.05% xylene cyanol FF, 0.05% bromophenol blue.
Wacker solution (~5ml): 25µl Wacker silane GF38, 150µl 10% acetic acid, 5ml EtOH.

2.13.1.2 **Sequencing reactions**

DNA dideoxy-sequencing was performed according to the method of Sanger *et al.* (1977) using either modified T7 DNA polymerase (Sequenase, Tabor and Richardson, 1987) or unmodified T7 DNA polymerase. Sequencing was performed essentially as described by the manufacturers, with minor modifications. 5 to 10ng of oligonucleotide were annealed to 0.5 to 1.0µg denatured plasmid template (see Section 2.8.5) or ssDNA (see Section 2.8.4) in 10µl 1xTM at 37°C for 30min. The labelling reaction was initiated by addition of 1µl 0.1M DTT, 2µl labelling mix, 2µl [$\alpha$-$^{35}S$] or 0.5µl [$\alpha$-$^{32}P$] dATP and 3U Sequenase were added to the annealing mix and incubated at room temperature for 5min. Alternatively 5µl Sequetide and 3U Sequenase were added to the annealing mix and processed in an identical manner. For each sequencing reaction four wells of a 96 microtitre plate were labelled G, A, T and C. 2.5µl of the

corresponding dideoxy-termination mix was added to the wells. 3.5µl labelling reaction was transferred to the side wall of each well and the plate covered with a titertek plate sealer (77-400-05 Flow Laboratories Ltd.) prior to a brief spin to 1,500rpm in an International Equipment Company Centra-4X centrifuge to mix the components. After incubation at 37°C for 5 to 30min, 4µl stop solution was added to each well. Samples were boiled for 2 to 3 min in a water bath. 2 to 3µl was loaded onto a sequencing gel (see below). When unmodified T7 DNA polymerase was used, the reactions were performed as above, except that the solutions decribed by the manufacturer were employed.

2.13.1.3 **Polyacrylamide gel electrophoresis**

Sequencing reactions were was resolved on a 6.0% acrylamide, 7.7M urea, 0.5 to 2.5x TBE gradient gel (Hong, 1987). Prior to use, both plates were washed with 100% EtOH. If the gel was to be transferred to 3MM paper (see below), both plates were coated with Repelcote™, then washed once with 100% EtOH, before the gel was poured. When the gel was to be oven dried, the front plate was treated as above, but the second plate was coated with ~5mL of Wacker solution, then washed five times with 100% EtOH, before the gel was poured. Electrophoresis was performed in 1xTBE at 60W. For normal length runs the BPB was used as a marker to terminate electrophoresis. In longer runs, XCFF was used as a marker to terminate electrophoresis. The sequencing gel was either transferred to 3MM paper and dried on a vacuum gel drier for ~90 min. Alternatively, it was soaked in 10% acetic acid for 30 min before drying onto the back plate in an oven at 140°C for 60 to 90 min. The gel was exposed to Kodak S 100 film for 16 to 72h at room temperature with an intensifying screen.

2.13.2 **TS extracts**

2.13.2.1 **Extract preparation**

100ml LB supplemented with TdR (50µg/ml), was inoculated

with 3ml overnight culture and incubated at 37°C, with vigorous shaking, until $OD_{600}$ ~0.6. 1ml cells were removed and stored on ice for a phenotype check (see below). The remaining cells were pelleted in a Sorvall GS3 rotor at 7,000rpm for 5min at 4°C. The supernatant was discarded and the pellet resuspended in 10ml ice-cold TES (50mM Tris-Cl pH7.4, 5mM EDTA, 50mM NaCl). The suspension was transferred to a 15mL Falcon tube and the cells repelleted at 3,000 rpm for 10 min at 4°C in a Beckman GPR benchtop centrifuge. The supernatant was discarded, the pellet volume estimated, frozen and thawed once, and resuspended in 2 vol sonication buffer (50mM Tris-Cl pH7.5, 10mM DTT, 0.1% Triton X-100). The sample was sonicated in a Dawe sonicator with micro-soniprobe for three 10s periods at ~150W on ice with 30s cooling between each period. The sonicate was transfered to a 1.5mL eppendorf tube and centrifuged in a Sorvall SS34 rotor, with suitable adaptors, at 10,000 rpm for 10 min at 4°C. The supernatant was recovered by aspiration and snap-frozen in a dry-ice/ethanol bath before storage at -70°C. For the phenotype check, the cells were streaked for single colonies on an LB plate supplemented with TdR and incubated at 37°C overnight. Single colonies were patched onto LB-TdR-Amp, M9-TdR and M9 plates and incubated at 37°C overnight.

2.13.2.2 **Protein concentration determination**

Protein concentrations were determined by the method of Bradford (1976) using the Bio-Rad protein assay kit. 10µl diluted extract, protein standard or $H_2O$ was mixed with 990µl Bio-Rad dye reagent (filtered through Whatman No.1 paper before use) and after 5 to 10min at room temperature, absorbance measured at 595nm. Each extract or protein standard was assayed in duplicate. BSA was used as a protein standard. A standard curve was made using concentrations of 0, 0.01, 0.02, 0.05, 0.1, 0.2 and 0.4mg/ml. Protein extracts were usually diluted 1:5 with $H_2O$ to give readings that could be readily determined from

the standard curve. Alternative dilutions were used when necessary.

### 2.13.3 TS assays

#### 2.13.3.1 Tritium release assay

This assay follows the release of radioactivity from 5-[$^3$H]dUMP to the solvent. Unreacted substrate is absorbed to charcoal, removed by centrifugation and the radioactivity in the supernatant determined (Roberts, 1966). TS extract was mixed with 20µl 2xassay mix (128mM NaF, 260mM Tris-Cl pH7.5, 64mM sucrose, 0.3%($^w/_v$) BSA, 0.25mM DTT, 38mM HCHO and 0.2mM tetrahydrofolate), 0.4µCi 5-[$^3$H]dUMP and $H_2O$ in a total volume of 40µl and incubated at 30°C for 15min. The reaction was terminated by addition of 15µl 33% TCA, 5µl 5mg/ml dUMP and 180µl charcoal suspension (2% charcoal in 0.1M $H_2O$. The sample was vortexed for 10s. After centrifugation for 2min, 500µl supernatant was removed and added to 3ml scintillation fluid (Ecoscint A) prior to liquid scintillation counting in a Beckman LS5000CE scintillation counter. A total counts control was included by using 5µL sonication buffer instead of extract and adding 940µL $H_2O$ instead of charcoal suspension. A measure of backgound counts was included by replacing extract with 5µL sonication buffer and this was subtracted from all experimental values. Assays were performed in duplicate on two or more independant extracts for each clone. The amount of extract used was adjusted so that less than 30% of the total counts are released. This ensures that assays are in the linear region of the reaction curve.

The amount of dTMP generated is directly equivalent to the amount of [$^3$H] released during the assay. The amount of 5-[$^3$H]dUMP added to the reaction in pmol can be calculated from its specific activity. The total counts represents the equivalent amount of product (dTMP).

i.e

$$\text{pmol dTMP produced} = \frac{\text{Sample CPM}}{\text{Total CPM}} \times \text{pmol 5-[}^3\text{H]dUMP added}$$

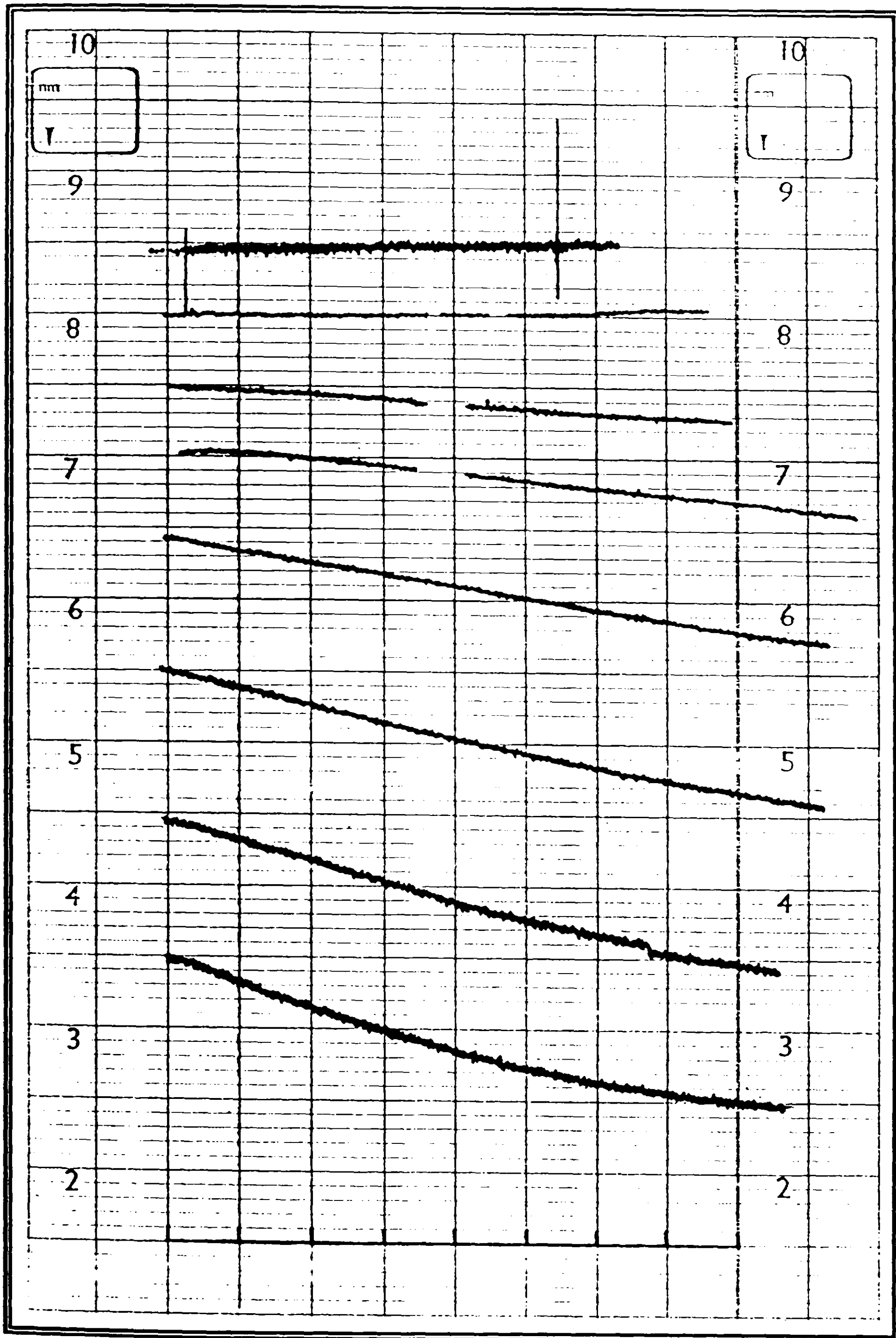



Fig. 2.1. **Debromination assay**.
Debromination assay reaction traces. Chart speed was 120s/cm, recorder range was 1.0. The reactions contain, in order from top to bottom of the figure: 5μl χ2913 extract (12mg/ml); zero extract; 0.5, 1.0, 2.0, 3.0, 4.0, 5.0μl χ2913/pAD768 extract (12mg/ml). Note that the reaction rapidly becomes non-linear at higher extract concentrations.

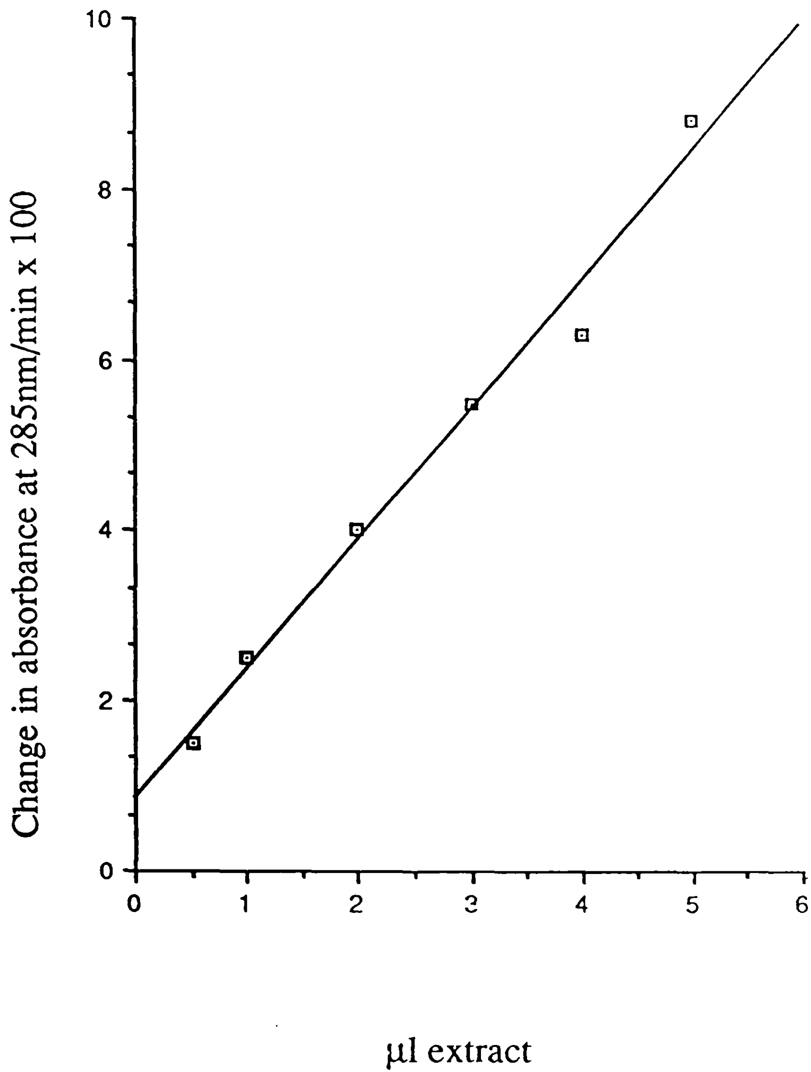



Fig. 2.2 **Debromination assay**.
The linearity of the debromination assay is shown on this graph of intial rate of absorbance change against amount of extract. Data points are taken from Fig. 2.1.

From the protein concentration, the specific activity of TS in the crude extracts can be expressed in terms of pmol dTMP produced per µg protein per minute.

2.13.3.2 **Debromination assay**

This assay follows the reduction in absorbance at 285nm caused by release of bromine from 5-BrdUMP (Garret *et al.*, 1979). It represents the first stage of the TS reaction, and as such does not require the presence of the folate co-factor. The reaction buffer was 200mM N-methylmorpholine hydrochloride pH7.4, 25mM $MgCl_2$, 1mM EDTA, 75mM ß-ME prepared at 2x strength and stored at 4°C. Reactions were run in Hellma quartz microcells containing 250µL 2x buffer, 5µl 4mM 5-BrdUMP (40µM final concentration), 1 to 5µl of crude extract and $H_2O$ in a final volume of 500µl. Absorbance change was monitored using a Cecil CE595 double beam spectrophotometer (zeroed on the cuvettes before adding extract) and displayed on a chart recorder. Reactions were run for 15 to 20min at room temperature (usually 25 to 30°C), and an extract containing WT VZV TS was included in each run. The linearity of the assay under these conditions was first confirmed. Fig. 2.1 shows traces from reactions with increasing amounts of active extract. Fig. 2.2 shows that the initial rates of absorbance change are proportional to the amount of extract added. To compare extracts initial rates were calculated and expressed as mOD/min/ug protein.

2.13.4 **Characterization of potential FUdR resistant TS variants**

To characterize variants that could not form single colonies on solid M9 media, but which otherwise showed evidence of drug resistance, the following assay was devised. The variants were expressed in the *E.coli* strain SØ928*thyA*/pCBTK3 and grown overnight in 10ml of M9-Amp-Cm medium at 37°C. This was used to inoculate 200ml of the same medium. The new culture was incubated at 37°C until the cells reached mid-log phase ($A_{600}$ = 0.4). At this

point, the culture was divided equally between two flasks, and FUdR added to a final concentration of 50 or 100μM (see Section 3.1). The growth of the bacteria was monitored at 30min intervals, with $OD_{600}$ readings taken at each time point. The final growth curves are the results of two independent experiments (see Figs. 3.8 and 3.9).

### 2.13.5 TK extracts

1ml overnight culture was pelleted by a 15s spin in a mini-centrifuge. The supernatant was discarded and the cells resuspended in 100μl TK lysis buffer (20mM Tris-Cl pH7.5, 2mM $MgCl_2$, 10mM NaCl, 6.5mM ß-ME, 0.5% Nonidet P 40). 5μL lysozyme (10mg/ml) was added and the cells incubated on ice for 10min. Three cycles of freeze /thawing (-70°C, 5min / 37°C, 5min) were performed. The lysed cells were centrifuged for 5min and the supernatant, which represented the TK extract removed and stored at -70°C. Protein concentrations were determined as above.

### 2.13.6 TK assays

10μl TK extract was mixed with 40μl TK assay mix (20mM sodium phosphate pH6.0, 1mM $MgCl_2$, 0.25mM ATP, 166μl/ml [methyl-$^3$H]TdR and incubated at 37°C for 60min. 10μl TK stop solution (10mM EDTA, 2mM TdR) was added and the solution boiled for 3min before cooling on ice for 3min. After a 3min spin, 50μl supernatant was spotted onto a folded DE81 disc (Whatman) and allowed to dry at RT for 2 to 3min. The discs were washed three times, 10min each time, in 4mM ammonium formate, 10μM TdR at 37°C. After two 1min washes in 100% ethanol, the discs were dried between two pieces of filter paper. The discs were placed in 3ml scintillation fluid and counted in a liquid scintillation counter as above. Assays were performed with 3 discs per extract, and two independant extracts were made for each clone.

## 2.14 Computing and Molecular Modelling

All DNA and protein sequence analysis was performed using the Sequence Analysis Software Package of the Genetics Computer Group of the University of Wisconsin run on the VAX/VMS computer in the Institute of Virology, University of Glasgow.

The computer generated graphics of the VZV TS structures were constructed using the UCSD Molecular Modelling System of run on the Silicon Graphics IRIS computer in the Department of Chemistry, University of Glasgow. The modelling was performed as described (see Sections 3.1 to 3.3). Photographs were taken directly from the computer screen at 50% brightness, on Kodacolour Gold 100 (daylight) film, with exposures of 1/30, 1/15 and 1/8s at f2.

# 3. RESULTS AND DISCUSSION

The Results and Discussion chapter is divided into four sections. The first three detail studies of VZV TS, whilst the final section deals with VZV TK. The studies on TS commence with a description of the development of protocols to isolate drug resistant variants of TS, continue with a section on the molecular modelling of a number of variants that have a Thy⁻ phenotype, and is completed by site-directed mutagenesis studies of two key structural regions of the enzyme. The work on TK includes site-directed mutagenesis and describes the identification of the potential evolutionary source of the herpesvirus TKs.

## 3.1 Drug resistance studies

The aim of the drug resistance studies was to isolate a mutant variant (or variants) of TS that was resistant to the effects of various inhibitory drugs. The definition of drug (or inhibitor) resistance in this case is that such a variant should be able to catalyse the reductive methylation of deoxyuridylate to form thymidylate as normal in the presence of a concentration of the active form of the inhibitor that prevents the wild type enzyme from catalysing this reaction. The isolation and characterisation of such a mutant would reveal information about the role that various amino acids play in catalysis. This information may enhance the understanding of the mechanisms of drug resistance observed during the treatment of various cancers where TS inhibitors are widely used (see Section 1.12). Finally, it may be useful in evaluating the role of TS in VZV infections and establish the use of TS as a target enzyme for antiviral chemotherapy.

Neither of the two reports of FU resistance in tumours, where a lesion in the TS locus has been demonstrated, have described the amino acid changes in the protein that are responsible for drug resistance (Heidelberger *et al.*, 1960a,b; Jastreboff *et al.*, 1983). The scarcity of the

| Mutagen | Screening protocol | No. of drug$^R$ clones isolated | No. of drug$^R$ clones sequenced |
|---|---|---|---|
| Ethyl methanesulphonate | 1 | 12 | 12 |
| Hydroxylamine | 1 | 2 | 2 |
| Hydroxylamine | 2 | 2 | 2 |
| Sodium bisulphite | 2 | 0 | 0 |
| *mut*D | 2 | 14 | 7 |

Table 3.1. **Initial Drug Resistance Screening Results.**
Number of drug resistant clones isolated and sequenced from initial experiments.

reports, only two in thirty years of FU use as a drug (Heidelberger, 1957), suggests that only a limited number of mutations in the TS gene may give rise to resistance. The isolation of a resistant variant of VZV TS by *in vitro* studies would thus require both highly efficient and efficacious mutagenesis and selection procedures.

In order to isolate drug resistant variants, it was decided to study VZV TS in an *E.coli* expression system which had previously been shown to express active enzyme (Thompson *et al.*, 1987). The aim was to generate a library of TS variants by a suitable random mutagenesis protocol which could be screened for drug resistance. The amino acid substitutions responsible for the drug resistant phenotype would then be determined by DNA sequencing, and the enzyme studied in more detail to establish the mechanism of resistance.

The results presented in this section describe the selection and optimization of suitable mutagenesis protocols and the refinement of a genetic selection system. We then go on to describe the use of these methods to screen a large number of clones for a drug resistant variant of VZV TS.

### 3.1.1 Initial experiments

Four methods of random mutagenesis were studied for application to wide scale screening for drug resistant variants of VZV TS. Several drug resistant clones were isolated from DNA randomly mutagenized by three methods (EMS, HA and the *E.coli* *mut*D strain CB158; see Sections 2.11.1 and 2.11.2). The results of these initial experiments are summarized in Table 3.1. The nature of screening protocols 1 and 2 (in that they were preliminary methods) meant that an accurate number of the clones screened could not be readily determined. From approximately 1,000 clones generated by each mutagenesis protocol, screening resulted in the isolation of 30 drug resistant variants. Twenty-one of these drug resistant clones were analysed by DNA sequencing and no nucleotide

changes were discovered in the TS gene of any isolate.

To assess the efficiency of the mutagenesis protocols, 200 clones generated from each mutagenesis method were assayed for growth complementation. 137 clones generated by bisulphite mutagenesis had a Thy⁻ phenotype. When 19 of these clones were analysed by *Eco*RI restriction enzyme digestion, 18 had different digestion patterns to that of the WT TS plasmid. This suggested that bisulphite was inducing excessive damage to the DNA rather than inducing point mutations. The TS gene of the single clone that had the same digest pattern was sequenced with oligos 3 to 7 and found to have no nucleotide changes. None of the other methods produced clones with a Thy⁻ phenotype.

These results suggested that the mutagenesis protocols were inefficient as shown by the fact that so few Thy⁻ clones were isolated, and that those that were isolated had no nucleotide changes in the TS gene. The drug screening methods were not suitable as judged by the isolation of a high number of 'false positive' clones, clones that were drug resistant, but had no nucleotide changes in the TS gene suggesting the presence of mutations at other loci.

On the basis of these experiments, two new methods of random mutagenesis were employed, together with a radically different method of screening for drug resistant clones. These new methods were first evaluated for efficiency and efficacy, and subsequently employed to screen large numbers of clones.

### 3.1.2 New drug screening protocols

To overcome the isolation of the high number of false positives clones obtained when screening for TS variants resistant to FU, the following protocols were adopted. They were devised to allow the use of FUdR, and the drug EYUdR (see Section 2.12.2).

#### 3.1.2.1 The use of FUdR rather than FU

Our analysis of the pathways for nucleotide metabolism in *E.coli*, together with previous studies in this area (see

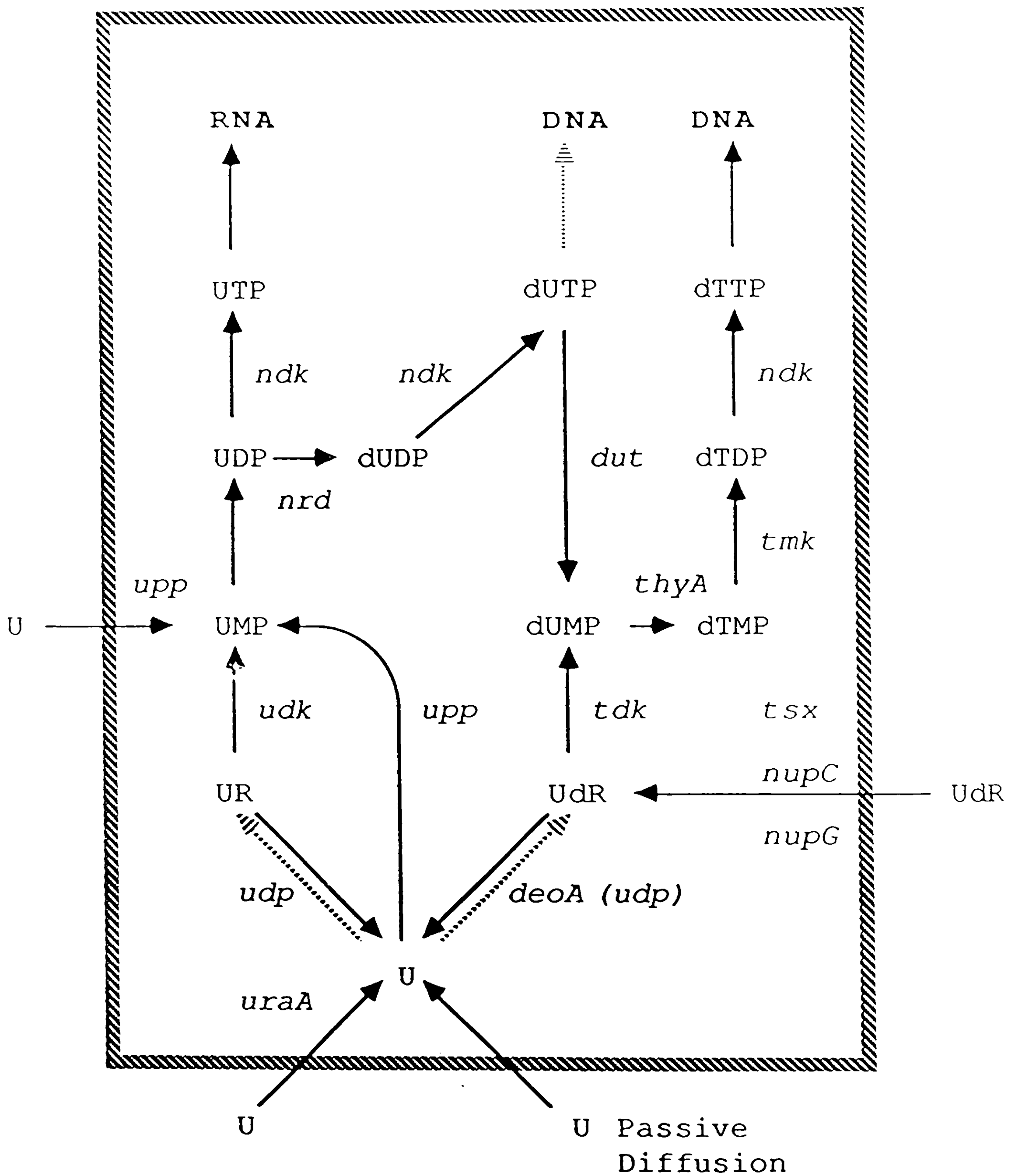



Fig. 3.1. **Nucleotide Metabolism.**
The pathways of *E.coli* nucleotide metabolism relevant to the drug screening protocols. The structural and biochemical similarity of uracil and FU means that most pathways for their metabolism are the same (see text for details).
The following abbreviations are used for the various nucleotides described in this scheme: U - uracil; UR, UMP, UDP and UTP - uridine and its mono-, di- and triphosphate derivatives. UdR, dUMP, dUDP and dUTP - deoxyuridine and its mono-, di- and triphosphate derivatives. dTMP, dTDP and dTTP - the mono-, di- and triphosphate derivatives of thymidine. The following abbreviations are used for the genes of various proteins: *uraA*, *nupC*, *nupG* and *tsx* - genes for proteins involved in nucleobase and nucleoside uptake. *udp* - uridine phosphorylase; *deoA* - thymidine phosphorylase; *upp* - uracil phosphoribosyl transferase; *udk* - uridine kinase; *tdk* - thymidine kinase; *tmk* - thymidylate kinase; *thyA* - thymidylate synthase; *nrd* - ribonucleotide reductase; *dut* - dUTPase; *ndk* - nucleoside diphosphate kinase.

Section 1.12), suggested that the use of FU was not suitable for the isolation of drug resistant variants of TS, and may contribute to the isolation of false positive clones. This is due to the fact that FU can kill bacterial cells in a number of ways (see Section 1.13.2). To inhibit TS, FU must be converted to FdUMP. As FU metabolism closely parallels that of uracil (Neuhard & Nygaard, 1987; O'Donovan & Neuhard, 1970), FU would not be expected to be converted to FUdR, an immediate precursor of FdUMP, to a great extent. This means that FdUMP is predominantly formed via the action of ribonucleotide reductase (see Fig. 3.1. and Section 1.13.2). Mutational events that affect any of the enzymes in this activation pathway would prevent the formation of FdUMP, and could result in the isolation of false positive clones.

FU can also be incorporated into RNA, which can result in cell death (see Section 1.13.2). If the incorporation of FU into RNA is the major cause of cell death, then it would be impossible to isolate a drug resistant variant of TS, as the enzyme is not the main target for the drug. In addition, FU can kill certain strains of E.coli by the inhibition of normal cell wall synthesis (see Section 1.13.2). This would also prevent the isolation of a drug resistant variant of TS. For these reasons, it was decided to use FUdR, as opposed to FU, for subsequent experiments.

#### 3.1.2.2 A *thyA* derivative of *E.coli* strain SØ928

Once inside the cell, FUdR can be phosphorylated to form FdUMP, in a reaction catalysed by TK, or degraded to form FU, by the action of thymidine and/or uridine phosphorylase (see Fig. 3.1). Obviously the use of an *E.coli* strain that could catalyse the breakdown of FUdR to form FU, would create the same problems as using FU directly. For this reason, an *E.coli* strain was sought that could not catalyse this reaction.

The ideal strain would have deletion mutations in the genes *deoA* and *udp*, which encode thymidine phosphorylase (TdRPase) and uridine phosphorylase (UdRPase) respectively

(see Fig. 3.1). Such a strain was not available, however the *E.coli* strain SØ928 was a more than adequate replacement. This strain has the *deoA* gene deleted and is also deficient in UdRPase activity (Valentin-Hansen *et al.*, 1978). The lack of these enzyme activities, and the fact that the metabolism of FU and uracil are very similar in *E.coli*, means that this strain cannot catalyse the formation of FU from FUdR (see Fig. 3.1; Neuhard & Nygaard, 1987).

As the breakdown of FUdR to FU is prevented in *E.coli* SØ928, FUdR can only kill the cells by the inhibition of TS and concomitant induction of thymineless death (see Section 1.13.2). In this case, FUdR would be converted to FdUMP in a single step; the phosphorylation of FUdR being catalysed by TK. Such a strain could then only become resistant to FUdR in one of two ways. The first, and the desired option, would be due to mutations in the TS gene. The second would be due to mutations in the TK gene that would result in inefficient phosphorylation of FUdR.

The *E.coli* strain SØ928 also lacks uracilphosphoribosyl transferase activity (Valentin-Hansen, *et al.*, 1978). If the cells regained UdRPase activity as the result of a spontaneous mutation, this gives the added advantage of preventing the formation of FUMP from FU, and the subsequent incorporation of FU into RNA (see Fig. 3.1 and Section 3.1.2.1).

To allow the VZV TS to be the target of drugs used in the drug screening protocols (rather than the *E.coli* TS), a *thyA* variant of the *E.coli* strain SØ928 was selected as described (see Section 2.7.2).

The new screening protocols facilitated an accurate measure of the number of clones that were being screened (see Section 2.12.2.3). The *E.coli* strain SØ928*thyA* transformed with either pAD403 or pAD768 was used to establish the concentration of FUdR to use for drug screening protocol 3. 2.5μM FUdR was found to inhibit single colony formation of cells expressing VZV TS encoded by the plamid pAD403, whereas a concentration of 4.0μM FUdR

was required when VZV TS was encoded by the plasmid pAD768. 10.0μM TFT (trifluorothymidine) was found to inhibit single colony formation of cells expressing VZV **TS** encoded by the plamid pAD403.

3.1.2.3 **Co-expression of VZV TK in SØ928*thyA***

The strategy of co-expressing VZV TK on a multi-copy plasmid in the same cells as VZV TS was chosen to reduce the possibility of false positive clones arising due to mutation of the *E.coli tdk* gene and to ensure that nucleoside phosphorylation would mimic the situation in VZV infected cells.

The active VZV TK gene was subcloned from the plasmid pVZVTK+5 (see Section 3.4.1) into the plasmid pCB105 to form pCBTK3. pCB105 contains the chloramphenicol acetyl transferase gene and can be selected for by using chloramphenicol. This vector is also compatible with the pBR322-based TS expression vectors and they can therefore stably co-exist in the same cell (Sherratt & Boyd, 1986). pCBTK3 was used to transform SØ928*thyA* and selected for using chloramphenicol.

The transformed cells, referred to as SØ928*thyA*/pCBTK3, were then transformed with pAD768 and selected for by using chloramphenicol and ampicillin. These transformants were referred to as SØ928*thyA*/pCBTK3/pAD768 and were used to establish the concentration of FUdR and EYUdR to use for drug screening protocols 4 and 5. An FUdR concentration of 4μM was sufficient to inhibit growth (irrespective of the presence of pCBTK3), whereas the concentration of EYUdR required to inhibit the formation of single colonies on solid media was 0.5μM. In the absence of pCBTK3, the concentration required to inhibit the growth is greater than 10μM. This confirms the need for the co-expression of VZV TK to phosphorylate this compound. This observation also suggested that *E.coli* TK differed in its substrate specificity to the herpesvirus TKs, and may resemble the cellular TKs in this respect (see Section 3.4.2). Recent sequence information pertaining to the *E.coli* TK gene has

strengthened this argument (Bockamp *et al.*, 1990; Black & Hruby, 1991; R. Thompson, J. E. Scott & P. T. Harrison, unpublished data).

#### 3.1.2.4 Drug evaluation

The initial studies to isolate a drug resistant variant employed FUdR. As this compound is a close structural analogue of dUMP (see Section 1.14.1), and so few instances of FUdR resistance have been reported, it was decided to screen for a drug resistant variant with an additional TS inhibitor, TFT.

During the course of these studies, we gained access to several compounds that have been shown to inhibit the growth of VZV (see Section 1.4.3). These compounds were evaluated as potential TS inhibitors. This was achieved by determining the concentration of each drug to prevent single colony formation of SØ928*thyA*/pCBTK3/pAD768. BVdU, ACV and AraA, all recognised as non-TS inhibitors, allowed single colony formation at 10µM and were not tested further. PYaraU and EYaraU, also known non-TS inhibitors did not inhibit growth at any of the concentrations tested. PYUdR, a weak TS inhibitor allowed growth at 20µM and was not tested further, however, EYUdR, a potent TS inhibitor, prevented single colony formation at 0.5µM. On this basis, EYUdR was selected as a replacement for TFT in the search for a drug resistant variant of VZV TS.

### 3.1.3 New mutagenesis protocols

Two new methods were employed for random mutagenesis, both requiring the use of vectors producing ssDNA (Geider, 1986) and the mutant enrichment procedure of Künkel (Künkel, 1985; Künkel *et al.*, 1987). The first method involves misincorporation mutagenesis (Lehtovaara *et al.*, 1988; Holm *et al.*, 1990; Shiraishi & Shimura, 1988), whilst the other method is based on 'spiked' oligonucleotide mutagenesis (McNeil & Smith, 1985; Ner *et al.*, 1988). Before use in screening for drug resistant clones, the efficiency of both methods was evaluated and optimized.

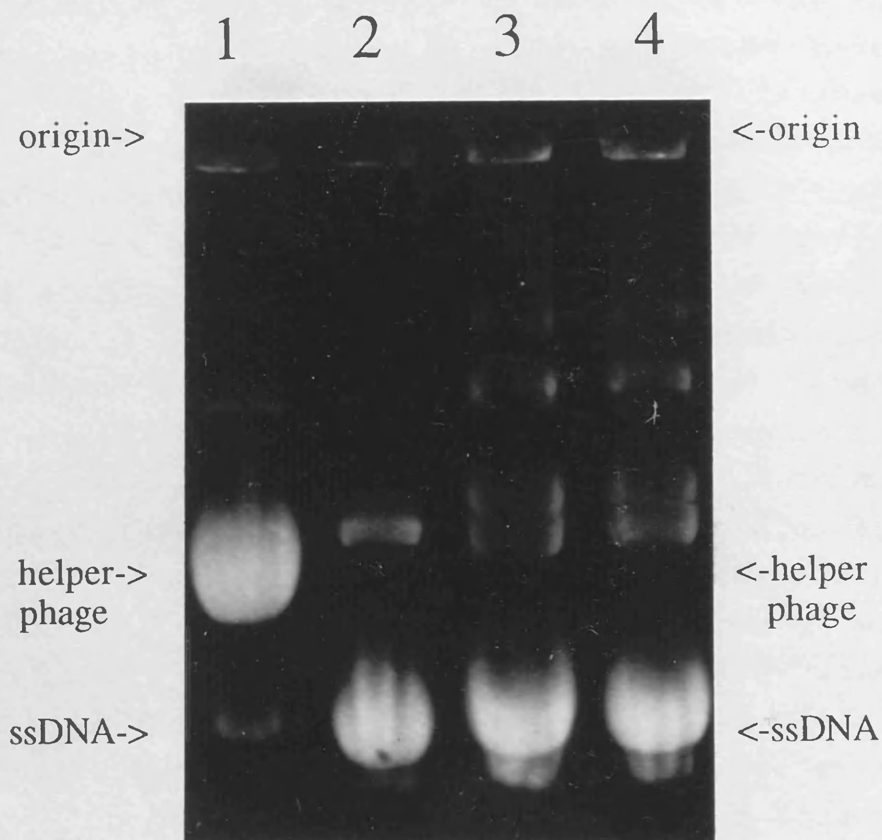



Fig. 3.2 **ssDNA**
This figure shows the effects of recloning of the VZV TS gene in the opposite orientation. Lane 1 shows helper phage DNA (top) and ssDNA (bottom) generated from the vector pAD113. Lanes 2, 3, and 4 shows helper phage DNA (top) and ssDNA (bottom) generated from the vectors pAD401, pAD403 and pAD404. pAD403 was used regularly for the production of ssDNA.

### 3.1.3.1 **ssDNA and plasmid constructs**

Previous attempts to clone VZV TS into M13 vectors have resulted in instability problems (R. Thompson, personal communication). To counter this, we decided to clone VZV TS into various plasmid vectors that contain the bacteriophage f1 origin of replication (Geider, 1986; Mead *et al.*, 1986).

TS was subcloned from pGL271 into pBS+ to construct the plasmid pAD113. The region of ssDNA containing the TS gene could be sequenced using oligos 3 to 7 because of the orientation of the TS gene, relative to the f1 origin of replication. However, ssDNA produced from this clone had a very low ratio of ssDNA to helper phage.

The low levels of ssDNA and high background of helper phage were thought to be causing problems with the synthesis of the second strand during mutagenesis (see below). In an attempt to improve the levels of ssDNA, and increase the ratio of ssDNA to helper phage, a variety of experiments were performed based on the observations of other groups (Geider, 1986; Mead *et al.*, 1986; Russel *et al.*, 1986; Vieira & Messing, 1987). These experiments included the propagation of pAD113 in different strains of *E.coli*, use of different helper phage, superinfection with helper phage at different stages of growth and the use of different multiplicities of infection. None of these methods increased the yields of ssDNA, or improved the ratio of ssDNA to helper phage.

TS was cloned into the vector pGEM3Zf+ in the same relative orientation to the f1 origin of replication as pAD113 to construct the plasmid pAD203, but the yields and ratio of ssDNA to helper phage were still low. The use of a strong termination signal downstream of the coding region has previously been shown to facilitate the cloning and expression of various foreign genes in *E.coli* (Gentz *et al.*, 1981). To see if the termination signal could improve yields of ssDNA, the transcriptional termination sequences of the *rrn*B operon (Gentz *et al.*, 1981; Brosius *et al.*, 1981; Amann & Brosius, 1985) were cloned downstream of the

TS gene to form the clone pAD301. No improvements in the yield of ssDNA or the ratio of ssDNA to helper phage were observed.

The construction of the plasmid pAD403, by subcloning the TS gene from pAD113 into pBS-, in the opposite orientation to the f1 origin of replication relative to pAD113, gave a dramatic increase in the ratio of ssDNA to helper phage. With pAD113 the ratio was approximately 1:5 compared to 3:1 for ssDNA production from pAD403 (see Fig. 3.2). Careful analysis revealed that two point mutations (A-48G; A-47T; numbering such that the A of the ATG codon of the initiating Met of TS is +1) had arisen in the -10 region of the promoter region linked to the TS gene during the construction of pAD403. These changes were corrected by site-directed mutagenesis using oligo 68, and the new clone was designated pAD768. The ratio and yields of ssDNA from pAD768 was identical to that observed for pAD403.

3.1.3.2 **Strand selection procedure**

The strand selection procedure for mutant enrichment (Künkel, 1985; Künkel *et al.*, 1987) was used for both random and site directed mutagenesis. The method uses an ssDNA template that contains a number of uracil residues in place of thymine (U-ssDNA; Künkel, 1985). U-ssDNA is produced in an *E.coli dut ung* strain. Such strains lack the enzymes dUTPase and uracil N-glycosylase. The lack of dUTPase activity leads to raised levels of dUTP (see Fig. 3.1) and results in increased incorporation of uracil into DNA in place of thymine. The lack of uracil N-glycosylase, which normally removes uracil from DNA, means that the uracil is stably incorporated. The U-ssDNA molecule can then be used as a normal template for the *in vitro* mutagenesis reactions as the coding potential of uracil is the same as that of thymine (Shlomai & Kornberg, 1978). The completely double stranded molecule, the product of the *in vitro* mutagenesis reaction, has a parental strand that contains both uracil and thymine residues, whilst the newly

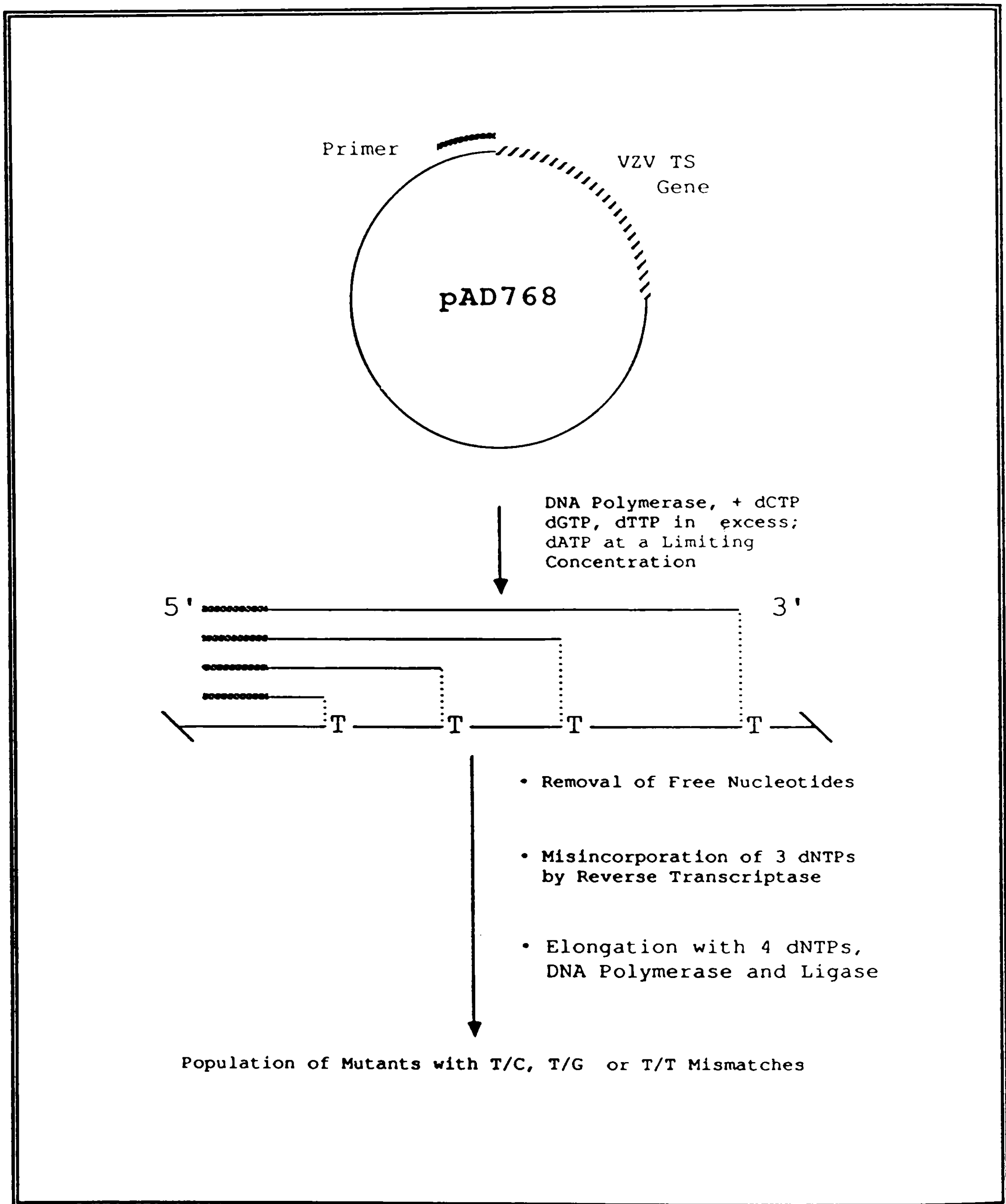



Fig. 3.3. **Misincorporation Mutagenesis.**
The principle of misincorporation mutagenesis is shown for an elongation reaction where dATP is the limiting nucleotide. See text for details.

synthesized mutant strand contains only thymine residues. When wild type *E.coli* cells (i.e. $Ung^{+}$) are transformed with this DNA, there is a strong *in vivo* biological selection against the replication of the parental template strand (Künkel *et al.*, 1987) that results in the production of predominantly mutant dsDNA molecules. The high efficiency of mutant production (typically >50%) allows the screening of site-directed mutants by DNA sequencing or the generation of highly mutated DNA libraries by random mutagenesis.

The highest levels of U-ssDNA was obtained from vectors in *E.coli* BW313 cells (compared to *E.coli* RZ1032 and CJ236). The strand selection efficiency was confirmed in that ssDNA (normal thymine containing) was capable of transforming both *E.coli* BW313 ($Ung^{-}$) and JM109 ($Ung^{+}$), whereas U-ssDNA could transform *E.coli* BW313, but not JM109.

### 3.1.4 **Misincorporation mutagenesis**

The principle of misincorporation mutagenesis is illustrated in Fig. 3.3 (Lehtovaara *et al.*, 1988; Holm *et al.*, 1990). A phosphorylated oligonucleotide is annealed upstream of the region of the gene targeted for mutagenesis. A so-called limited elongation reaction is performed with three dNTPs in excess and the fourth at a limiting concentration. This generates all possible 3'-ends of the elongated primer over the desired target region. Four separate reactions are performed for each primer, each of which generates a population of molecules terminating just before a given type of base, due to limiting amounts of nucleotides (these four populations are analogous to those synthesized in dideoxy-sequencing, except that the next base to be incorporated is known and the 3'-terminii can be further elongated in the following misincorporation reaction).

Point mutations are introduced to the four molecular populations by enzymatic misincorporation of the three mismatching nucleotides under conditions where proof-

reading does not occur. This reaction can be catalysed with reverse transcriptase (Lehtovaara *et al.*, 1988) or with the Klenow fragment of DNA polymerase I if α-thio dNTPs are used (Shiraishi & Shimura, 1988). The final stage is the complete synthesis of the second strand of the molecule.

3.1.4.1 **Use of the Klenow fragment of DNA polymerase I**

Oligo 56 was used as the primer to initiate the complete misincorporation mutagenesis reaction, using α-thio dATP as the limiting nucleotide, the Klenow fragment of DNA polymerase I to catalyse the misincorporation step and U-ssDNA made from pAD403 as described. DNA generated from this experiment was used to transform *E.coli* XL-1*thyA* cells, and the transformants screened by the TS growth complementation assay. Of 434 transformants analysed, only 13 non-complementing mutants were isolated. Analysis by *Eco*RI restriction enzyme digestion revealed that only five had an identical restriction digest pattern to that of pAD403. The TS gene of these clones was sequenced using oligos 59 to 62, but only two clones contained nucleotide changes in the TS gene. One clone had a single nucleotide change (C305T) that corresponded to the substitution of Ser102 by Leu (S102L), whilst a second clone had several changes; the A516G nucleotide substitution caused no amino acid changes, whilst the insertion of three C residues between nucleotides 1496 and 1497 resulted in two amino acid changes. Asp174 was replaced by Ala (D174A), and an additional amino acid, His, was inserted between amino acids 174 and 175 (174H175).

Despite the isolation of two Thy⁻ mutants with nucleotide changes in the TS gene, the efficiency of mutagenesis (estimated by the number of mutations in the TS coding region that resulted in a Thy⁻ phenotype) was only 0.46% (2/434). This suggests that the true mutagenesis efficiency (the number of clones containing mutations in the TS gene irrespective of phenotype) is also very low.

| Clone | Nucleotide Changes | Codon Changes | Amino Acid Changes | Phenotype |
|---|---|---|---|---|
| A | T536A | G**T**A to G**A**A | V179E | Thy$^-$ |
| B | T565C | **T**TT to **C**TT | F189L | Thy$^-$ |
| F | T559G | **T**GT to **G**TT | C187G | Thy$^-$ |
| M | T589A<br>T592C | **T**CC to **A**CC<br>**T**GC to **C**GC | S197T<br>C198R | Thy$^-$ |
| N | No Changes | No Changes | No Changes | Thy$^+$ |
| O | T627A | CT**T** to CT**A** | L209 Silent | Thy$^+$ |
| P | T468A | AT**T** to AT**G** | I156 Silent | Thy$^+$ |
| Q | T627G | CT**T** to CT**G** | L209 Silent | Thy$^+$ |
| R | T573A | GT**T** to GT**A** | V191 Silent | Thy$^+$ |

Table 3.2. **TS Mutants.**
Nucleotide, amino acid and codon changes of mutant clones, and their phenotypes. The nucleotide change T536A indicates that T536 was substituted by A. The effect of the nucleotide substitution on the codon are shown in **bold**. The amino acid change V179E indicates that V179 was replaced by E.

3.1.4.2 **Use of reverse transcriptase**

As a consequence of the results shown above, it was decided to use U-ssDNA generated from the plasmids pAD403 and pAD768 as a template for mutagenesis and reverse transcriptase to catalyse the misincorporation reaction.

Misincorporation mutagenesis was performed using U-ssDNA made from pAD403, oligo 56 as the primer and dATP as the limiting nucleotide. Mutagenised DNA was used to transform *E.coli* SØ928*thyA* and the transformants were screened by the TS growth complementation assay. Of 150 clones screened, 20 (13.33%) had a Thy$^-$ phenotype. The TS gene of three of these clones (A, B and F) was sequenced using oligos 59 to 62. The nucleotide changes and corresponding amino acid substitutions are shown in Table 3.2. In addition, a further six clones (M to R), were selected at random from the transformants, prior to phenotypic screening, and sequenced. The phenotypes of these clones, the nucleotide changes they possess and the corresponding amino acid substitutions are also shown in Table 3.2. The mutations described for these clones are in accordance with the predicted pattern of misincorporation and are randomly distributed throughout the region of the TS gene targeted by oligo 56.

To further assess the efficiency of mutagenesis, a reaction was performed using U-ssDNA made from pAD768, oligos 59, 60 and 62 and dATP as the limiting nucleotide as described. Mutagenised DNA was used to transform *E.coli* SØ928*thyA* and of 260 transformants assayed by the TS growth complementation assay, 35 (13.46%) had a Thy$^-$ phenotype. Analysis of 12 of these clones by *Eco*RI restriction enzyme digestion revealed that all 12 had a pattern identical to that of pAD768.

In summary, the use of reverse transcriptase for the misincorporation mutagenesis protocol has been shown to be superior to the Klenow fragment of DNA polymerase I. A high level of transformation was established with DNA generated by this method with a misincorporation reaction initiated from oligos 56 and 59 to 62. The efficiency of

```
                                                                      -35
                                                                     |----|
TTGCGCCGACATCATAACGGTTCTGGCAAATATTCTGAAATGAGCTGTTGACAATTAATC
AACGCGGCTGTAGTATTGCCAAGACCGTTTATAAGACTTTACTCGACAACTGTTAATTAG

     -10                                                  RBS
    |----|                                               |--|
ATCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAACAGACC
TAGCCGAGCATATTACACACCTTAACACTCGCCTATTGTTAAAGTGTGTCCTTTGTCTGG
                                     |-------------------- HELIX A -----------------|
M  G  D  L  S  C  W  T  K  V  P  G  F  T  L  T  G  E  L  Q  Y  L  K  Q  V  D  D  I  L  R  Y  G  V  R
ATGGGAGACTTGTCATGTTGGACAAAGGTGCCGGGTTTTACGTTAACCGGCGAACTTCAGTACTTAAAACAAGTGGATGATATTTTAAGGTATGGAGTTCGG
TACCCTCTGAACAGTACAACCTGTTTCCACGGCCCAAAATGCAATTGGCCGCTTGAAGTCATGAATTTTGTTCACCTACTATAAAATTCCATACCTCAAGCC

                        |^^^^^^^^^^^^^^^ SHEET i ^^^^^^^^^^^^^^^|                           |--
K  R  D  R  T  G  I  G  T  L  S  L  F  G  M  Q  A  R  Y  N  L  R  N  E  F  P  L  L  T  T  K  R  V  F
AAACGCGATCGAACAGGAATCGGAACGTTATCTTTATTTGGAATGCAAGCTCGATACAATTTGCGAAATGAATTTCCTCTTTTAACTACAAAGCGTGTTTTT
TTTGCGCTAGCTTGTCCTTAGCCTTGCAATAGAAATAAACCTTACGTTCGAGCTATGTTAAACGCTTTACTTAAAGGAGAAAATTGATGTTTCGCACAAAAA
     ''''''''''''''''''''(R01 - 57-mer)''''''''''''''''''||||||,,,,,,,,,,,(R02 -54-mer),,,,,,,,,,,,,,

--------------- HELIX B ----------------|          |------- HELIX C ------|              |- HELIX D -
W  R  A  V  V  E  E  L  L  W  F  I  R  G  S  T  D  S  K  E  L  A  A  K  D  I  H  I  W  D  I  Y  G  S
TGGAGGGCCGTCGTGGAAGAGTTGTTATGGTTTATCCGCGGGTCAACCGATTCCAAAGAACTCGCCGCTAAAGATATACACATATGGGATATATACGGATCG
ACCTCCCGGCAGCACCTTCTCAACAATACCAAATAGGCGCCCAGTTGGCTAAGGTTTCTTGAGCGGCGATTTCTATATGTGTATACCCTATATATGCCTAGC
,,,,,,,,,,,,(R03 - 42-mer),,,,,,,,,,|||||||||''''''''''''''(R04 - 57-mer)'''''''''''''''''''''''

-----|                                                 |--- HELIX G ----|
S  K  F  L  N  R  N  G  F  H  K  R  H  T  G  D  L  G  P  I  Y  G  F  Q  W  R  H  F  G  A  E  Y  K  D
AGCAAATTTCTAAATAGGAATGGCTTCCATAAAAGACACACGGGGGACCTTGGCCCCATTTACGGCTTCCAGTGGAGACATTTTGGAGCGGAATATAAAGAC
TCGTTTAAAGATTTATCCTTACCGAAGGTATTTTCTGTGTGCCCCCTGGAACCGGGGTAAATGCCGAAGGTCACCTCTGTAAAACCTCGCCTTATATTTCTG

                                |------------ HELIX H -------------|                |^^^^ SHEET v ^^^|
C  Q  S  N  Y  L  Q  Q  G  I  D  Q  L  Q  T  V  I  D  T  I  K  T  N  P  E  S  R  R  M  I  I  S  S  W
TGTCAATCAAACTATTTACAGCAAGGAATCGATCAGCTGCAAACTGTTATAGATACAATTAAAACAAACCCAGAAAGCCGACGAATGATTATATCGTCTTGG
ACAGTTAGTTTGATAAATGTCGTTCCTTAGCTAGTCGACGTTTGACAATATCTATGTTAATTTTGTTTGGGTCTTTCGGCTGCTTACTAATATAGCAGAACC
                                                        ''''''''''''''''(R05 - 48-mer)'''''||||||||||||

  |------ HELIX I -------;           |^^^^^^ SHEET iv ^^^^^^|                  |^^^^^ SHEET iii ^^^^^^^
N  P  K  D  I  P  L  M  V  L  P  P  C  H  T  L  C  Q  F  Y  V  A  N  G  E  L  S  C  Q  V  Y  Q  R  S
AATCCAAAGGATATCCCCTTAATGGTACTACCTCCATGTCACACGTTATGTCAGTTTTACGTTGCAAACGGTGAATTATCCTGCCAAGTATACCAGAGATCG
TTAGGTTTCCTATAGGGGAATTACCATGATGGAGGTACAGTGTGCAATACAGTCAAAATGCAACGTTTGCCACTTAATAGGACGGTTCATATGGTCTCTAGC
|||,,,(R06 - 48-mer),,,,||||||||||||'''''''(R07 - 45-mer)''''''||||||,,,,,,,,,(R08 - 48-mer),,,,,,,,,,

^^|   |---------------------------- HELIX J -------------------------------|            |^^^^^^^^^^^
G  D  M  G  L  G  V  P  F  N  I  A  G  Y  A  L  L  T  Y  I  V  A  H  V  T  G  L  K  T  G  D  L  I  H
GGGGATATGGGCCTTGGGGTACCGTTCAACATTGCTGGATATGCACTTCTTACCTACATAGTAGCGCATGTTACAGGACTTAAAACCGGAGATTTAATTCAT
CCCCTATACCCGGAACCCCATGGCAAGTTGTAACGACCTATACGTGAAGAATGGATGTATCATCGCGTACAATGTCCTGAATTTTGGCCTCTAAATTAAGTA
|||||||||''''''''''(R09 - 54-mer)'''''''''||||||||||||,,,,,,,,(R10 - 54-mer),,,,,,,,,,,,|||||||||''''''

^^^^^^^^^ SHEET ii ^^^^^^^^^^|          |----- HELIX K -----|
T  M  G  D  A  H  I  Y  L  N  H  I  D  A  L  K  V  Q  L  A  R  S  P  K  P  F  P  C  L  K  I  I  R  N
ACAATGGGGGATGCACATATTTACTTGAATCATATAGATGCTTTAAAAGTGCAGCTAGCTCGATCCCCAAAACCTTTTCCTTGCCTTAAAATTATTCGAAAT
TGTTACCCCCTACGTGTATAAATGAACTTAGTATATCTACGAAATTTTCACGTCGATCGAGCTAGGGGTTTTGGAAAAGGAACGGAATTTTAATAAGCTTTA
''''''''''''(R11 - 54-mer)''''''''''''''

V  T  D  I  N  D  F  K  W  D  D  F  Q  L  D  G  Y  N  P  H  P  P  L  K  M  E  M  A  L  *
GTAACAGATATAAACGACTTTAAATGGGACGATTTTCAGCTTGATGGATATAATCCACACCCCCCCCTAAAAATGGAAATGGCTCTTTAA
CATTGTCTATATTTGCTGAAATTTACCCTGCTAAAAGTCGAACTACCTATATTAGGTGTGGGGGGGGATTTTTACCTTTACCGAGAAATT
```

Fig. 3.4. **Regions of VZV TS Covered by Spiked Oligos.** The regions of the VZV TS gene target by the spiked oligos and the relationship of these areas to the amino acid sequence and the various elements of predicted secondary structure of VZV TS targeted by spiked oligo mutagenesis. The oligo sequences are indicated in **bold** and areas of overlap by ***bold italics***. They are also underscored by commas (' or ,) with areas of overlap indicated by vertical lines (|). The various elements of secondary structure are also shown. The ribosomal binding site (RBS) and -35 and -10 boxes of the promoter are also indicated.

mutagenesis can be estimated to be at least 50%, that is at least every other clone generated will be expected to have a single nucleotide substitution, a level necessary for the wide scale screening for a drug resistant variant of VZV TS. Moreover, the wide and random spread of the mutants demonstrates the efficacy of this method. A more detailed discussion of the mutagenesis efficiency is given in Section 3.1.6.3).

### 3.1.5 Spiked oligonucleotide mutagenesis

Most methods of random mutagenesis, including the misincorporation method described above, can be susceptible to mutagenic 'hot spots', that is the clustering of mutations in certain regions of the gene, whilst other regions may be left unmutated (McPherson, 1991). The method of 'spiked' oligonucleotide mutagenesis avoids such hot spots and is probably the most truly random mutagenesis protocol available (Blacklow & Knowles, 1991). The procedure is essentially the same as with site-directed mutagenesis in that oligos are used to generate nucleotide changes, and the template DNA is prepared, and selected against, in exactly the same way. However, the oligonucleotides have been synthesized by a protocol that gives a defined probability of incorporating one of the three 'wrong' nucleotide bases at each position along the oligo. This is achieved by spiking each phosphoramidite mix with a known concentration of the other three mixes (Ner *et al.*, 1988; McNeil & Smith, 1985).

11 oligos (R01 to R11) were synthesized using nucleotide mixes spiked with 1.33% of the other three nucleotide mixes (see Section 3.1.7.1). The oligos were designed to cover most of the elements of secondary structure of VZV TS (see Fig. 3.4) and introduce an average of two nucleotide changes. The rationale of introducing two nucleotide changes per oligo is two-fold. First, it serves to increase the variation at individual codons. When one base of the codon is changed, an average of 5.7 (of the 19) other amino acids can be encoded (and these alterations

| Oligo | Nucleotide Changes | Codon Changes | Amino Acid Changes | Phenotype |
|---|---|---|---|---|
| R01 | A121C | **A**TC to **C**TC | I41L | Thy$^-$ |
| R02 | T162G | AA**T** to AA**G** | N54K | Thy$^-$ |
| R04 | A288G | AT**A** to AT**G** | I96M | Thy$^-$ |
| R07 | A546G<br>T549G<br>C562A | CC**A** to CC**G**<br>TG**T** to TG**G**<br>**C**AG to **A**AG | P182Silent<br>C183W<br>Q188K | Thy$^-$ |

Table 3.3. **Spiked Oligo Mutagenesis.**
Nucleotide, amino acid and codon changes of mutant clones, and their phenotypes. The nucleotide change A121C indicates that A121 was substituted by C. The affect of the nucleotide substitution on the codon are shown in **bold**. The amino acid change I41L indicates that I41 was replaced by L.

tend to be conservative), but when two bases are changed, the average number of other amino acids that can be accessed rises sharply to an average of 15.7 (Hermes *et al.*, 1990). Secondly it allows the potential for a number of double amino acid variants to be synthesized. A detailed discussion of the factors controlling the number of changes is given in Section 3.1.7.1.

Prior to use in conjunction with the drug screening protocols, the level of mutagenesis efficiency of this method had to be established. Oligos R01 through R11, and U-ssDNA made from pAD403 were used as described. Mutagenized DNA was used to transform *E.coli* SØ928*thyA*. Of 80 transformants assayed by the TS growth complementation assay, 14 (17.5%) had a Thy⁻ phenotype. Three of the Thy⁺ clones were selected at random and the region of the TS gene corresponding to the oligo used to introduce mutations was sequenced. Two had no detectable nucleotide changes whilst the clone generated from oligo R11 had a single silent nucleotide change (A702T), which did not change the Gly residue at position 234.

Analysis of the phenotype of a further 99 clones revealed that 25 (25.25%) had a Thy⁻ phenotype. The distribution of clones with a Thy⁻ phenotype was spread equally amongst the 11 oligos.

To see if the Thy⁻ phenotypes were the result of mutations in the TS gene or an incomplete filling-in reaction, eight clones were analysed by *Eco*RI restriction enzyme digestion. Five shared an identical restriction pattern to that of pAD403. The TS gene of four of these was sequenced using oligos 59 to 62. All four clones contained point mutations in a region that corresponded to the oligo used to generate that particular clone in the initial mutagenesis reaction (Table 3.3). Refinements to the mutagenesis protocol, that included the use gene 32 protein and U-ssDNA generated from pAD768 as a template for spiked oligo mutagenesis alleviated the problems of an incomplete filling-in reaction.

In summary, these results have demonstrated the efficacy

| Clone | Source | Growth on FUdR2 plates | Growth on TFT10 plates |
|---|---|---|---|
| G56 | FUdR2 plate | single colonies* (1) | single colonies* (2) |
| G56 | TFT10 plate | N.T. | single colonies* (3) |
| T56 | TFT10 plate | N.T. | no growth |
| pAD403 | FUdR2 plate | poor growth* (4) | N.T. |
| pAD403 | TFT10 plate | N.T. | poor growth* (5) |
| pAD403 | M9 plate | no growth | no growth |

Table 3.4. **Initial characterization of drug resistant clones.** Ability of various clones generated by misincorporation mutagenesis to grow in the presence of various inhibitors. An asterisk (*) represents that the clone was processed further, with numbers in parentheses representing the subsequent clone number. Poor growth indicates that single colonies were not formed, but some growth was observed. N.T. indicates not tested.

| Clone | Growth on FUdR plates 2.0μM | Growth on FUdR plates 0.4μM | Growth on TFT plates 10.0μM | Growth on TFT plates 2.0μM |
|---|---|---|---|---|
| G56-1 | poor growth | single colonies | poor growth | single colonies |
| G56-2 | poor growth | single colonies | poor growth | single colonies |
| G56-3 | poor growth | single colonies | poor growth | single colonies |
| T56-4 | no growth | no growth | no growth | no growth |
| T56-5 | no growth | no growth | no growth | no growth |
| pAD403 | no growth | no growth | no growth | no growth |

Table 3.5. **Further characterization of drug resistant clones.** Growth of various drug resistant clones after retransformation of *E.coli* SØ928*thyA* cells on solid media in the presence of various inhibitors. Poor growth indicates that single colonies were not formed, although some growth was observed.

and efficiency of mutagenesis (>50%) of this method. As with the misincorporation mutagenesis method, the distribution of the mutations is random, with both nucleotide transitions and transversions observed. The isolation of a mutant with three changes further demonstrates the efficacy of this method.

### 3.1.6 Screening clones for drug resistance generated by misincorporation mutagenesis

The new methods of mutagenesis have been shown to be highly efficient and reliable. The aim now was to assess their worth in conjunction with the new drug screening protocols.

#### 3.1.6.1 Success of drug selection scheme

The new selection scheme was shown to be successful by the isolation of a plasmid linked drug resistant clone with a single nucleotide change in the bacterial promoter region. 1,400 clones generated by misincorporation mutagenesis (using U-ssDNA made from pAD403 and oligo 56 as described) were screened for resistance to FUdR and TFT by method 3. When spread onto FUdR2 and TFT10 plates, clones derived from the reaction where dGTP was the limiting nucleotide for the elongation reaction (G56), grew as a lawn, whereas all other clones and a pAD403 control culture failed to grow or grew to a significantly lesser extent than G56. EYUdR was not available during the initial experiments. Bacteria taken from this lawn were streaked for single colonies on FUdR2 and TFT10 plates (Table 3.4). Five clones (indicated with an asterisk*) were tested for a plasmid linked drug resistant phenotype, but only the three clones derived from the G56 reaction showed plasmid linked drug resistance (see Table 3.5). The TS gene of one of these clones (G56-1) was sequenced using oligos 59 to 62. Complete sequencing of the TS coding region revealed no nucleotide changes. Analysis of the upstream sequences revealed a single nucleotide change, A to C at position -53 (numbering is such that +1 is A of the ATG methionine

| | |
|---|---|
| Sequence of -10 region of pAD403 | TCGTATGTTG |
| Sequence of -10 region of Drug[R] G56-3 clone | TAGTATGTTG |

Fig. 3.5. **Promoter changes.**
Comparison of the sequences at and surrounding the -10 region of the bacterial promoter of pAD403 and the drug resistant clone G56-3. Nucleotide substitution is shown in **bold**. -10 region is underlined. Note that pAD403 contains a promoter mutation. The consensus sequence for the -10 region is TATAAT.

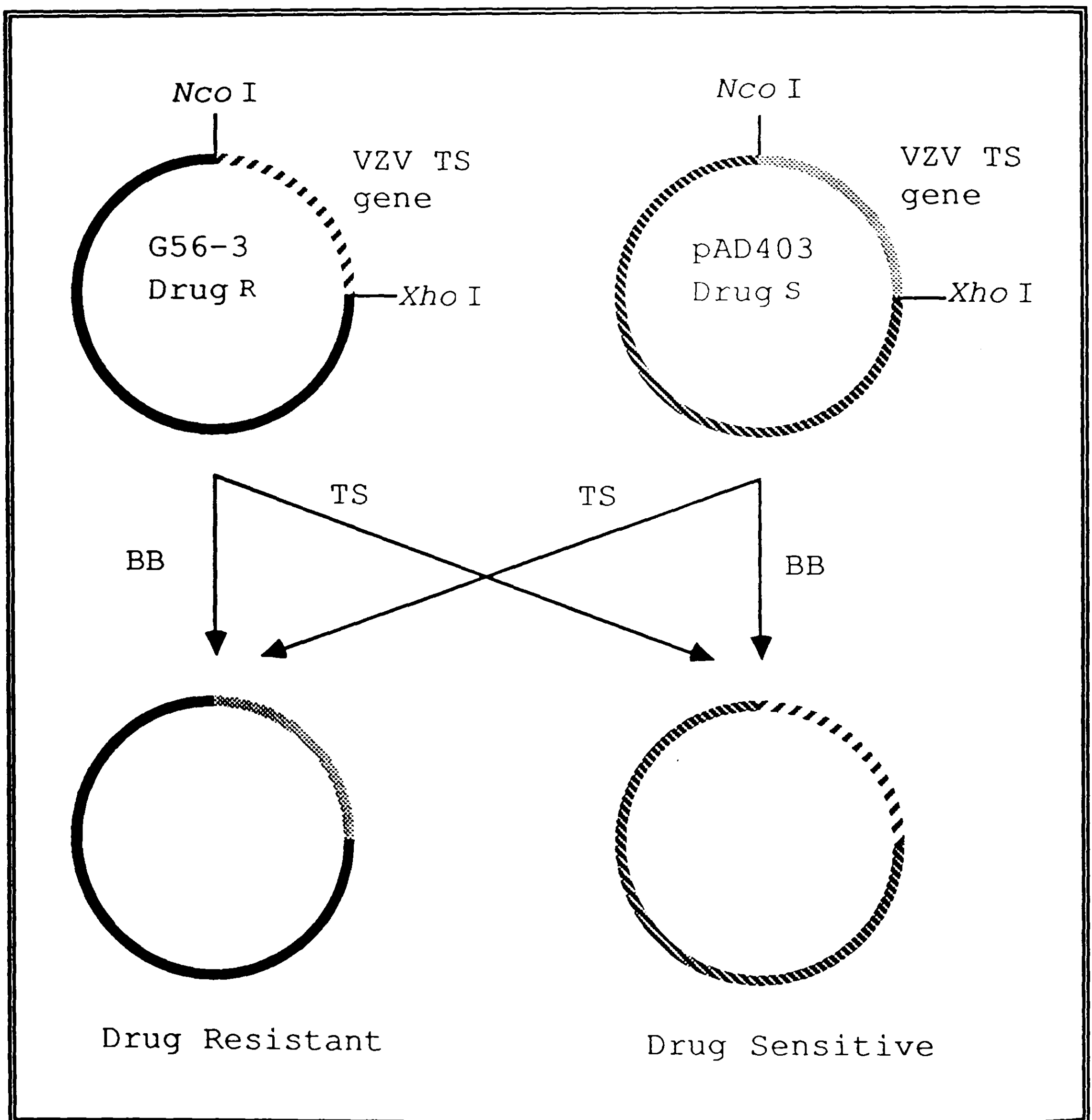



Fig. 3.6. **Drug resistance clones.**
Confirmation that the drug resistance is caused by mutations in the promoter region rather than in the TS coding region. The *Nco*I site spans the initiating ATG codon of the TS coding sequence. BB indicates the vector backbone.

initiation codon), immediately upstream of the -10 region of the bacterial promoter as shown in Fig. 3.5.

To confirm that the drug resistance was linked to the nucleotide change in the bacterial promoter, two constructs were made. pAD403 and the drug resistant G56-3 clone were digested with *Nco*I and *Xho*I. This digest generates two DNA fragments from each plasmid, a TS containing fragment, the initiating ATG codon being part of the *Nco*I restriction site, and a vector backbone fragment. The four fragments were isolated. The TS containing fragments of pAD403 and G56-3 were ligated to the reciprocal backbone fragment (see Fig. 3.6). The resulting DNA was used to transform the *E.coli* strain SØ928*thyA* and the transformants assayed for drug resistance. The constructs made from the G56-3 backbone and WT TS fragment formed single colonies on FUdR0.4 plates. The constructs made from the wild type backbone and the G56-3 TS fragment failed to form single colonies when streaked onto FUdR0.4 plates. This confirms that the drug resistance was linked to a mutation outside of the TS coding region. This is most likely the identified mutation in the bacterial promoter (see Fig. 3.5), but the possibility of additional mutations that affect plasmid copy number cannot be ruled out.

In addition, these results showed that bacteria that could grow as a bacterial lawn at 2.0µM FUdR could only form single colonies at 0.4µM FUdR, whilst the same clones that could grow as a lawn at 10µM TFT could only form single colonies at 2.0µM TFT. This observation lead to a minor alteration of the drug screening protocols (Methods 4 and 5; see Section 2.12.2). The WT pAD403 clone and two clones derived from the T56 reaction (where dTTP was the limiting nucleotide), could not form single colonies at these drug concentrations.

### 3.1.6.2 **Wide scale screening using methods 4 and 5**

At this point the mutations in the promoter of pAD403 were identified and subsequently corrected to make the vector pAD768 (see Section 3.1.3.1). For all the

subsequent experiments, U-ssDNA was made from pAD768. Misincorporation mutagenesis was performed using the five oligos, 56, and 59 to 62, as described. Approximately 10,000 clones were screened against both FUdR and EYUdR from DNA libraries generated from each oligo by either method 4 or 5, and the following resistant clones were isolated.

Clones generated from oligo 56 where dGTP was the limiting nucleotide (G56 clones) grew as a lawn in the presence of 10μM EYUdR, compared to virtually no growth of pAD768 on the same plates. Bacteria from both plates were streaked for single colonies on EYUdR0.5, 1, 2, 5 and 10 plates. Bacteria from the G56 reaction formed single colonies at concentrations of up to 10μM EYUdR, whilst the wild type pAD768 clone failed to form single colonies above a concentration of 0.5μM EYUdR.

As a continuation of drug screening method 4, DNA was isolated from both clones to see if drug resistance was plasmid linked. The G56 clone retained its resistance to EYUdR whilst the wild type pAD768 clone retained sensitivity to this drug. Both remained sensitive to FUdR. The TS gene of the G56 derived isolate was sequenced with oligos 59 to 62 but revealed no nucleotide changes. In addition, no sequence changes were seen in the upstream promoter region of this clone.

In addition to the G56 EYUdR$^R$ clone, transformants derived from the misincorportion mutagenesis reaction using oligo 60 with dCTP (C60 clones) or dTTP (T60 clones) and the pAD768 control each formed approximately 10 single colonies on EYUdR10 plates. When the C60, T60 and pAD768 resistant clones were tested for chloramphenicol resistance, the majority were sensitive. This strongly suggests that loss of the plasmid pCBTK3 was the cause of drug resistance. By way of contrast, the G56 EYUdR$^R$ clone was resistant to both ampicillin and chloramphenicol, resistance to the latter drug showing that the plasmid expressing the VZV TK was present and that resistance to EYUdR was not due to loss of this plasmid. The isolation

of chloramphenicol sensitive mutants led to the inclusion of this antibiotic in all further experiments that used the drug screening protocols.

### 3.1.6.3 **Summary and statistical analysis**

Screening of approximately 10,000 clones for each oligo resulted in the isolation of only a single EYUdR resistant clone. This clone had no changes in the VZV TS coding region or known promoter elements. Moreover, the cells expressing this clone had a copy of the pCKTK3 plasmid, suggesting the expression of VZV TK, which should ensure the phosphorylation of EYUdR. The mechanism of drug resistance in this clone remains unresolved, but could involve mutations in the origin of replication that control plasmid copy number. A plasmid with an increased copy number would be expected to over-express TS in much the same way as MTX resistant strains of *L.casei* (see Section 1.14.2). However, since it did not contain mutations in the TS gene, it was not processed further.

This method of mutagenesis can generate all the possible single nucleotide changes over a given target area (Lehtovaara *et al.*, 1988; Holm *et al.*, 1990). To estimate the number of clones that need to be screened in order to give a high level of confidence (>99%) of covering all possible point mutations, both the mutagenesis efficiency and efficacy must be determined.

The mutagenesis protocol is highly efficacious with mutants distibuted randomly across the target area, and sequencing studies suggesting an efficiency of ~83% (a mutant in 5 out of 6 clones; see Table 3.2 - clones M through R).

If the mutagenesis efficiency is 100%, then the number of clones that will need to be screened to give a high degree of confidence of identifying all possible single nucleotide changes is given by the following equation (Hutchinson *et al.*, 1986).

$$N = [\ln(1-p^{1/s})/\ln(1-1/s)] + s/2 \quad \text{(Equation 1)}$$

Where p = Degree of confidence (0.99)
N = Number of clones to be screened
s = Number of possible single nucleotide changes

Since each oligo (56 and 59 to 62) can be expected to mutagenize approximately 250 nucleotides (Lehtovaara *et al.*, 1988; Holm *et al.*, 1990), then for any of 3 nucleotide replacements at each position this gives a value for s of 750. This means that 8,785 clones must be screened in order to give a 99% level of confidence for evaluating all the possible single nucleotide substitutions. Experimentally, at least 10,000 clones were analysed in each case. A mutagenesis efficiency of 83% reduces this figure to 8,300 clones. By equation (2) (Hutchinson *et al.*, 1986), this gives a 98.1% degree of confidence that all the single nucleotide subtitutions were seen.

$$p = [1-e^{[\ln(1-1/s)(N-s/2)]}]^s \quad \text{(Equation 2)}$$

If the mutagenesis efficiency was as low as 66%, corresponding to 6,667 mutant clones being screened, this still results in an acceptable degree of confidence (84.4%) of screening all the possible single point mutations.

In summary, this data suggests that all the possible single nucleotide substitutions were made in the VZV TS gene, and when screened for resistance to FUdR or EYUdR, none of these changes conferred a drug resistant phenotype.

### 3.1.7 Screening clones generated by spiked oligo mutagenesis

5,500 clones (500 clones per oligo), generated from U-ssDNA made from pAD768 and using all 11 spiked oligos, were screened by method 4 for resistance to EYUdR. Clones derived from oligo R02 grew significantly better than pAD768 and clones derived from the other oligos on EYUdR10 plates. When streaked on to FUdR20 and EYUdR10 plates, the oligo R02 derived clones formed single colonies whereas pAD768 and clones derived from the other oligos did not.

| Clone | FUdR | | | EYUdR | | |
|---|---|---|---|---|---|---|
| | 5μM | 10μM | 20μM | 1μM | 2μM | 4μM |
| R01 | + | - | N.T. | + | - | N.T. |
| R02 | + | - | N.T. | + | - | N.T. |
| R03 | + | - | N.T. | + | - | N.T. |
| R05 | + | - | N.T. | + | - | N.T. |
| R06 | + | + | + | + | - | - |
| R07 | + | + | + | + | - | - |
| R08 | + | - | - | + | + | + |
| pAD768 | + | - | N.T. | + | - | N.T. |

Table 3.6. **Drug Resistant Variants.**
Initial characterization of clones derived from spiked oligo mutagenesis. The growth of clones R01 to R03, and R05 to R08 in liquid M9 medium, supplemented with various concentrations of FUdR and EYUdR. + indicates growth; - indicates no growth; N.T. indicates not tested.

| Clone | Solid Medium | | Liquid Medium | |
|---|---|---|---|---|
| | FUdR (3μM) | EYUdR (0.5μM) | FUdR (20μM) | EYUdR (4μM) |
| R06 | single colonies (1) | single colonies (2) | growth | growth |
| R07 | single colonies (3) | no growth | growth | growth |
| R08 | single colonies (4) | single colonies (5) | growth | growth |
| pAD768 | no growth | no growth | no growth | no growth |

Table 3.7. **Further characterisation.**
Clones isolated from screening DNA generated from spiked oligo mutagenesis. Growth clones R06 to R08 in solid and liquid M9 medium, supplemented with various concentrations of FUdR and EYUdR. Numbers in parentheses indicate clones processed further (see text for details).

The drug resistant R02 derived clone formed single colonies when streaked onto LB-Amp/Cm, confirming the presence of pCBTK3. However, further analysis **of this** clone demonstrated that it did not have **a plasmid** linked phenotype, and this clone was not process**ed further**.

A further 11,000 clones (1,000 **clones per** oligo), generated from U-ssDNA made from pAD768 were screened by method 4 for resistance to both FUdR and EYUdR. All were sensitive. Another 3,500 clones (500 clones per oligo), generated from U-ssDNA made from pAD768 and using the doped oligos R01 to R03 and R05 to R08 were screened by screening method 5 for resistance to both FUdR and EYUdR. All the cultures grew in the presence of FUdR (5µM) or EYUdR (1µM) (see Table 3.6). When sub-cultured into fresh drug containing medium, cultures derived from pAD768 and oligos R01 to R03 and R05, failed to grow in the presence of FUdR (10µM) or EYUdR (2µM). Cultures derived from oligos R06 and R07 grew in the presence of FUdR (10µM) but failed to grow in media containing EYUdR (2µM), whilst the culture derived from oligo R8 grew in the presence of EYUdR (2&4µM; Table 3.6). When drug screening method 5 was continued, clones R06, R07 and R08 formed single colonies in the presence of the two drugs and grew in liquid culture at higher drug concentrations as shown in Table 3.7.

Mini-prep DNA was made from the five drug resistant clones that grew as single colonies (see Table 3.7) and digested with *Eco*RI. The restriction pattern of the TS expressing plasmid was the same as the wild type pAD768 plasmid. The restriction pattern of the plasmid pCBTK3 of the drug resistant clones 1 and 2 (numbering as in Table 3.7) were identical to that of pCBTK3 from drug sensitive control cells and purified **plasmid**, whilst the restriction pattern of clones 3 to 5 **were different**. When DNA of all these clones was used to **retransform** ***E.coli*** SØ928*thyA/* pCBTK3, transformants containing **DNA from** clones 3 to 5 were sensitive to both FUdR and EYUdR, whilst transformants containing DNA from clones 1 and 2 were resistant to FUdR and EYUdR. These two plasmid linked drug

resistant clones (1 and 2) grew in the presence of ampicillin and chloramphenicol, confirming the presence of the pCBTK3 plasmid. The TS gene of these two resistant clones was sequenced using oligos 59 to 62 but no nucleotide changes were detected.

#### 3.1.7.1 Statistical analysis

This method of mutagenesis can generate any number of point mutations within the given target region, depending upon the number of changes in the oligo. Before discussing the results in more detail a statistical analysis of this method of mutagenesis will be described.

To calculate the degree to which the oligos should be spiked ($\alpha$) to give the desired number of changes (c) for an oligo of a defined length (n), the following equation is used (Ner *et al.*, 1988; McNeil & Smith, 1985).

$$\alpha = c/n \qquad \text{(Equation 3)}$$

Where $\alpha$ = Fraction to which each oligo mix should be spiked with the other three phosphoramidite mixes
n = Length of oligo
c = Desired number of changes

The value for $\alpha$ was calculated for n = 51 (the average length of the 11 spiked oligos) and c = 2. This gives a value for $\alpha$ of 0.04. On this basis, the oligonucleotide synthesis mixes were spiked as follows. 0.04 vols of a fresh vial of each mix were removed, and replaced by 0.013 vols of the other 3 mixes. The four vials, each contained 96% of the original precursor and 1.33% of the other three nucleotide precursors. The oligonucleotides were then synthesized as normal (the spiking of the oligonucleotides mixes and subsequent synthesis of the oligos was kindly performed by John McLauchlan).

From the binomial distribution, the probability of the number of changes in each oligo can be calculated by the following equation (Hermes *et al.*, 1990).

42-mer

| c | s | N | P | N* |
|---|---|---|---|---|
| 1 | 126 | 1,248 | 0.32 | 4,696 |
| 2 | 7,749 | 108,910 | 0.27 | 485,986 |
| 3 | 309,960 | 5,500,269 | 0.15 | 44,178,871 |
| 4 | 9,066,330 | 191,566,940 | 0.06 | 3,846,725,606 |

57-mer

| c | s | N | P | N* |
|---|---|---|---|---|
| 1 | 171 | 1,746 | 0.23 | 9,149 |
| 2 | 14,364 | 210,752 | 0.27 | 940,439 |
| 3 | 790,020 | 14,758,887 | 0.21 | 84,675,200 |
| 4 | 31,995,810 | 717,669,530 | 0.12 | 7,205,517,374 |

Table 3.8. **Statistical analysis of spiked oligo mutagenesis.** The values of s, N, P and N* for a 42-mer and 57-mer for different values of c.

c = Number of base changes per oligo.
s = Number of different mutations per oligo (see Equation 5).
N = Number of clones to be screened to give a 99% degree of confidence of analysing all possible clones.
P = Expected fraction of clones containing c changes (calculated for c = 2).
N* = Number of clones to be screened (N) multiplied by 1/P multiplied by 1/E, where E = mutagenesis efficiency (see text for details).

$$P_{(c,n,a)} = [n!/(n-c)!\ c!]\ [a]^c\ [1-a]^{n-c} \qquad \text{(Equation 4)}$$

As this method of mutagenesis can introduce more than one change, the number of different mutations (s), each oligo can generate can be calculated by the following equation, derived from equation 4 (Hermes *et al.*, 1990).

$$s = [n!/(n-c)!\ c!]\ [3]^c \qquad \text{(Equation 5)}$$

From Equation 1, the value of N, the number of clones that must be screened to give a 99% degree of confidence of analysing all possible variants, can be determined. However, this figure must be corrected to account for the mutagenesis efficiency (E; ~83% in our experiments), and the fraction of the clones that contain c changes (P) as shown in the legend to Table 3.8. The values of P, N and $N^*$ for two representative oligos are shown in Table 3.8.

3.1.7.2 **Summary**

A number of drug resistant clones were isolated although none had mutations in the VZV TS gene. In several instances, the plasmid encoding the VZV TK was found to have undergone some kind of rearrangement such that it could no longer express an active form of TK. This would account for the resistant phenotype observed. For the two clones where the TK encoding plasmid appeared normal by *Eco*RI restriction enzyme analysis, other mechanisms of resistance must be considered. The VZV TK gene may contain mutations that inactivate the enzyme and account for the drug resistant phenotype. However, as the gene is expressed on a multicopy plasmid, every copy would be required to harbour this mutation, thus making this option unlikely. Mutations that alter the copy number of the TS expressing plasmid are also possible.

At this point it seemed that the drug selection scheme was likely to reveal high numbers of drug resistant clones that had mutations in the VZV TK gene rather than the VZV TS gene. Furthermore, our statistical analysis performed

for this method of mutagenesis had shown that an inordinately large number of clones would have to be screened to ensure that all possible mutations were tested.

Clearly these values were not ach ved, the typical number of clones being screened was approximately 2,000. For a 42-mer this only gives a 4.7% degree of confidence of screening all the possible 1 base substitution mutants (<<1.0% for a 57-mer). Similar calculations were performed for the number of clones required to detect all the two, three and four base substitutions (Table 3.8). Screening for all the three and four base changes was clearly prohibitive. An extension of the screening procedures to encompass all the two base substitution mutants was also discounted for the following reasons. The screening of clones generated by misincorporation mutagenesis has already shown that a single base substitution is unlikely to confer drug resistance to VZV TS (see Section 3.1.6). To screen all of the clones necessary to cover all the permutations of two substitutions would require of the order of 10,000,000 clones to give a 99% chance of covering all the possible mutants, which in itself is a considerable, though feasible, task.

However, problems were emerging with the drug resistance screen, namely the continued isolation of false positive clones. In addition, the screening of this number of clones would still only give access to an average of ~15.7 other amino acids at each position (Hermes *et al.*, 1990). Although it is possible that a single amino acid change is responsible for the drug resistant phenotype of TS previously reported (see Section 1.14.1), there is no guarantee that a single amino acid replacement would confer resistance to VZV TS. In addition, there is no certainty that the equivalent mutation in TS from two different sources would have the same effect on activity (see Section 3.4.1). Together, these arguments lead to the termination of this part of the project.

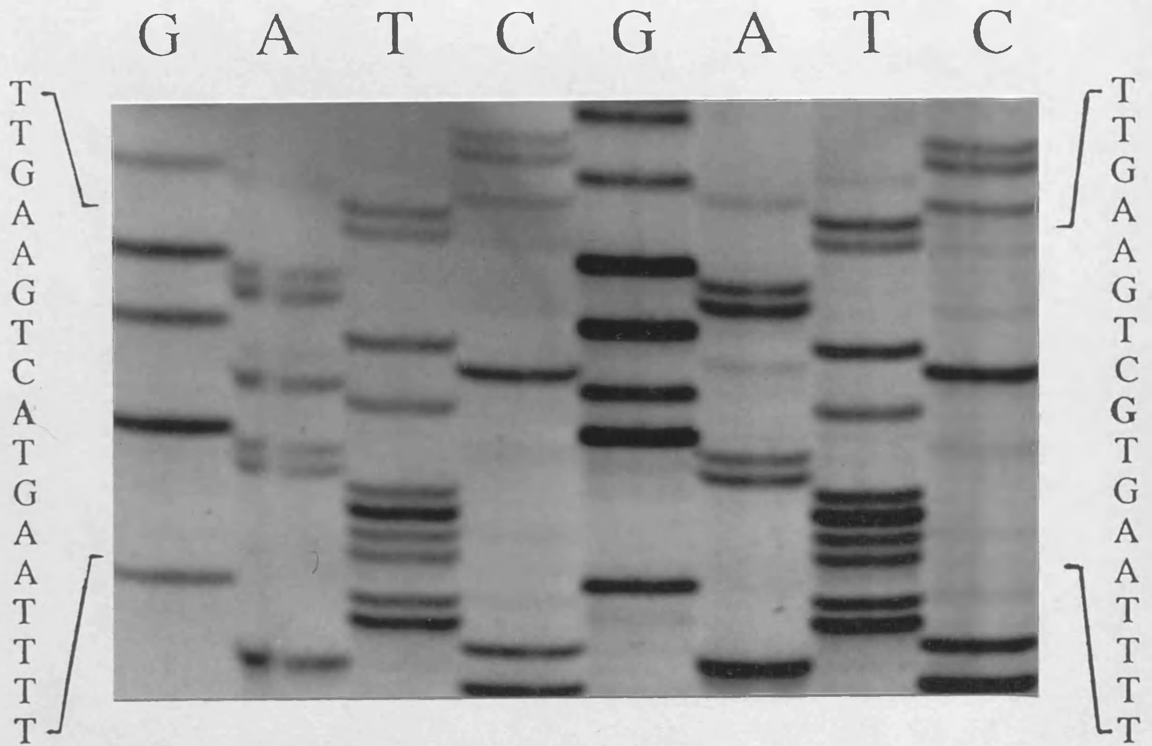



Fig. 3.7. **Sequencing gel.**
This figure shows a portion from a representative sequencing gel. The four tracks on the left are of the wild type VZV TS gene (pAD768) and the four tracks on the right are from the Y21H variant of VZV TS. The tracks were loaded G, A, T, C. Both clones were sequenced with oligo 62. The nucleotide substitution, A61G is indicated in **bold**. The sequence shown corresponds to the non-coding strand of the gene.

### 3.1.8 Site-directed mutagensis studies

The failure of the random mutagenesis studies to isolate a variant of VZV TS that is resistant to FdUMP suggests that only a very limited number of mutations, if any, may give rise to a drug resistant enzyme. Indeed, only one instance of a variant of TS that showed decreased affinity for FdUMP, but retained catalytic activity has been reported (Jastreboff *et al.*, 1983; see Section 1.14.1). In this case, the amino acid changes responsible for the decrease in affinity were not determined.

A second report has shown that the TS expressed in a human cell line resistant to FUdR contains a single amino acid substitution, Y33H (Barbour *et al.*, 1990). These authors reported that the variant enzyme has decreased affinity for FdUMP, but did not show conclusively that it retains catalytic activity (see Section 1.14.1). However, as this variant constitutes the only known amino acid substitution that may confer FdUMP resistantce to TS, it was decided to make the same amino acid substitution at the equivalent residue of VZV TS. By way of a control, and to see if the human variant retains catalytic activity as suggested (Barbour *et al.*, 1990), the same mutation was introduced into human TS.

#### 3.1.8.1 Mutant construction and enzyme activities

Both mutants were constructed by site-directed mutagenesis. Tyrosine 21 of VZV TS (the equivalent of Y33 of human TS; see Fig 1.6) was replaced by histidine in pAD768 using oligo 70 to create the mutant Y21H. Tyrosine 33 of human TS was replaced by histidine in pAD876 using oligo 72 to create the mutant Y33H. The TS gene of Y21H was sequenced with oligos 3 to 7 and revealed no nucleotide changes apart from those introduced by site-directed mutagenesis (see Fig. 3.7). The region of the TS gene of Y33H that could be sequenced by oligo 74 revealed no nucleotide changes apart from those introduced by site-directed mutagenesis. The enzymes were expressed in the *E.coli thyA* strain $\chi$2913. The TS activities measured by

| Expressed enzyme | Tritium release assay pmols dTMP/min/ug | % WT |
|---|---|---|
| VZV WT | 105.8 ± 10.9 | 100.0 |
| VZV Y21H | 5.7 ± 1.3 | 5.3 |
| Human WT | 28.8 ± 2.7 | 100.0 |
| Human Y33H | 11.0 ± 2.7 | 38.2 |

Table 3.9. **TS assays.**
TS activity of the VZV and human wild type and variant enzymes measured by the tritium release assay. All enzymes were expressed in *E.coli* $\chi$2913. Activities are given as mean ± standard error of the mean.

| Expressed Enzyme | Growth in M9 medium | | |
|---|---|---|---|
| | Patch | Single Colonies | Liquid Culture |
| VZV WT | + | + | + |
| VZV Y21H | + | - | + |
| Human WT | + | + | + |
| Human Y33H | + | - | + |

Table 3.10. **Growth complementation assays.**
The growth complementation of *E.coli* $\chi$2913 expressing the TS variants is shown for single colonies from LB-Amp-TdR plates which were patched or streaked on to solid M9 medium, or used to inoculate liquid M9 medium. + indicates growth; - indicates no growth.

| Expressed Enzyme | FUdR concentration (μM) | | |
|---|---|---|---|
| | 1.0 | 5.0 | 10.0 |
| VZV WT | + | + | + |
| VZV Y21H | + | - | - |
| Human WT | + | - | - |
| Human Y33H | + | +/- | - |

Table 3.11. **Assay of drug resistance.**
The growth of *E.coli* SØ928 *thyA*/pCBTK3 expressing the TS variants i liquid minimal medium supplemented with various concentrations of FUdI + represents growth; +/- represents poor growth; - represents no growth.

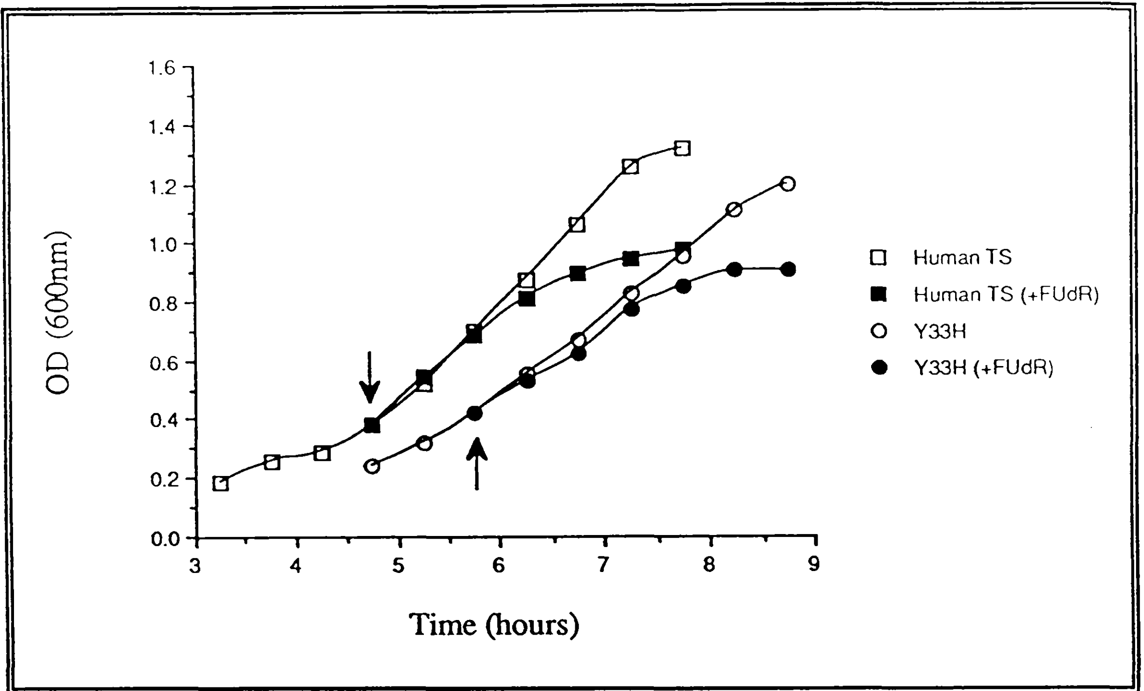



Fig. 3.8 **Characterization of Y33H.**
Effect of 50µM FUdR on the growth of *E.coli SØ928thyA* cells expressing the wild type and Y33H variant of human TS. Arrows indicate the addition of FUdR to a final concentration of 50µM.

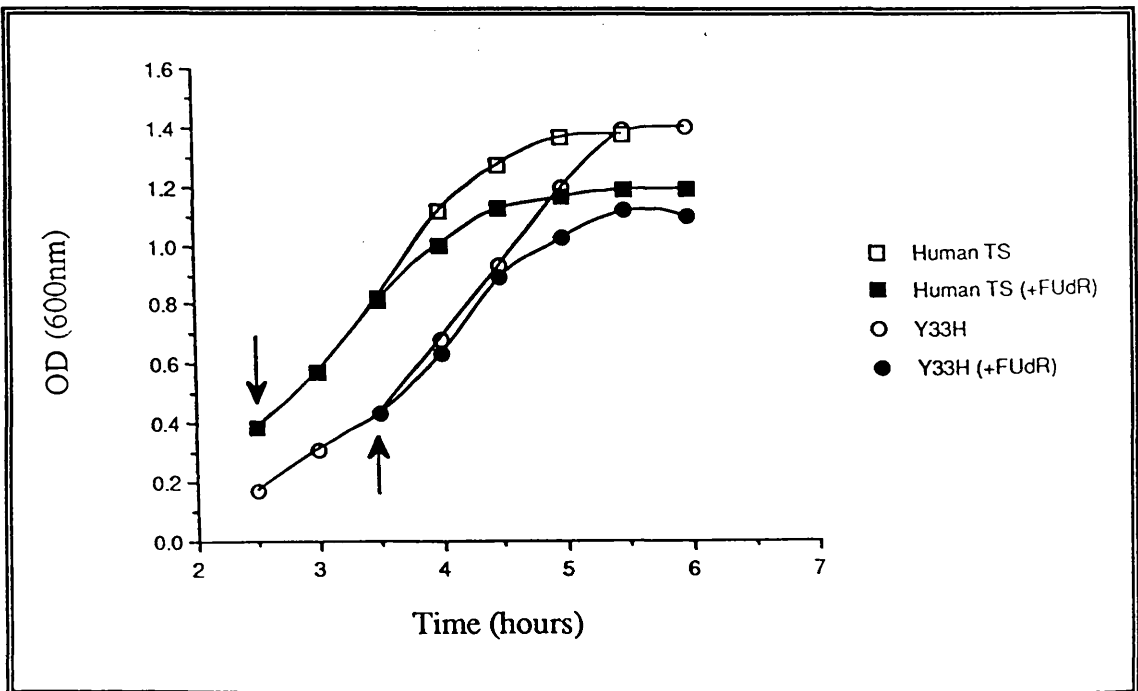



Fig. 3.9. **Characterization of Y33H.**
Effect of 100µM FUdR on the growth of *E.coli SØ928thyA* cells expressing the wild type and Y33H variant of human TS. Arrows indicate the addition of FUdR to a final concentration of 100µM.

the tritium release assay are shown in Table 3.9. The mutants were also scored by the TS growth complementation assay (see Table 3.10). The inability of the two mutants to form single colonies on M9 medium resulted in the use of alternative methods to assess the drug resistance of these clones.

The mutant and wild type clones were expressed in SØ928*thyA*/pCBTK3 (see Section 3.1) and scored for the ability to grow in liquid M9 medium supplemented with FUdR (see Table 3.11). The VZV mutant Y21H is clearly not drug resistant in that at the lowest concentration that the mutant does not grow, the wild type still grows. However, cells expressing the human mutant Y33H grow poorly in the presence of 5μM FUdR compared with cells expressing the wild type enzyme that do not grow at all.

To further characterize the apparent drug resistance shown by the human variant Y33H, the affect of adding FUdR at the mid-log stage of growth of SØ928*thyA*/pCBTK3 expressing the wild type human TS and the Y33H variant was analysed (see Section 2.13.4 and Figs. 3.8 and 3.9). Both the human wild type TS and Y33H variant grew at similar rates, although the mutant takes longer to commence exponential growth. 50μM FUdR has an almost identical inhibitory effect on both wild type human TS and the Y33H variant. When 100μM FUdR is added, cells expressing the Y33H mutant do not reach the same final cell density as cells expressing the wild type enzyme. Moreover, cell lysis is clearly visible in the cultures of cells expressing the Y33H variant in the presence of 100μM FUdR. This suggests that the Y33H enzyme is slightly more susceptible to the effects of FUdR than the wild type human TS.

#### 3.1.8.2 **Molecular modelling**

The Y21H mutant was modelled on the VZV TS structure to investigate why the histidine substitution leads to the 20-fold drop in enzyme activity. In the wild type enzyme structure, the aromatic ring of the Y21 side chain is very tightly packed against the side chains of M207, V211 and

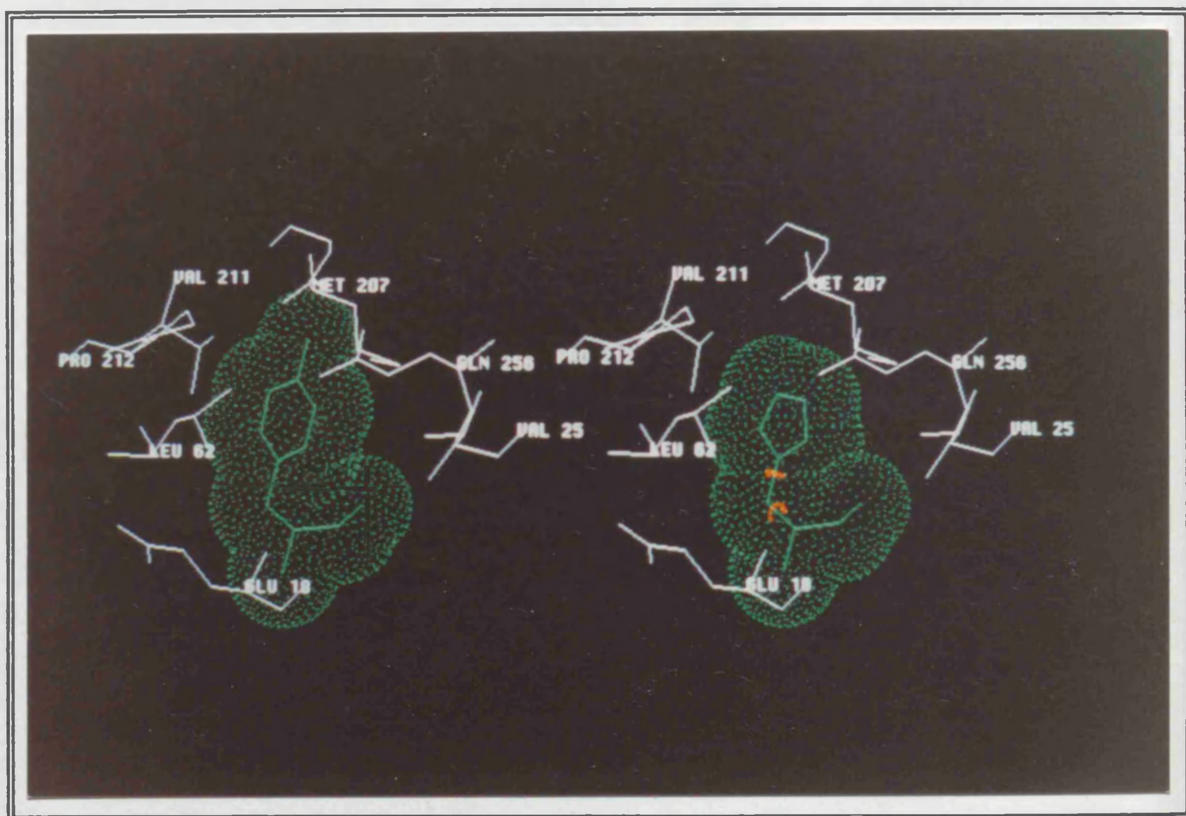



Fig. 3.10. **Y21H**
The wild type enzyme is shown on the left with the mutant on the right. Amino acid 21 is shown in green. The red arrows illustrate that rotation about the $C_\alpha$-$C_\beta$ bond and $C_\beta$-$C_\gamma$ bond of the H21 side chain has been performed. Residues M207, V211 and P212 that form the bend between β-strand iii and the J-helix are shown at the top of the picture. Several other neighbouring amino acids are also shown.

P212 that form the bend between ß-strand iii and the J-helix, and also against the side chains of E18, L22, V25, L62 and Q256 (Fig. 3.10). All of these residues are very highly conserved in all the known TS sequences (see Fig 1.6). In addition, the hydroxyl group of the Y21 side chain points in the direction of the side chain of M297, however, the role of this hydroxyl group, if any, is not clear from the model. No potential H-bonds are apparent, and it would seem to be the presence of the bulky aromatic ring of the side chain of Y21, rather than its hydroxyl group that is the conserved feature.

When a histidine residue is modelled at this position of VZV TS, the major disruption is the loss of hydrophobic contacts to the junction between ß-strand iii and the J-helix. In addition, the distance between the delta carbon of M297 and the side chain of amino acid 21 increases from 5.6 to 7.7 Ångströms.

The introduction of a charged residue and the creation of a large gap (and possibly the loss of a hydroxyl group) in a buried region of the protein in which movement is essential for catalysis, may explain why this mutant loses activity. Exactly why the human variant retains catalytic activity is not apparent from the VZV model. It should however be noted that the N-terminus of the human TS is 12 amino acids longer than the VZV TS, and as the residue being studied here (Y21) is close to the N-terminus, the conformation that these residues adopt may account, at least in part, for the observed phenotypic differences.

#### 3.1.8.3 **Summary**

Tyrosine 21 (Y33 in human TS) is almost totally conserved in all TS sequences reported to date, the single exception being replacement by phenylalanine in *L.lactis* (see Fig 1.6). Such a substitution (Tyr to Phe) is generally accepted as a conservative substitution that would be expected to have little effect on the overall structure of a protein (Taylor, 1986; Bordo & Argos, 1991; Rennell *et al.*, 1991; John Overington, personal communication).

However, the more disruptive Tyr to His substitution would be expected to have a greater affect on enzyme activity.

The mutant Y21H of VZV TS clearly does not have a drug resistant phenotype. This mutant should have been detected during the misincorporation mutagenesis procedure, as it requires only a single point mutation to change Y21 (codon **TAT**) to H21 (codon **CAT**). By way of contrast, the Y33H variant of human TS differs from the VZV Y21H enzyme in that it retains significant catalytic activity (Table 3.9). This mutation has been proposed as the major determinant of the FUdR resistant phenotype of a human colorectal tumour cell line (Barbour *et al.*, 1990), and the data presented here go some way to supporting this claim. The Y33H variant enzyme has detectable TS activity, but the assays used here gave conflicting results concerning the drug resistant character of this mutant.

The observation that the same mutation in TS from different sources can have different affects on enzyme activity, is not unusual (see Section 1.11), and is most probably the result of local differences in accomodation of the amino acids side chains (see Sections 1.9, 3.2 and 3.3). Purification of both the VZV and human TS mutants, and detailed kinetic analysis could lead to a better understanding of these differences.

## 3.2 Mutants created by random mutagenesis

During the assessment of the efficiency of the random mutagenesis protocols, a number of clones were identified that failed to complement the growth of an *E.coli thyA* strain. The aim of this section is to exploit the availability of the VZV TS models, in an attempt to determine why the mutations disrupt enzyme activity. In addition the comparison of the activities of some of these mutants with that of the wild type enzyme establishes a link between *in vivo* complementation and enzyme activity.

| Expressed enzyme | Tritium release assay pmol/min/µg | % WT |
|---|---|---|
| WT | 105.8 ± 10.9 | 100.0 |
| S102L | 0.9 ± 0.3 | 0.9 |
| C187G | <0.05 | <0.05 |
| F189L | 4.9 ± 0.5 | 4.7 |

Table 3.12. **TS assays.**
Tritium release assay of three of the mutants isolated by random mutagenesis. Activities are given as the mean ± standard error of the mean. All variants were derivatives of pAD768 expressed in the *E.coli* strain $\chi$2913.

| Clone | Solvent accessibility |
|---|---|
| I41L | buried |
| N54K | buried |
| I96M | buried |
| S102L | buried |
| D174A/174H175 | buried |
| V179E | accessible |
| C183W/Q188K | accessible |
| C187G | buried |
| F189L | buried |
| S197T/C198R | buried |

Table 3.13. **Mutants of VZV TS.**
The amino acid changes in the mutants isolated during the random mutagenesis protocols and the solvent accessibilty of their side chains in the wild type enzyme. 174H175 represents the insertion of histidine between residues 174 and 175.

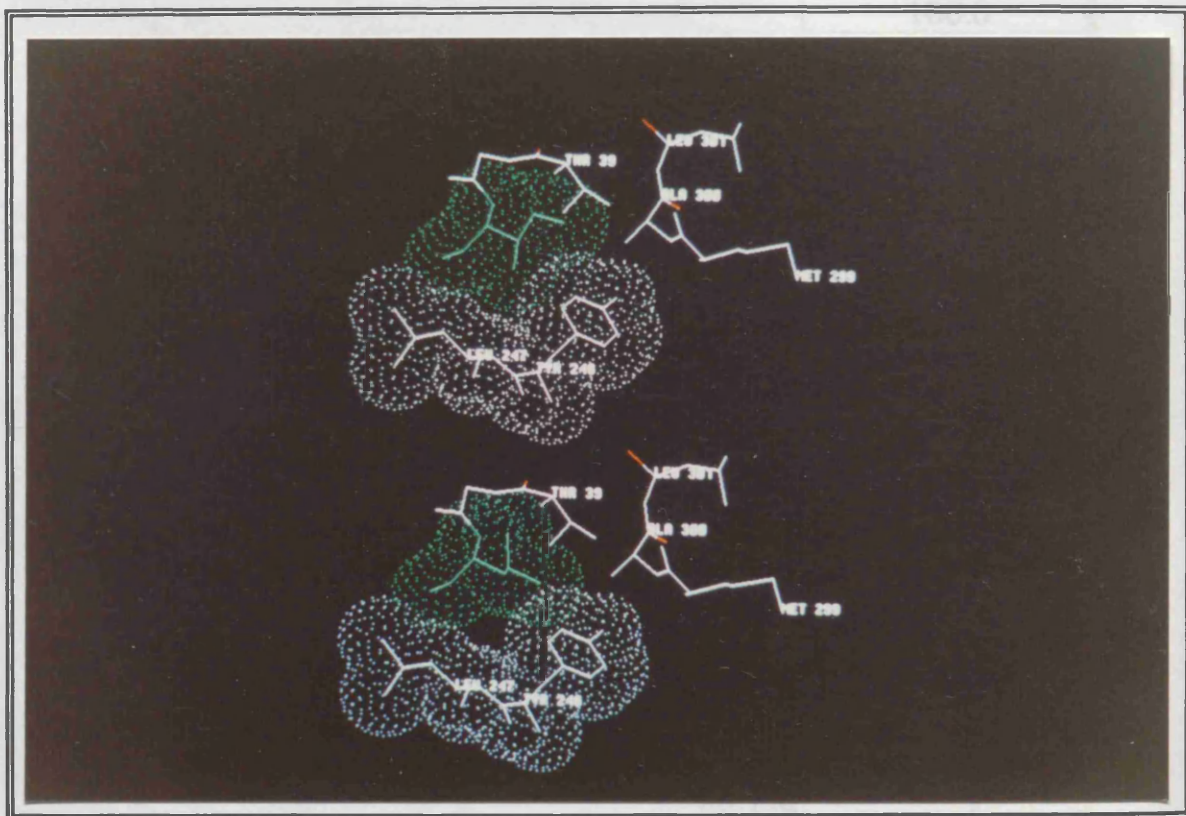



Fig. 3.11. **I41L**
The model of the wild type VZV TS is shown at the top of this figure and the mutant is shown at the bottom. Residue 41 is highlighted in green in both cases. Van der Waals radii are shown for I/L41, Y246 and L247. The backbone carbonyl groups of several residues are highlighted in red. The most striking feature is the gap created between Y246 and L247.

### 3.2.1 Mutant enzyme activities

The activity of three of the mutant enzymes in the tritium release assay are shown in **Table 3.12**. These results suggest that a specific activity of between 1% and 5% of the wild type level in crude extracts results in a Thy$^-$ phenotype. This value is slightly higher than the values obtained in more detailed studies of *L.casei* TS (~1%; Climie *et al.*, 1990) and *E.coli* TS (0.97% to 1.25%; Michaels *et al.*, 1990).

### 3.2.2 Modelling Studies

The interpretation of the phenotype of these mutants has been facilitated by the availability of the models of VZV TS generated by Celia Schiffer and Robert Stroud (personal communication) based on the *E.coli* TS structures determined by the same group (Perry *et al.*, 1990; Finer-Moore *et al.*, 1990; Montfort *et al.*, 1990). A full list of the mutants studied is given in Table 3.13. The classification of the residues in relation to their state of solvent accessibility is also shown.

Several different classes of possible explanation for loss of activity are suggested in the following sections. These include steric clashes with other amino acid side-chains and the polypeptide backbone, creation of gaps in the structure which may lead to instability, or the stabilization of an unfavouable structure, and disruption of possible H-bonds. In each particular case, the enzyme is thought to be incapable of accomodating these changes, without invoking structural abberations that ultimately lead to loss of activity. A brief description is given for each mutant, and in several cases this is illustrated with representations of the enzymes tertiary structure.

#### 3.2.2.1 I41L

Isoleucine 41 packs neatly into a cleft created by the hydrophobic side-chains of Y246 and L247 to form part of a flat hydrophobic surface, which is at right angles to the orientation of the backbone carbonyl groups of T39 and

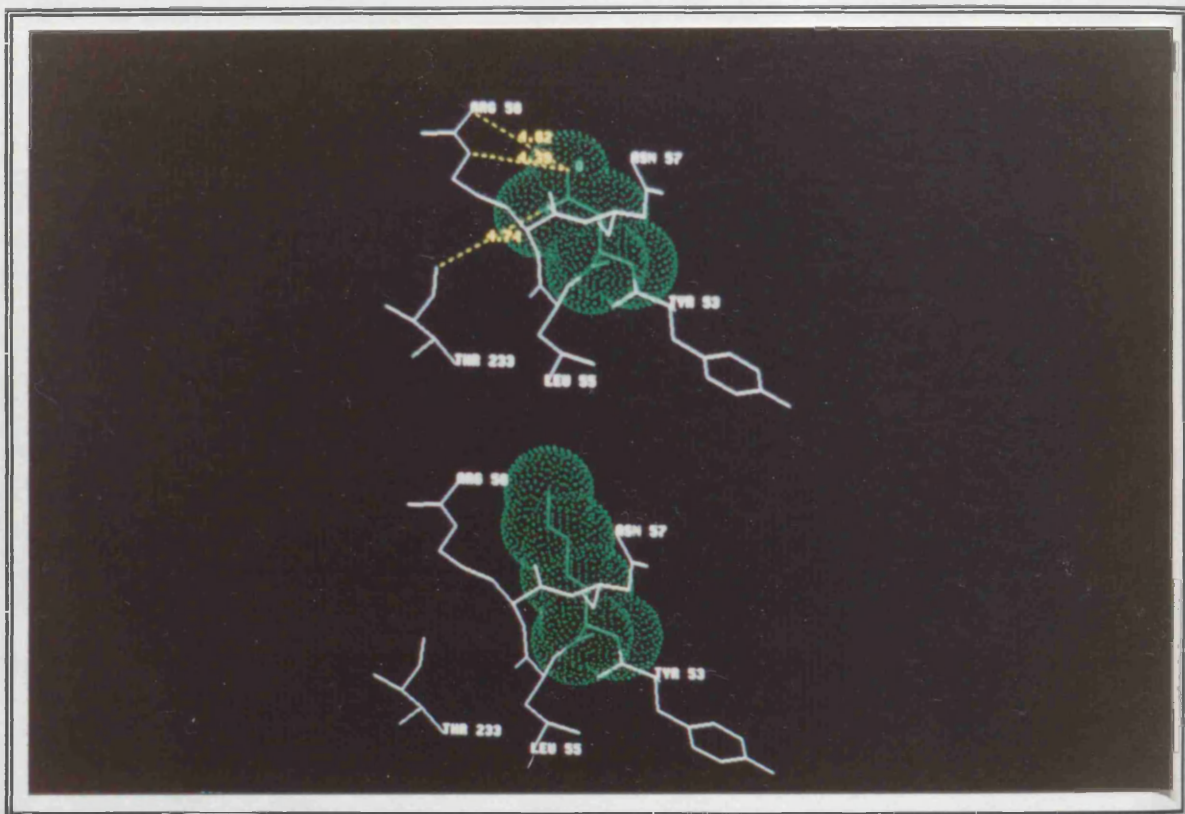



Fig. 3.12. **N54K**
The wild type enzyme is shown at the top and the mutant is shown at the bottom. Residue 54 is shown in green with van der Waals radii. The $O_\delta$ of the side chain of N54 is labelled. Potential H-bonds between the side chain of N54 and side chain of R56 and backbone carbonyl group of T233 are represented with a dashed yellow line, the distance shown in Å. As H atoms are not shown, the length of the H-bonds are increased by ~1.2Å.

M299, and the carboxyl group of the C-terminal residue (Fig. 3.11). When a leucine residue is modelled in place of isoleucine, its side-chain projects out of this plane and leaves a gap in the cleft of Y246 and L247 (see Fig. 3.11). This gap may lead to the incorrect formation of this region of the protein which could disrupt the orientation of the surrounding residues.

A comparison with other TS sequences at the position equivalent to VZV 41 shows that only threonine and valine are found in this position (except for phage ø3T and *L.lactis*, which have an alanine at this position). Both of these amino acids, like isoleucine, have a branched ß-carbon, one arm of which contains a lone methyl group. Both valine and threonine can be modelled in this position, easily filling the cleft between Y247 and L248, without disrupting the planar hydrophobic surface that is interupted by the side chain of leucine. This confirms the need for the presence of a methyl group in the cleft between these residues, and its absence due to the substitution by leucine is probably the major factor that contributes to enzyme inactivation. This observation strongly suggests that the backbone carbonyl groups of T39 and M299, and the carboxyl group of the C-terminal residue may play some role in catalysis or substrate binding. Interestingly, the side-chain hydroxyl group of threonine points in exactly the same direction as the other three carbonyl/carboxylate groups, augmenting the possiblity of H-bond interactions involving these groups.

3.2.2.2 **N54K**

Asparagine 54 resides at the carboxy end of ß-strand i and is naturally seen substituted only by serine and aspartate in the known TS sequences. Analysis of the VZV TS model reveals that the N54 side chain may form hydrogen bonds with the side chains of T233, R56 and N57 (see Fig. 3.12). Substitution of N54 with lysine would prevent hydrogen bond formation. In addition, a lysine side chain would clash sterically with the side chain of R56 and the

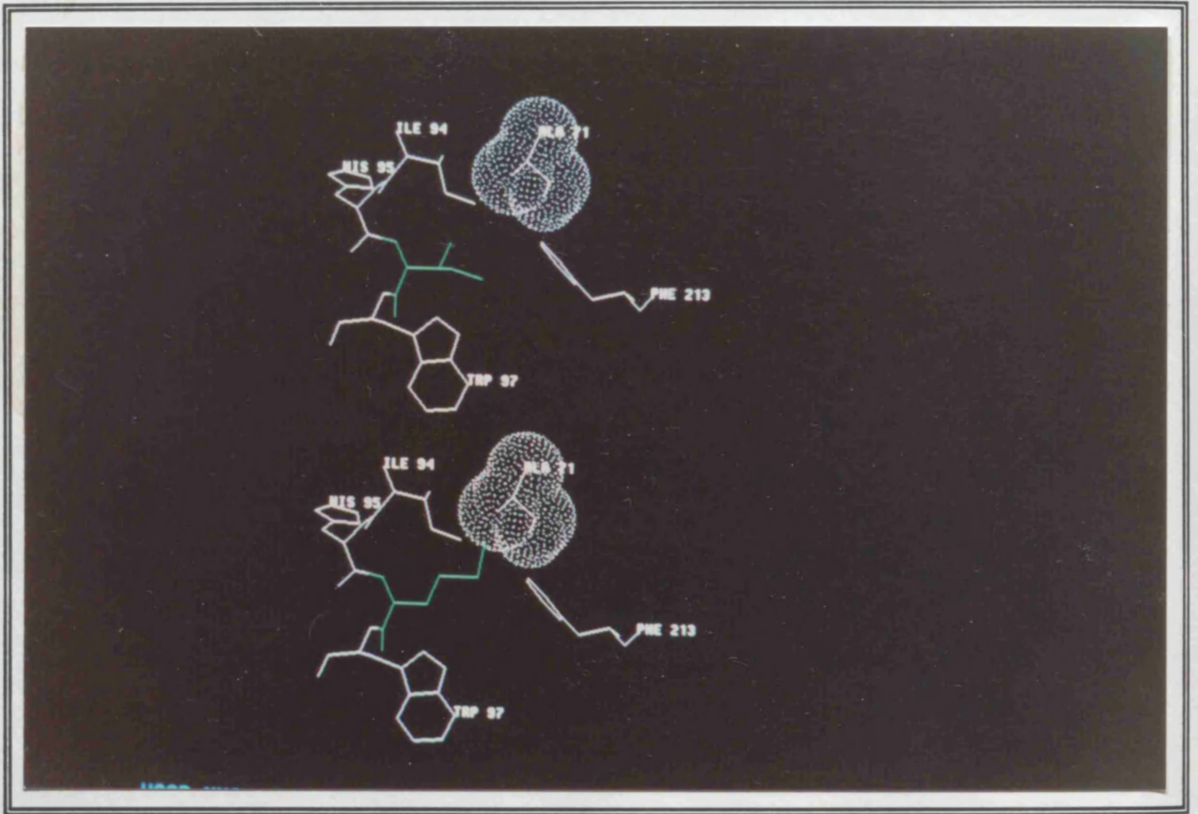



Fig. 3.13. **I96M**
The wild type enzyme is shown at the top and the mutant is shown at the bottom. Residue 96 is depicted in green and van der Waals radii are shown for A71.

resulting juxtaposition of two positive charges would also be destabilizing. The protein is probably unable to accomodate this amino acid substitution without disrupting the structure of ß-strand i which is essential for dimer formation and movement against other elements of secondary structure during catalysis.

3.2.2.3 **I96M**

Isoleucine 96 is a highly conserved residue substituted only in TS from phage T4 and *L. lactis*. This residue forms part of a conserved loop between helices C and D, forming close contacts with other residues in this loop. Substitution by methionine would cause a steric clash with the side-chain of alanine 71 in helix B (see Fig. 3.13). If a glycine residue is modelled at position 71, the methionine subtitution no longer clashes sterically. Interestingly, a glycine is found at the equivalent position in all other eukaryotic TS sequences, suggesting that the equivalent Ile to Met substitution may be tolerated in TS from other sources. Further analysis of the VZV TS model reveals that the substitution of I96 with other hydrophobic amino acids could result in an inactive enyzme. A valine residue at position 96 would leave an unfavourable gap, whereas leucine at this position would create the same steric problems as methionine. These observations may partly explain the conserved nature of this residue.

3.2.2.4 **S102L**

The side chain of S102 sits in a pocket created by the residues G101, S103, K104 and D118 and is within hydrogen-bonding distance with the side-chain amino group of K104. This amino acid is highly conserved in the eukaryotic TS sequences, substituted only in *Plasmodium* spp. by threonine. Indeed, a threonine residue can be modelled at this position of VZV TS, preserving the potential to from a H-bond with K104, thus it seems likely that such a contact is made in all eukaryotic TSs. A leucine residue

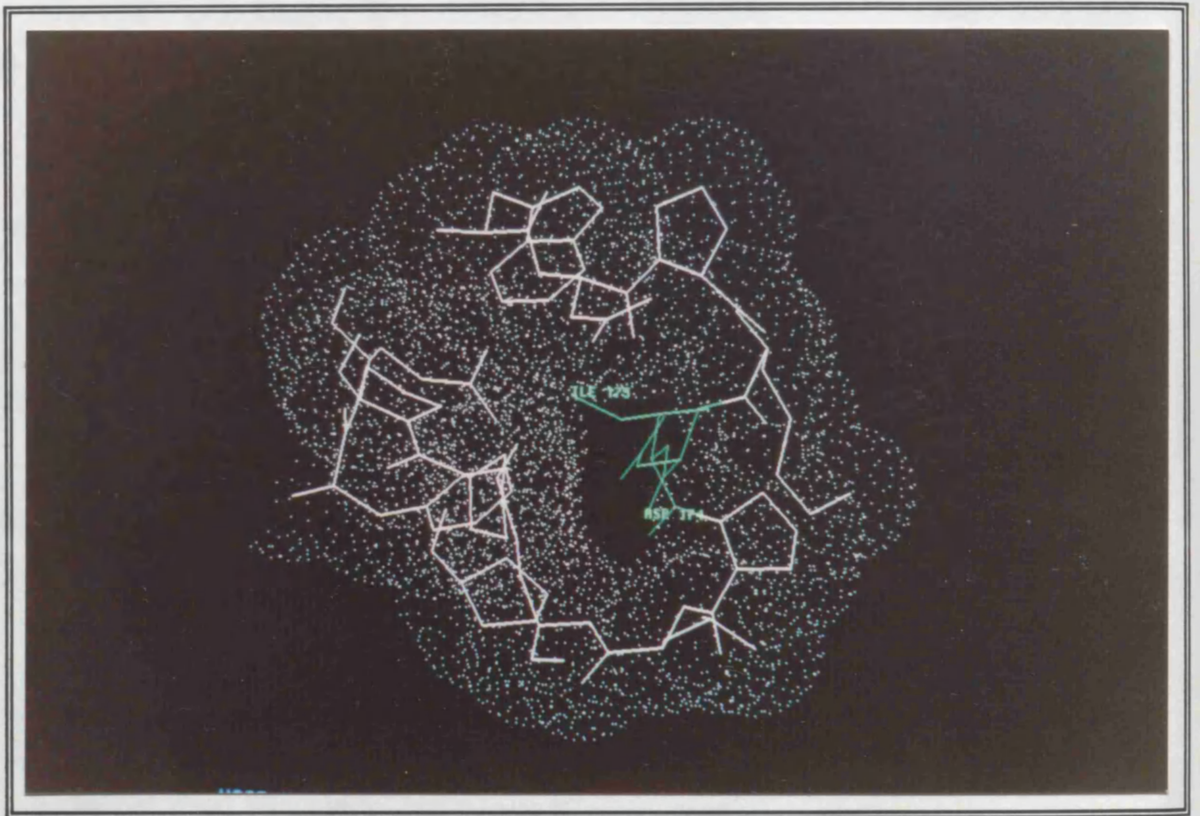



Fig. 3.14. **D174A/174H175**
Only the wild type enzyme is shown in this case as the mutant could not be accurately modelled. Residues D174 and I175 are coloured in green. The closely packed neighbouring residues are shown with their van der Waals radii, commencing with W170 at the top of the figure, and continuing toward the C-terminus in a clockwise direction.

modelled at this position not only interferes sterically with Q117 and D118, but removes the possibility of a H-bond with K104, due to the loss of the side-chain hydroxyl group. Unfortunately, this region of VZV TS differs in sequence, and presumably structure, from prokaryotic TSs, and these differences may affect the accuracy of the corresponding region of the VZV model (as it is based on the *E.coli* structure). For this reason, no representation of this mutant is shown here and, the affects of the amino acid substitution must be interpreted with additional caution.

3.2.2.5 **D174A/174H175**

The two residues at this position (D174 and I175) are located in helix I and conform to the consensus sequences (D/E/A) (I/V/L/F) respectively (see Fig. 1.6). Our analysis of this region reveals that residues W170 through K173 which form a very highly conserved loop that joins ß-strand v and helix I, form a surface perfectly complementary to residues D174 and I175 (see Fig. 3.14). It also illustrates beautifully how three highly conserved (P172, P176, P181) and one invariant proline residue (P182) 'wrap' themselves around these two residues forming a compact buried region of the protein. The D174A mutation is not that disruptive *per se*, as shown by modelling. However, the insertion of a histidine residue cannot be modelled in this region without gross disruption of the surrounding structure. It is therefore likely that the disruption of the precise packing of residues in this region is the cause of enzyme inactivation. It should be recalled that P182 immediately precedes the active site C183 residue.

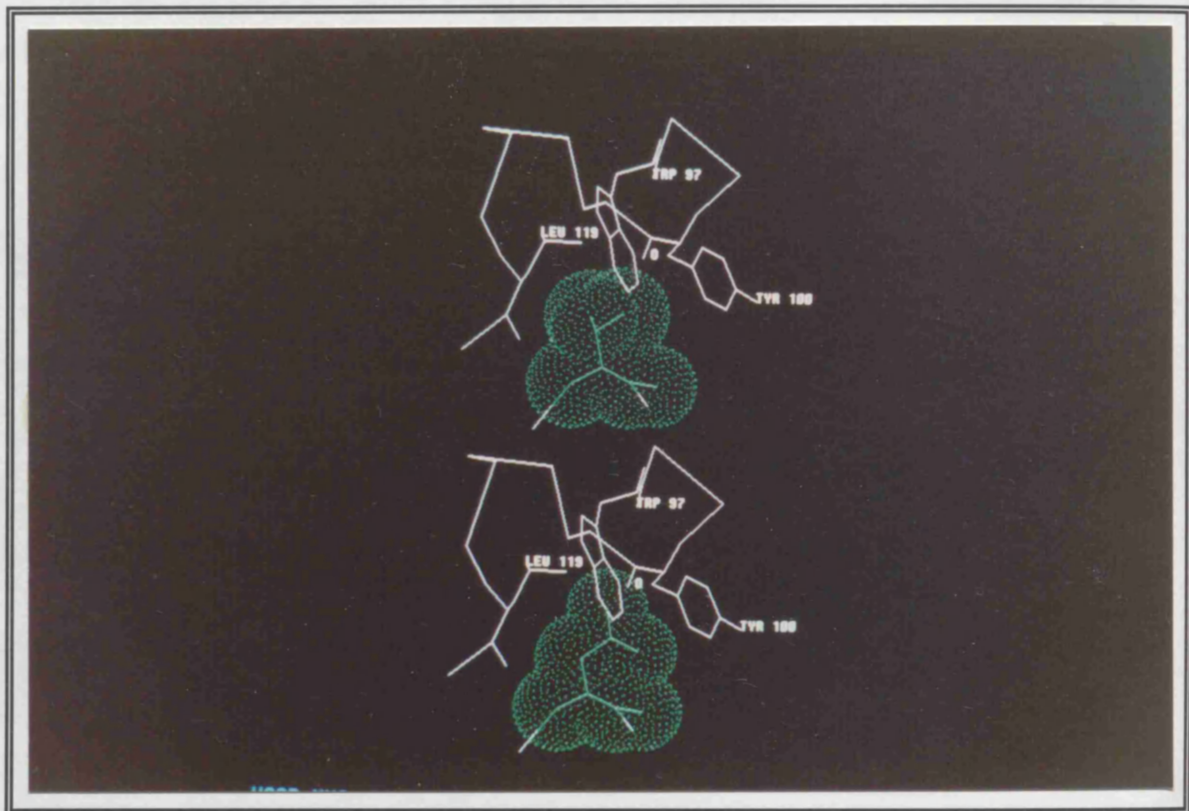



Fig. 3.15. **V179E**
The wild type enzyme is shown at the top and the mutant is shown at the bottom. Amino acid 179 is shown in green with van der Waals radii. The steric clash of the glutamate side chain of the mutant with the backbone carbonyl group of Y100 is clearly seen.

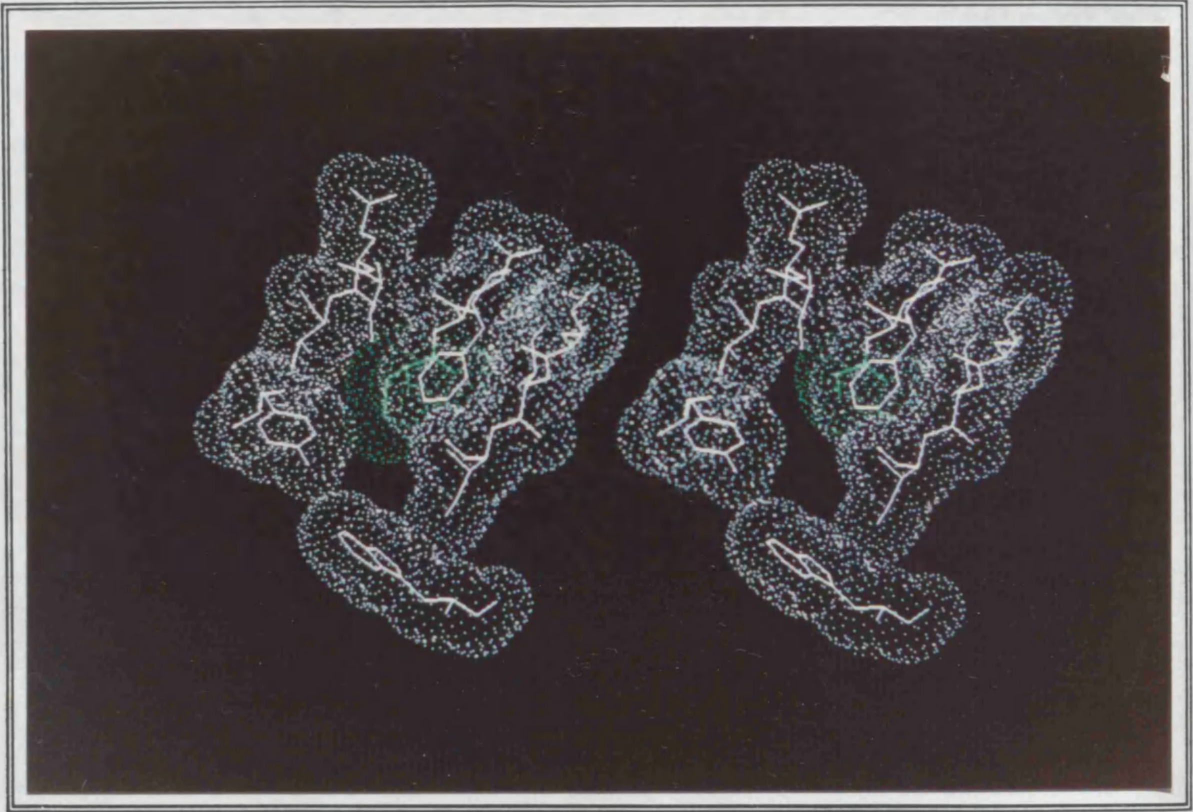


Fig. 3.16. **C187G**
The wild type enzyme is shown on the left and the mutant is on the right. Residue 187 is depicted in green. Residues from β-strands iii, iv and v are shown, as is W97, with van der Waals radii shown for all residues.

3.2.2.6 **V179E**

This residue lies at the end of helix I, and is replaced in other TS sequences by the amino acids alanine and glycine which in itself suggests the requirement for a small amino acid at this position (see **Fig.** 1.6). V179 fits snugly into a small pocket created by Y100, L119, I122 and L180 (see Fig. 3.15). The side chain of E179 cannot fill this pocket and clashes with the backbone of Y100 from helix D. Although one arm of the valine side chain points towards the solvent, modelling of the glutamate side chain such that it is exposed to the solvent, could only be achieved by unacceptable distortion of its side chain.

3.2.2.7 **C183W/Q188K**

The two substitutions in this mutant were not modelled as the lack of the active site cysteine residue alone is almost certainly the cause of the loss of enzyme activity (see Section 1.11).

3.2.2.8 **C187G**

This residue lies in the middle of ß-strand iv and its side-chain points into the core of the TS monomer, clearly suggesting that it has no interactions with the second subunit of the dimer. It is not a well conserved residue, although in eukaryote TS sequences, only amino acids with small side-chains are found at this position (see Fig. 1.6). C187 is packed against the side chains of residues C198, V200 and Y218, and is also close to F189 (see Fig. 3.16). Substitution by glycine creates a gap that could lead to inefficient hydrophobic packing in this region, which could ultimately result in enzyme inactivation.

The factors that affect structural accomodation have recently been quantified (see Section 1.9; Lim & Sauer, 1991; Perry *et al.*, 1990). One of the observed rules governing accomodation is that, in general, for a reduction of three or more heavy atom equivalents (carbon, nitrogen, oxygen or sulphur), a second site or covariant mutation is

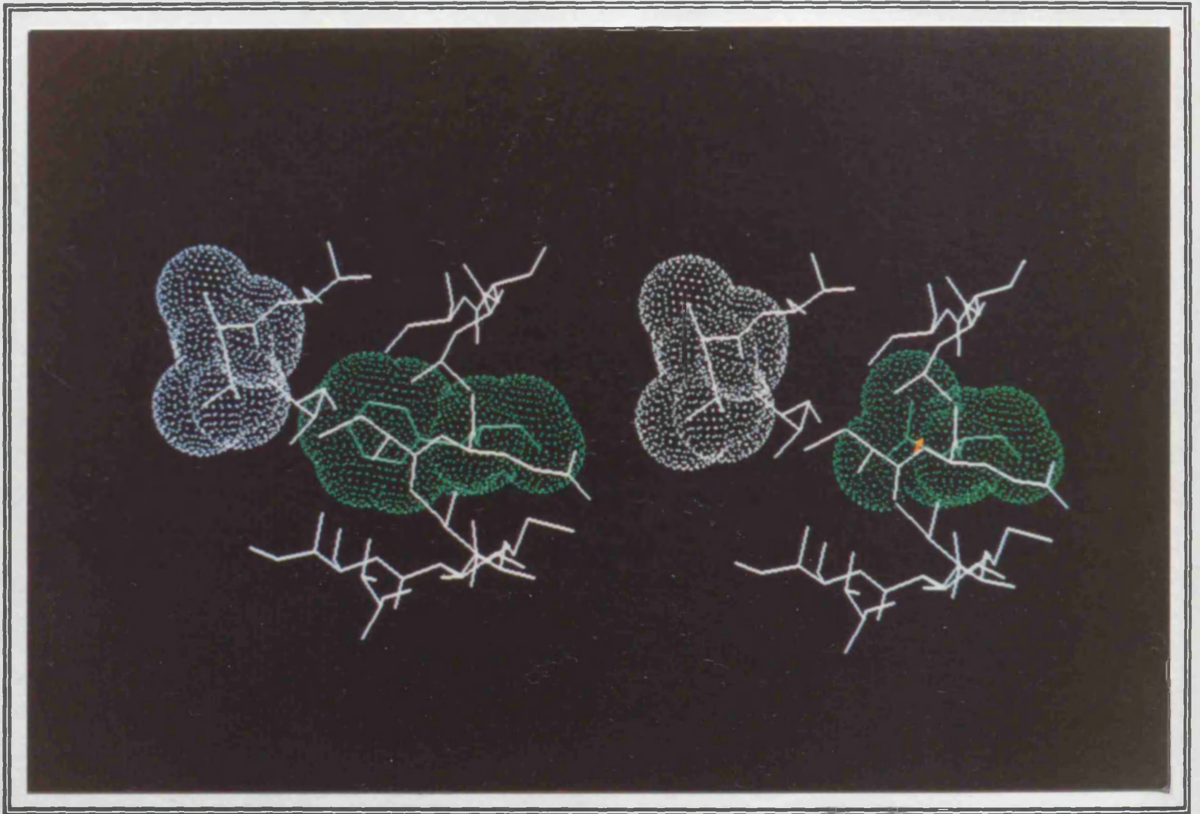


Fig. 3.17. **F189L**
The wild type enzyme is shown on the left and the mutant is shown on the right. Residue 189 is shown in green with van der Waals radii. L196 and S197 are also shown with van der Waals radii. The red arrow indicates that rotation about the $C_\beta$-$C_\gamma$ bond of the L189 side chain was performed. A prominent gap is seen in the F189L mutant that is normally filled by the side chain of F189 in the wild type enzyme.

required. In this case, only two heavy atom equivalents have been lost, but this may still be sufficient to cause the need for a second site mutation.

Two additional observations add weight to this proposal. Firstly, a glycine residue is seen at this position in *C.fasiculata* TS suggesting that the presence of a glycine alone is not disruptive to the ß-strand *per se*. Secondly, the gap created in this substitution in VZV TS, may well be accomodated in *C.fasiculata* TS. A detailed analysis of the VZV TS model suggests that residue 200 may be covariant with residue 187. Indeed, when V200 is replaced with leucine (seen at the equivalent position of *C.fasiculata* TS) the leucine side chain can fill the gap created by the original C187G substitution.

3.2.2.9 **F189L**

The aromatic side-chain of F189 falls neatly into a pocket created by residues L196, S197 and C198 from ß-strand iv, and L221, T222 and V225 from the J-helix (Fig. 3.1). Modelling reveals that a Leu residue cannot fill this pocket, which presumably is important for the correct packing of the core residues. Such a view correlates with the observation from a detailed analysis of the core packing of the lambda repressor, that a decrease in the hydrophobic core of greater than two methylene group equivalents causes disruption of the core and results in loss of protein function (Lim & Sauer, 1989; 1991). F189 is invariant, save for phage ø3T where it is replaced by tryptophan. Tryptophan was modelled such that its aromatic side-chain fitted into this pocket of the VZV TS, but this resulted in a steric clash with one arm of the side-chain of V152. Interestingly, the equivalent residue is Leu in phage ø3T, which, when modelled at position 152 of VZV TS (with tryptophan substituted for phenylalanine at position 189), facilitates the accomodation of W189. This suggests that these two residues may be an example of a covariant pair (see Section 1.6).

These results suggest that the positioning of a large

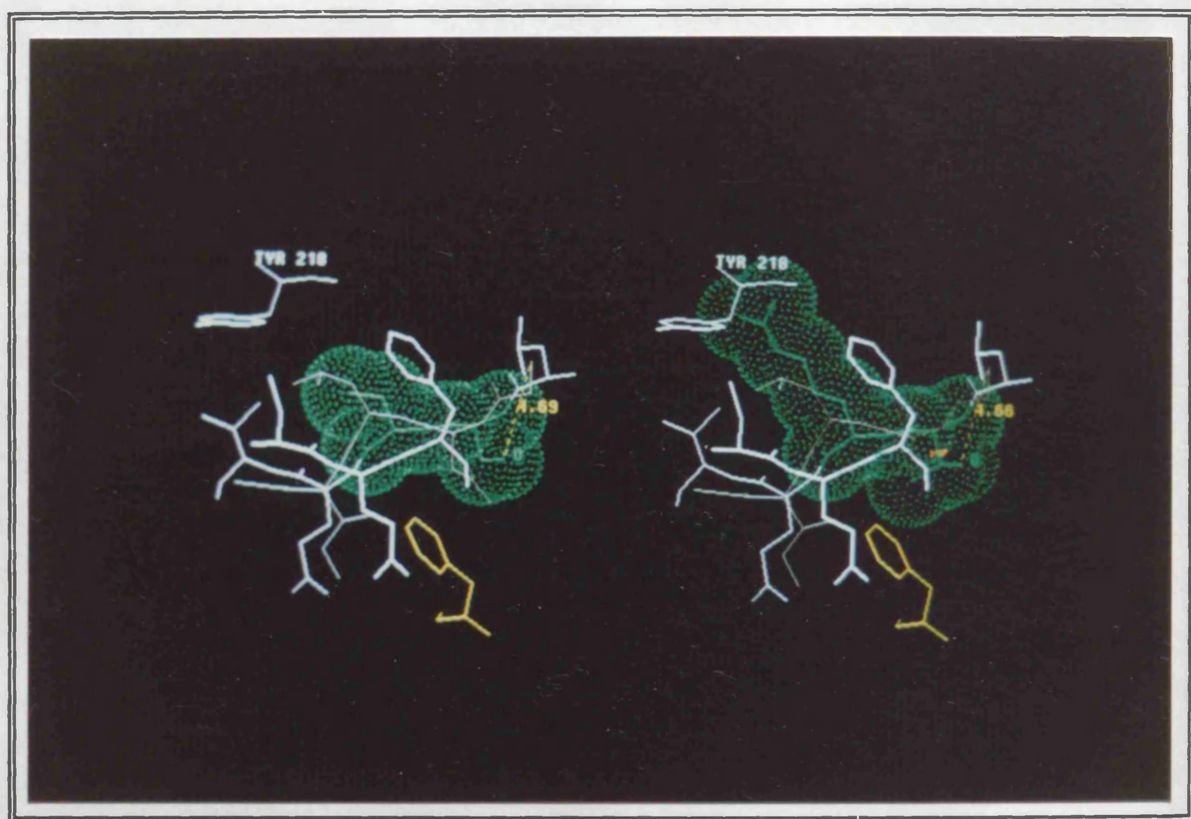



Fig. 3.18. **S197T/C198R**
The wild type structure is shown on the left and the mutant on the right. Residues 197 and 198 are coloured in green and shown with their van der Waals radii. F47' from the second subunit is coloured in yellow. The H-bond between the side chain hydroxyl group of S/T197 to the backbone NH of G234 is shown by the dashed yellow line, and the distance depicted in Å. As H atoms are not shown on the model, the length of the H-bonds is increased by ~1.2Å.

side chain in the groove created by residues 196 to 198 may be essential for the correct folding or stability of this region. This analysis may explain why the substitution of phenylalanine by leucine, often considered a conserved change (Bordo & Argos, 1991; Taylor, 1986; John Overington, personal communication), in this case leads to enzyme inactivation.

#### 3.2.2.10 **S197T/C198R**

These two residues are completely conserved in all eukaryotic TS sequences reported to date. Indeed the tight packing of their side chains in the interior of the protein, surrounded by highly conserved residues, provides some reason for the high degree of conservation. Replacement of S197 with threonine preserves a H-bond from the backbone amino group of G234 to the serine/threonine side chain hydroxyl group, but the additional methyl group of T197 clashes with the side chain of F47' and the backbone carbonyl group of D235 (see Fig. 3.18). The more disruptive C198R substitution results in a major steric clash with the invariant Y218 residue. Together these steric clashes probably account for the observed enzyme inactivation.

### 3.2.3 **Summary**

These analyses show the combined power of modelling, phylogenetic comparisons and random mutagenesis studies. They provide arguments as to why some residues are highly conserved, suggest a variety of ways in which substitutions lead to inactivation and describe possible amino acid changes that could be made to probe the way in which TS can accomodate mutations.

Many of the mutants are inferred to lead to inactivation due to steric clashes or the creation of gaps in core regions of the protein. Steric clashes must be overcome in order to form a stable protein structure. In some cases, the protein may be unable to accomodate these clashes producing an unstable enzyme unable to catalyse the normal

```
                          43       i       54
                           |^^^^^^^^^^^^|
VZV                DRTGiGTlSlFG.mqarYnLrne.FP
HVS                DRTGtGTlSiFG.tqsrFsLene.FP
HVA                DRTGvGTlSvFG.mqsrYsLekd.FP
Human              DRTGtGTlSvFG.mqarYsLrde.FP
Mouse              DRTGtGTlSvFG.mqarYsLrde.FP
C.albicans         DRTGtGTkSlFappqlrFdLsndtFP
P.carinii          DRTGtGTlSvpapsplkFsLrnktFP
S.cerevisiae       DRTGtGTlSlFappqlrFsLrddtFP
C.fasciculata      DRTGvGTiSlFG.aqm.FslrdnqLP
L.amazonensis      DRTGvGTlSlFG.aqmrFsLrdnrLP
L.major            DRTGvGTiSlFG.aqmrFsLrdnrLP
L.tropica          DRTGvGTiSlFG.aqmrFsLrdnrLP
P.chabaudi         DRTGvGvlSkFG.ymmkFnLsey.FP
P.falciparum       DRTGvGvlSkFG.yimkFdLsqy.FP
E.coli             DRTGtGTlSiFG.hqmrFnLqdg.FP
B.subtilis         DRTGtGTiStFG.yqmrFnLreg.FP
L.lactis           DgqmanskyvtG.sfvtYdLqkgeFP
L.casei            DRThtGTySiFG.hqmrFdLskg.FP
Tn4003             DRThtGTiSkFG.hqlrFdLtkg.FP
Phage T4           DRTGtGTialFG.sklrWdLtkg.FP
Phage phi3T        DgTpahTlSvms.kqmrFdnse..vP
```

Fig. 3.19. **Alignment of β-strand i sequences.**
The alignment of the residues in and around β-strand i of the 21 known TS sequences. Residues shown in capitals are conserved in at least 17 of the sequences. *C.albicans*, *P.carinii* and *S.cerevisiae* are the three yeast sequences referred to in the text. For the full legend, see Fig. 1.6.

reaction. Alternatively, the required repacking may lead to the re-orientation of residues that play a critical role in catalysis. In cases where the amino acid substitution appears to cause no steric clashes or readily detectable disruption to the final structure, the phenotype may reflect disruption of the folding pathway (Rennell *et al.*, 1991).

### 3.3 Studies of key structural regions of TS

Two regions of TS were chosen for analysis by site-directed mutagenesis. The first was a buried region, ß-strand i, whilst the second, the carboxy-terminal tail, lies on the surface of the protein.

#### 3.3.1 ß-Strand i

ß-strand i has several important roles to play in the function of TS (Matthews *et al.*, 1989, 1990b; Finer-Moore *et al.*, 1990). However the amino acid sequences show some degree of redundancy in TS from different organisms (see Fig. 3.19). To investigate these differences, several mutants were constructed in this region of VZV TS, and their effects on enzyme activty determined.

##### 3.3.1.1 Mutant construction and enzyme activities

All of the mutants were constructed by site-directed mutagenesis. Glycine 48 was replaced by alanine in pAD403 using oligo 57 to create the mutant G48A. The mutant G48A/48P49/M49P was constructed using oligo 58, whilst oligo 67 was used for mutant 58T59. Mutant G48A/48P49/M49P/58T59 was constructed using oligo 67 and template DNA made from pAD403 containing the G48A/48P49/M49P mutations. All four mutants complemented the growth of the *E.coli* strain $\chi$2913. The mutants were initially identified by slot blot hybridization or by direct DNA sequencing. The TS gene of each mutant was completely sequenced with oligos 3 to 7 and revealed no nucleotide changes apart from those introduced by site-directed mutagenesis. The TS coding

| Expressed enzyme | Tritium release assay pmol/min/μg | % WT |
|---|---|---|
| pAD768 (WT) | 105.8 ± 10.9 | 100.0 |
| G48A | 50.8 ± 7.1 | 48.0 |
| G48A/48P49/M49P | 32.6 ± 3.3 | 30.8 |
| 58T59 | 73.2 ± 3.7 | 69.2 |
| G48A/48P49/M49P/58T59 | 80.4 ± 3.4 | 76.0 |

Table 3.14. **TS assays.**
Tritium release assays of the wild type VZV TS and variants with mutations in β-strand i. Activities are given as mean ± standard error of the mean. All variants were derivatives of pAD768 expressed in the *E.coli* strain χ2913.

| Codon | Amino acid | Codon | Amino acid | Codon | Amino acid |
|---|---|---|---|---|---|
| CGA | Arg | AAA | Lys | AGC | Ser |
| GGA | Gly | ACA | Thr | AGG | Arg |
| TGA | Stop | ATA | Ile | AGT | Ser |

Table 3.15. **Possible revertants of G48R.**
The nine codons shown above represent all the possible revertants of the mutant G48R using ethyl methanesulphonate.

sequence of each mutant was excised from pAD403 and transferred to the pAD768, as this plasmid has the wild type -10 region of the bacterial promoter sequence (see Section 3.1.3.2). Enzyme activities were determined in crude extracts made from *E.coli* $\chi$2913 cells, and the results are shown in Table 3.14.

3.3.1.2 **Predicted role of G31 in *E.coli* TS**

ß-strand i moves against ß-strand ii during the formation of the ternary complex and is one of the three ß-strands that play host to the so-called ß-kink of the sheet (Matthews *et al.*, 1989, 1990b; Finer-Moore *et al.*, 1990). A feature of this ß-kink is the hydrogen bonding network that starts from the side chain of an arginine residue (R52') in ß-strand i of the second subunit, traverses ß-strands i, ii and iii of the first subunit and ends at the backbone carbonyl group of the active site cysteine residue (C183), resulting in the correct orientation of the side chain for catalysis (Matthews *et al.*, 1989; see Fig. 3.20). When this structure was first described for *E.coli* TS, Matthews and co-workers (1989) proposed that two highly conserved glycine residues, G31 (G48 in VZV) in ß-strand i and G204 (G241) in ß-strand ii must always be found at these positions because the observed dihedral angles are only allowed for glycine (Ramachandran & Sasisekharan, 1968). A mutant of VZV TS, G48R/M49I had previously been identified that failed to complement the growth of the *E.coli* strain 2913 (R. Thompson and J. E. Scott, personal communication). In experiments using EMS to induce revertants that regained the ability to complement the growth of the *E.coli thyA* strain, only enzymes with glycine at residue 48 were found. This shows that the M49I change in itself had little effect on activity, as measured by the TS growth complementation assay.

The G48R mutation was due to a change in the codon from GGA to AGA. Single mutations, induced by EMS, could change this to several alternate codons, resulting in substitution

of a variety of amino acids, but not alanine (see Table 3.15). Of 40 independent complementing revertants isolated, all had a glycine at position 48 (R. Thompson and J. E. Scott, personal communication) suggesting that a residue with a side chain equivalent in size or larger than that of serine cannot be tolerated at this position.

The phylogenetic comparison immediately suggests that there may not be an absolute requirement for glycine at this position as alanine is found in the yeast TS sequences and a serine residue is seen in the bacteriophage ø3T at the equivalent position (see Fig. 3.19).

### 3.3.1.3 **A glycine residue at position 48 is not absolutely essential for VZV TS activity**

On the basis of these observations, it was decided to substitute G48 of VZV TS with alanine. This mutant (G48A) had approximately half the enzyme activity in the tritium release assay compared to the wild-type enzyme (see Table 3.14). This demonstrates that a glycine residue at this position is not absolutely essential for enzyme activity. This is in contrast to the proposal of Matthews and co-workers (1989). The model of VZV TS cannot accomodate an alanine residue at this position due to steric clashes (data not shown). This suggests that the mutant form of TS undergoes some conformational changes at this position to accomodate the methyl side chain of alanine. The changes in the structure that allow this accomodation cannot presently be modelled from the crystal structure of *E.coli* TS or the model of VZV TS due to the lack of access to suitable modelling software. In addition, more recent work by Matthews group (Michaels *et al.*, 1990) has shown that substitution of G204 (G241) of *E.coli* TS with alanine or serine (but none of 10 other amino acids) produces active enzyme, showing that the glycine in ß-strand ii is similarly not absolutely required.

### 3.3.1.4 The additional residue in ß-strand i of yeast TS and a VZV TS mutant is accomodated by a ß-bulge

The yeast sequences contain an additional amino acid immediately adjacent to the glycine in ß-strand i (see Fig. 3.19). As mentioned previously, ß-strand i is the only buried element of secondary structure that has a variable number of amino acids (Hardy *et al.*, 1987; see Fig 1.6 ). The presence of this additional residue may be linked with the replacement of glycine by alanine in this region of the yeast TS enzymes. To investigate this possibility, the mutant G48A/48P49/M49P was constructed. This mutant may serve as a reasonable 'mimic' of the yeast ß-strand i sequence. It bears the closest resemblance to the *S.cerevisiae* sequence, differing by three conserved amino acids (A51L, Y53F and N54S) out of the 13 that make up ß-strand i (see Fig. 3.19). The activity of this mutant (30% of the wild type enzyme) clearly demonstrates that the VZV enzyme can tolerate this additional amino acid (see Table 3.14).

From the structural information available for TS from *L.casei* and *E.coli* it seems unusual that ß-strand i could tolerate an additional amino acid, as this would be expected to interfere with the movement and precise orientation of this element of secondary structure. However, yeast TSs have naturally accomodated an additional residue in the centre of ß-strand i, and we have shown that VZV TS can tolerate an extra residue at this position.

Recent experiments by various workers have shown that several other proteins are capable of accomodating amino acid insertions. Sondek and Shortle (1990) have shown that when glycine or alanine residues were inserted at 20 randomly selected sites in staphylococcal nuclease, insertions only at five sites totally inactivated the enzyme, whilst the majority of insertions reduced activity by approximately two-fold. Many of these tolerated insertions fell within α-helices or ß-strands. Other groups have recorded similar results with other proteins (Barany, 1985, 1987; Tanese & Goff, 1988). The resulting rearrangements

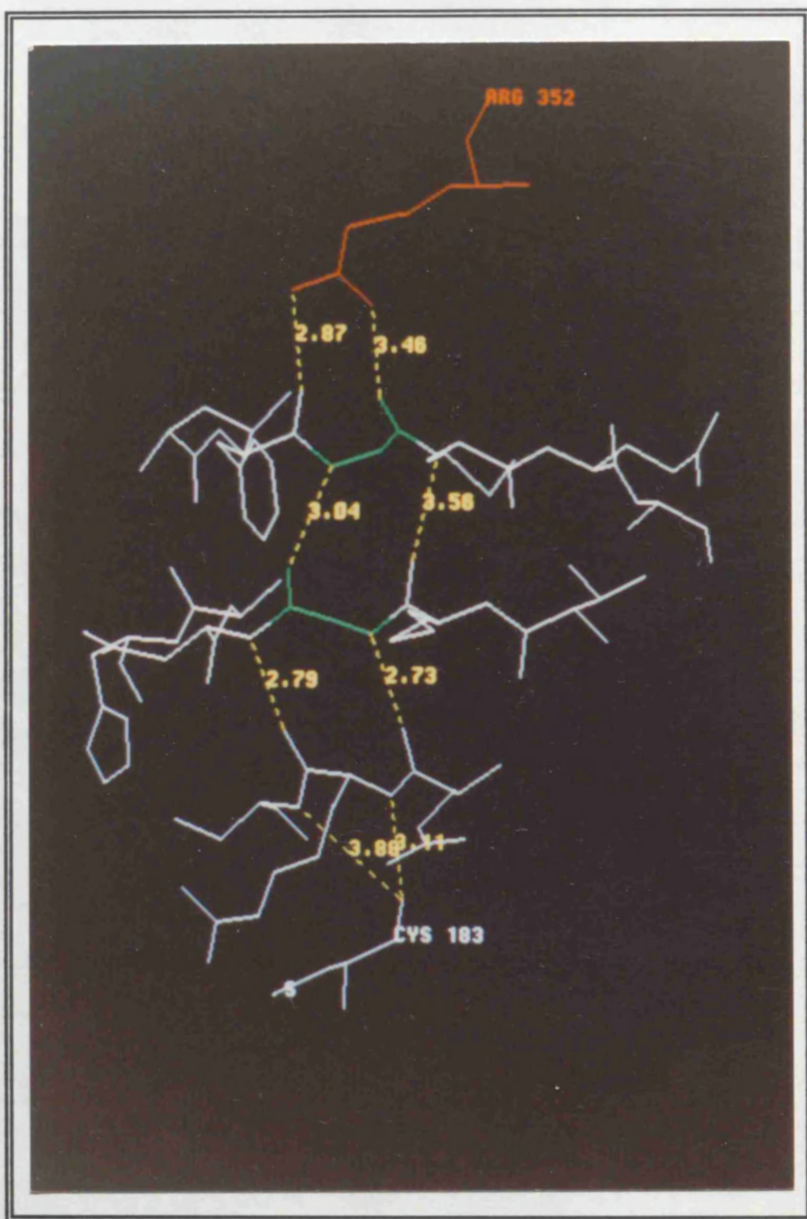



Fig. 3.20. **β-kink of VZV TS model.**
This figure shows the hydrogen bonding network of the β-sheet of TS, with residues from β-strands i, ii and iii running horizontally across the figure. The residue shown in red is R52' from the second subunit. G48 and G241 are shown in green. Hydrogen bonds are represented by yellow dashed lines. The active site residue, C183, is also shown and the sulphur atom indicated by the letter S.

that are made to accomodate these changes have not yet been described. It should be noted for these proteins, the amino acid insertions are tolerated in elements of secondary structure that are not significantly buried. This is in contrast to the insertion of an extra amino acid in ß-strand i of TS, an element of secondary structure which is buried in the core of the protein.

Initially, it was not clear how an additional residue could be accomodated in this region of the enzyme (see Fig. 3.20), however, several observations, described below, are consistent with the formation of a ß-bulge (Richardson *et al.*, 1978). The insertion of an extra residue to ß-strand i of VZV TS should not disrupt the H-bond interactions between the side chain of R52' from the second subunit to the backbone carbonyl groups of residues 47 and 48 of the first (nor the reciprocal interactions between R52 and 47' and 48') as the residue is inserted into ß-strand i of both subunits.

ß-strand i should still be able to move during catalysis if its general structure is maintained, which seems likely for the following reasons. If one half of the strand were to undergo a register shift, caused by the insertion of an extra residue, then the side chains of residues pointing into the core would be reorientated to point into the dimer interface and vice-versa. This would require extensive repacking to accomodate at least five residues, which in itself is unlikely. Moreover, if this was the way in which the yeast enzymes have evolved to accomodate this insertion, then extensive sequence changes would be expected in the regions that pack against the ß-strand i residues; such sequence changes are not apparent (see Fig. 3.19).

Taken together, these observations, and experimental data, suggest that some local accomodation of the additional proline residue takes place. This is most likely achieved by the formation of a ß-bulge, a structure with which proline residues are often associated (Richardson *et al.*, 1978; see Fig. 3.21). Several other

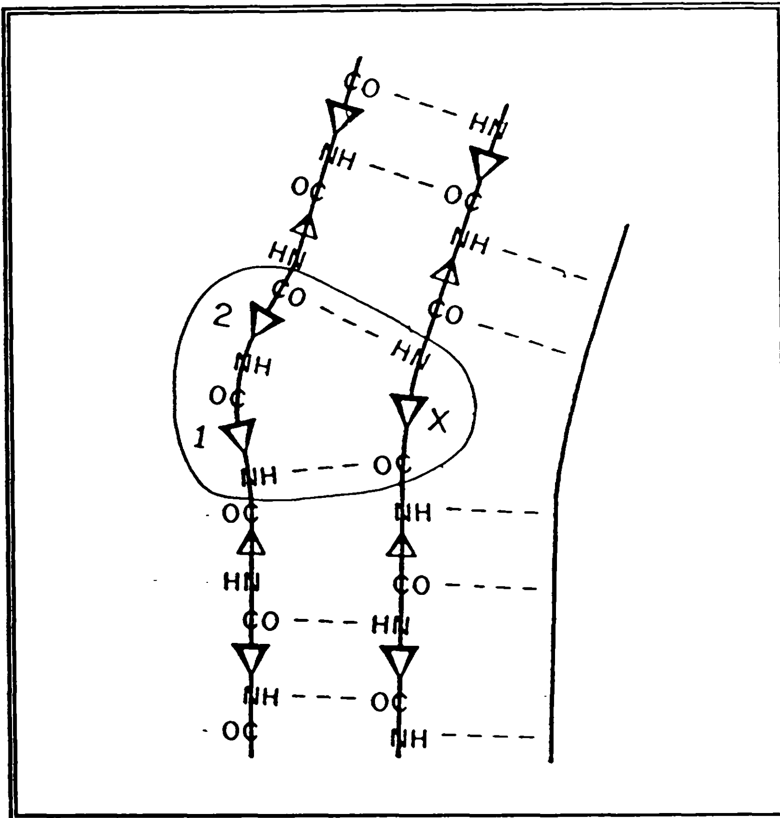



Fig. 3.21. **A classic β-bulge.**
A classic β-bulge in anti-parallel β-sheet. Small triangles represent side chains below the sheet, and larger triangles represent side chains above it (taken from Richardson & Richardson, 1989). Residue number 1 would represent G48 and residue number 2 would be the additional proline seen in the β-bulge of the VZV TS mutant G48P/48P49/M49P. X represents G241 of β-strand ii.

features of ß-bulges are consistent with this proposal. ß-bulges are common in anti-parallel ß-sheets (but rare in parallel ß-sheets); ß-strands i and ii of TS are anti-parallel. ß-bulges are usually located between a close pair of interstrand H-bonds rather than a wide pair; in TS the presence of the ß-kink would mean that the ß-bulge would occur between two intermediary spaced H-bonds. ß-bulges do not induce a register shift of the ß-strand, an observation compatible with the argument that such a conformational change is highly unlikely. In general, ß-bulges are considered the ideal way to accomodate insertions, especially when located at the end or edge of a ß-sheet (Richardson *et al.*, 1978), as is the case for ß-strand i in TS.

The conformation of a ß-bulge requires that the single residue on the opposite strand adopts a similar dihedral angles to those observed for the polyproline conformation (near phi= -60°, psi= 140°). The glycine residue situated at this position in ß-strand ii of TS could readily adopt this conformation (Ramachandran & Sasisekharan, 1968).

#### 3.3.1.5 Identification of a covariant insertion adjacent to the ß-bulge in ß-strand i

Immediately adjacent to the carboxy-end of ß-strand i, an additional amino acid is seen in both yeast (threonine) and *Leishmania* spp. (arginine) TS sequences. To determine whether this insertion is linked to the additional residue in the ß-strand i, two mutants were constructed. The first mutant, containing a threonine between residues 58 and 59 (58T59), reduces activity to approximately 70% of wild type TS when measured by the tritium release assay (see Table 3.14). This suggests that the extra amino acid can be accomodated by VZV TS. This region is part of a surface loop between ß-strand i and helix B so the presence of an additional hydrophilic amino acids, such as threonine or arginine (which are often found on the surface of proteins) could possibly be expected not to be too disruptive.

Recently, the consequences of the addition of extra

residues has been quantitated in a study of single amino acid insertion mutagenesis of staphylococcal nuclease (Sondek & Shortle, 1990). They found that 11 out of 11 glycine and/or alanine insertions at seven different sites in four different surface loops only reduced enzyme activity by two- to five-fold, suggesting that loop regions in general are tolerant to amino acid insertions.

An additional mutant containing the above insertion, together with the ß-strand i mutations (G48A/48P49/M49P/58T59), has approximately 2.5 times the activity of the G48A/48P49/M49P mutant, and slightly more than the 58T59 mutant (see Table 3.14). This data suggests that the two positions are covariant, that is, both may contribute towards the stable accomodation of the other.

#### 3.3.1.6 Summary

These experiments have shown that a glycine residue at position 48 is not absolutely essential for enzyme activity. Also, the additional residue found in ß-strand i of yeast TS sequences can be accomodated in VZV TS. The construction of this latter mutant can be thought of as a crude example of the recruitment of an element of secondary structure from one enzyme to the backbone of a second. Alternatively, it can be viewed as an illustration of the plastic response of the VZV enzyme to insertions in this region. The suggestion that the additional residue in the loop region is covariant with the ß-strand i insertion supports this notion.

### 3.3.2 Carboxy-terminus

The removal of the C-terminal residue of *L.casei* TS by carboxypeptidase has been shown to inactivate the enzyme (Aull *et al.*, 1974). The creation of the equivalent *L.casei* mutant by site-directed mutagenesis has confirmed that the C-terminal residue is essential for catalysis (Climie *et al.*, 1990). This has been presumed to be due to the disruption of H-bonds between the C-terminal carboxyl group and the side chains of W85, R23 and T24, and to the

| **Expressed enzyme** | **Tritium release assay pmol/min/μg** | **% WT** | **Debromination assay mOD/min/μg** | **% WT** |
|---|---|---|---|---|
| pAD768 (WT) | 105.9 ± 10.9 | 100.0 | 0.19 ± 0.01 | 100.0 |
| L301Δ | 3.5 ± 0.6 | 3.3 | 0.16 ± 0.02 | 84.2 |
| E298P | 6.0 ± 0.8 | 5.7 | 0.22 ± 0.07 | 115.8 |
| E298P/ L301Δ | 6.3 ± 0.8 | 6.0 | 0.24 ± 0.03 | 126.3 |
| A300E/ L301Δ | 1.5 ± 0.3 | 1.4 | 0.13 ± 0.01 | 68.4 |

Table 3.16. **TS assays.**
TS assays (tritium release and debromination) for the wild type VZV TS and variants with mutations in the C-terminal region. Activities are given as means ± standard error of the mean. All variants were derivatives of pAD768 expressed in the *E.coli* strain $\chi$2913.

PABA ring of the cofactor (Montfort *et al.*, 1990; Finer-Moore *et al.*, 1990; Matthews *et al.*, 1990b; see Section 1.8). To investigate the nature of these interactions in VZV TS, a number of mutants were constructed in this region. All of the mutants were designed before the VZV TS model became available.

#### 3.3.2.1 **Mutant construction and enzyme activities**

All of the mutants were constructed by site-directed mutagenesis. The C-terminal leucine residue was replaced by a stop codon in pAD403 using oligo AR1 to form L301Δ, whilst A300E/L301Δ was constructed using oligo 69. Mutant E298P was made in pAD403 using oligo AR2, whilst the mutant E298P/L301Δ was made using the same oligo, but using template DNA made from pAD403 containing the L301Δ mutation. The mutants were initially identified by DNA sequencing. The TS gene of each mutant was completely sequenced with oligos 3 to 7 and revealed no nucleotide changes apart from those introduced by site-directed mutagenesis. The TS coding sequence of each mutant was excised from pAD403 and transferred to pAD768, prior to the determination of enzyme activities in crude extracts made from *E.coli* χ2913 cells.

The *L.casei* TS C-terminal deletion mutant was originally shown to be inactive by the tritium release assay (Aull *et al.*, 1974). However, the same mutant has been shown to have essentially wild-type debromination activity (Dan Santi, personal communication). This suggests that the C-terminal interactions are independent of the first stage of the covalent binding of dUMP, and the activation of the C5-H (or C5-Br) bond. For this reason, the activities of all of the mutants were measured by both the tritium release and debromination assays (Table 3.16).

#### 3.3.2.2 **The C-terminal interactions of *L.casei* and *E.coli* TS are essentially identical**

The lack of a crystal structure for the ternary complex of *L.casei* TS has meant that interpretation of mutagenesis

| | |
|---|---|
| VZV | **K**296MEM**A**L301 |
| *L.casei* | **K**311APV**A**V316 |
| *E.coli* | **K**259APV**A**I264 |

Fig. 3.22. **C-terminal sequences.**
Sequence alignment of the six C-terminal six amino acids of TS from VZV, *L.casei* and *E.coli*.

data from *L.casei* TS must be performed on the basis of the equivalent complex of *E.coli* TS. The use of the *E.coli* structure to describe the interactions at the C-terminus of *L.casei* TS is valid due to the extensive structural homology observed between the *L.casei* and *E.coli* native TS structures (Perry *et al.*, 1990). The only sequence difference in the region of interest lies in the C-terminal amino acid of *L.casei* (V316) and *E.coli* (I264). This difference (in molecular terms, a single methyl group) is probably of little significance since the replacement of V316 by several hydrophobic residues (i.e. A/V/L/M), has no deleterious effect on activity, suggesting that this amino acid difference is of little consequence when comparing the structural and functional equivalence of this region of these two enzymes (Climie *et al.*, 1990). For the purposes of the following discussion, all the interactions described for *L.casei* TS, are based on the *E.coli* TS structure.

#### 3.3.2.3 The C-terminal interactions of VZV and *L.casei* TS are significantly different

To assess the role of the C-terminal residue in the VZV TS catalysed reaction, the mutant L301Δ was constructed. The debromination activity of this mutant is essentially the same as the wild type enzyme, suggesting that the C-terminal tail does not play a significant role in the initial stages of catalysis. However, when TS activity was measured by the tritium release assay, this mutant was active, albeit 30-fold lower than WT (see Table 3.16). This is at least 30-fold higher than the activity of the *L.casei* V316 mutant (Aull *et al.*, 1974), suggesting both structural and functional differences between this region of TS. These differences could lie in the C-terminal tail, a region with which the C-terminal tail interacts or a combination of both.

It was decided to investigate the differences in the C-terminal tail. A comparison of the final six amino acids of VZV TS with those of *L.casei* and *E.coli* TS shows four amino substitutions (see Fig. 3.22). The most significant

appears to be the glutamate in VZV TS (E298) which is replaced by proline in *L.casei* TS (P313) and *E.coli* (P263). To investigate the role that E298 plays in the conformation of VZV TS, two mutants were constructed. The mutant E298P was made as a mimic of the *L.casei* and *E.coli* C-terminal tail structures. This mutant shows a 15- to 20-fold decrease in tritium release activity compared to the wild type enzyme, but has essentially wild type debromination activity. The second mutant, E298P/L301Δ, was constructed as a mimic of the *L.casei* I316 mutant. The E298P/L301Δ mutant also shows a 15- to 20-fold decrease in tritium release activity, and a slight increase in debromination activity.

Clearly, the activity of these mutants suggests that neither serves as accurate mimics of the *L.casei* TS C-terminal tail. Moreover, it suggests that the structural differences between VZV and *L.casei* TS that account for the observed differences in activity of their respective C-terminal deletion mutants, do not reside solely in the C-terminal tail.

### 3.3.2.4 Comparison of the C-terminal interactions of VZV and *L.casei* TS at a molecular level

As the results presented above have shown, the C-terminal interactions in VZV TS are different from those found in *L.casei* TS. The availablity of the VZV TS models and the *E.coli* TS crystal structures has allowed us to perform a detailed analysis of the putative interactions made between the C-terminal tail of VZV TS and the rest of the protein, and the C-terminal tail of *E.coli* TS and the rest of the protein. The aim of this section is to describe the identification of the molecular basis of the structural and functional differences observed between VZV and *E.coli/L.casei* TS.

The interactions at the C-terminus of the TS ternary complex from *L.casei* include four H-bonds. These are formed between the carboxylate group of the C-terminal amino acid and the side chains of residues R23, T24 and

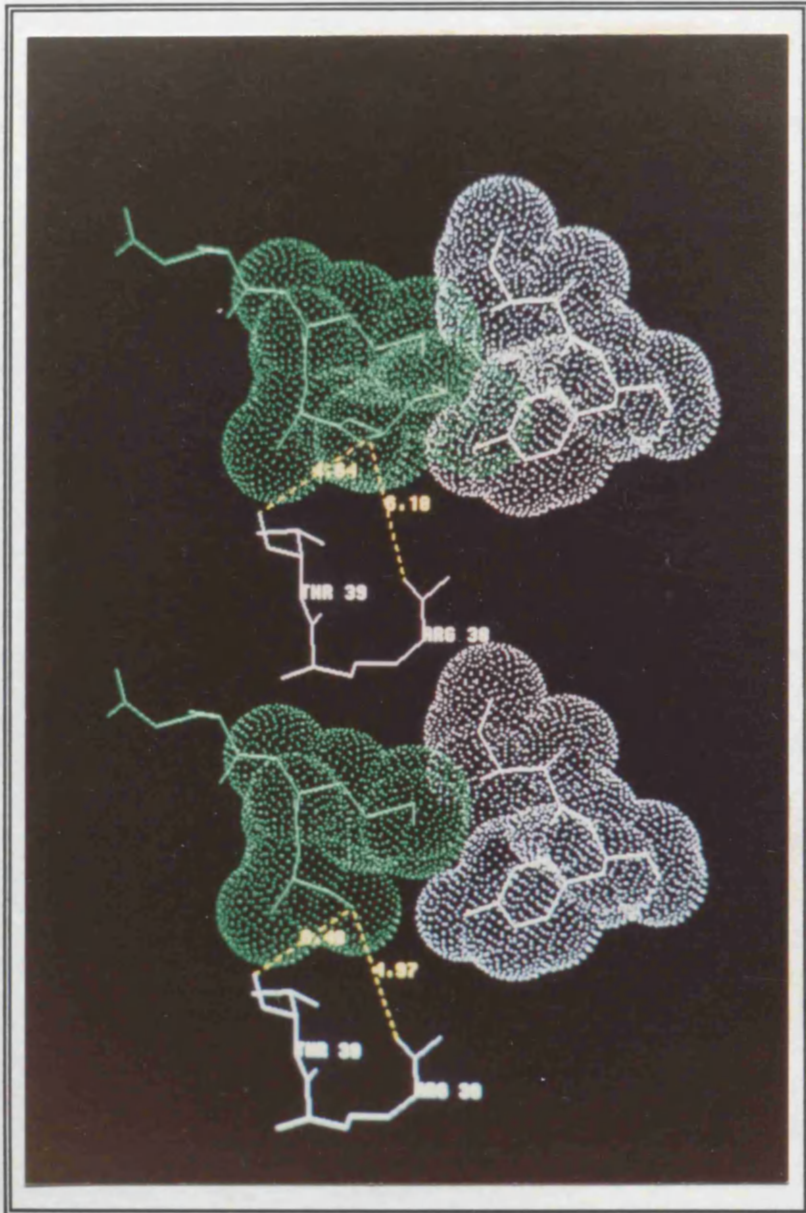



Fig. 3.23. **C-terminal interactions of VZV TS.**
This figure shows the interactions that occur at the C-terminus of VZV TS, with the wild type enzyme is shown at the top and the mutant L301Δ shown at the bottom. Residues E298 to L301 (A300 in the mutant) are coloured in green with van der Waals radii shown for M299, A300 and L301. The carboxylate group of the C-terminal residue is not shown. R38 and T39 are shown in white, with H-bonds shown in yellow. Residues I99 and Y100 are shown in white with van der Waals radii.

W85, and also to the PABA ring of the cofactor (Matthews *et al.*, 1990b; Finer-Moore *et al.*, 1990; Monfort *et al.*, 1990). Our comparison of the model of the VZV TS ternary complex and the TS ternary complex from *L.casei* reveals that the H-bonds formed between the carboxylate group of the C-terminal amino acid and the side chains of residues R38 and T39 (the equivalent amino acids of R23 and T24 in *L.casei* TS) are virtually identical in both enzymes. The H-bonds to the para-aminobenzoicacid ring of the cofactor cannot be modelled by us at present (due to the lack of access to suitable computer software), but the high degree of similarity between the *L.casei* structure and VZV model around the immediate area to which the cofactor binds suggests they will be essentially the same.

However, a H-bond equivalent to the one between the C-terminal residue of *L.casei* TS and W85, is not found in the VZV TS model. This is because the residue equivalent to W85 from *L.casei* TS is Y100 in VZV TS. The hydroxyl group of Y100 (VZV) can act as both a H-bond donor and acceptor. However, we could not position the side chain of Y100 such that its hydroxyl group was in a position to form a H-bond with the carboxylate group of the C-terminal amino acid of VZV TS, without significant disruption to the rest of the protein structure.

On this basis, we predict that a different set of interactions will occur. In VZV TS the side chains of M299 and L301 pack tightly against the side chain of I99 and Y100 forming a strong hydrophobic interaction in the ternary complex. In *L.casei* TS these interactions appear to be much less significant. This is due, in part, to the smaller size of the V314 side chain of the residue V314 in *L.casei* compared to the larger M299 side chain in VZV TS. This suggests that the VZV enzyme is stabilized more by the hydrophobic contacts across the face of TS, between the C-terminus and the D-helix, than the predominantly H-bond contacts observed in *L.casei* (see Fig. 3.23).

### 3.3.2.5 Differences in activity of the *L.casei* and VZV TS C-terminal deletion mutants explained in molecular terms

Our findings may well explain why removal of the C-terminal residue from *L.casei* TS completely abolishes TS activity (when measured by the tritium release assay), but the VZV TS mutant L301Δ retains 3 to 4% activity compared to the wild type enzyme. In the mutant form of both enzymes, weak H-bonds can still be made from the carboxylate group of the newly exposed C-terminal residue (A315 in *L.casei*, A300 in VZV) to R23 and T24 (R38 and T39 in VZV) and possibly to the para-aminobenzoic acid ring of the cofactor. This suggests that the difference in activity is independent of these interactions.

In the *L.casei* TS mutant I316 , the carboxylate group of A315 can no longer form a H-bond to W85 as the distance is too great. As this is the only major observed difference between the wild type and mutant *L.casei* TS, it would suggest that this H-bond interaction may greatly contribute to the stabilization of the ternary complex. This also suggests that its subsequent disruption may be the major factor leading to enzyme inactivation.

VZV TS however, does not form such a H-bond, so the removal of the C-terminal residue (L301) might be expected to invoke a different response. Indeed, we have shown this to be the case, in that the L301 mutant of VZV TS is not totally devoid of TS activity as determined by the tritium release assay (see Table 3.16). This may be interpreted as a decrease in, but by no means an abolition of, the hydrophobic contacts between the C-terminal tail and residue Y100 located in the D-helix.

An alternative hypothesis must also be considered. Although the hydroxyl group of Y100 cannot be modelled in an orientation conducive to H-bond formation in the wild type structure, in the L301Δ mutant, the side chain of Y100 may be able to reposition itself to form a H-bond with the newly exposed carboxyl group, and could stabilize the enzyme.

3.3.2.6 **The C-terminal interactions in VZV TS are unique**

That the interactions found in VZV TS differ greatly from those found in *L.casei* and *E.coli* TS begs the question, "What interactions take place in other TSs?". The answer may not be simply that TS from prokaryotic sources rely on a H-bond whilst TS from eukaryotic sources rely on hydrophobic interactions. Indeed, VZV TS appears to be unique in this respect. A comparison of all the TS sequences at the position equivalent to Y100 reveals the presence of asparagine in phage T4 TS and all the eukaryotic TSs, whilst tryptophan is found at this position in TS from all other organisms. Our examination of the structure of *E.coli* TS has shown that replacement of W83 (W85 in *L.casei*) with Asn allows the formation of a H-bond to the carboxylate group of the C-terminal residue, similar in length to that in wild type *E.coli* TS, with little disruption to the surrounding structure. This suggests that a H-bond could be formed in all other TS enzymes.

We have also found that the replacement of W83 with Asn in *E.coli* TS results in the formation of a gap in this region. However, in all the TS sequences where Asn is found at this position, Met is seen at the antipenultimate position in the sequence (see Fig. 1.6). Modelling shows that the side chain of this residue can fill this gap, suggesting these positions may be covariant (see Section 1.9).

The activities of the mutants E298P and E298P/L301Δ may be explained in terms of this observation. The mutant E298P has approximately 6% of the wild type level of TS activity when measured by the tritium release assay, but essentially wild type levels when measured by the de-bromination assay. This in itself suggests that the C-terminal interactions are different from other enzymes, as TS from HVA and *Tn*4003 has a Pro residue at this position (see Fig. 1.6).

The presence of a Pro residue at this position may make the C-terminal tail curve further in towards the active site due to the restrictions in the conformations that a

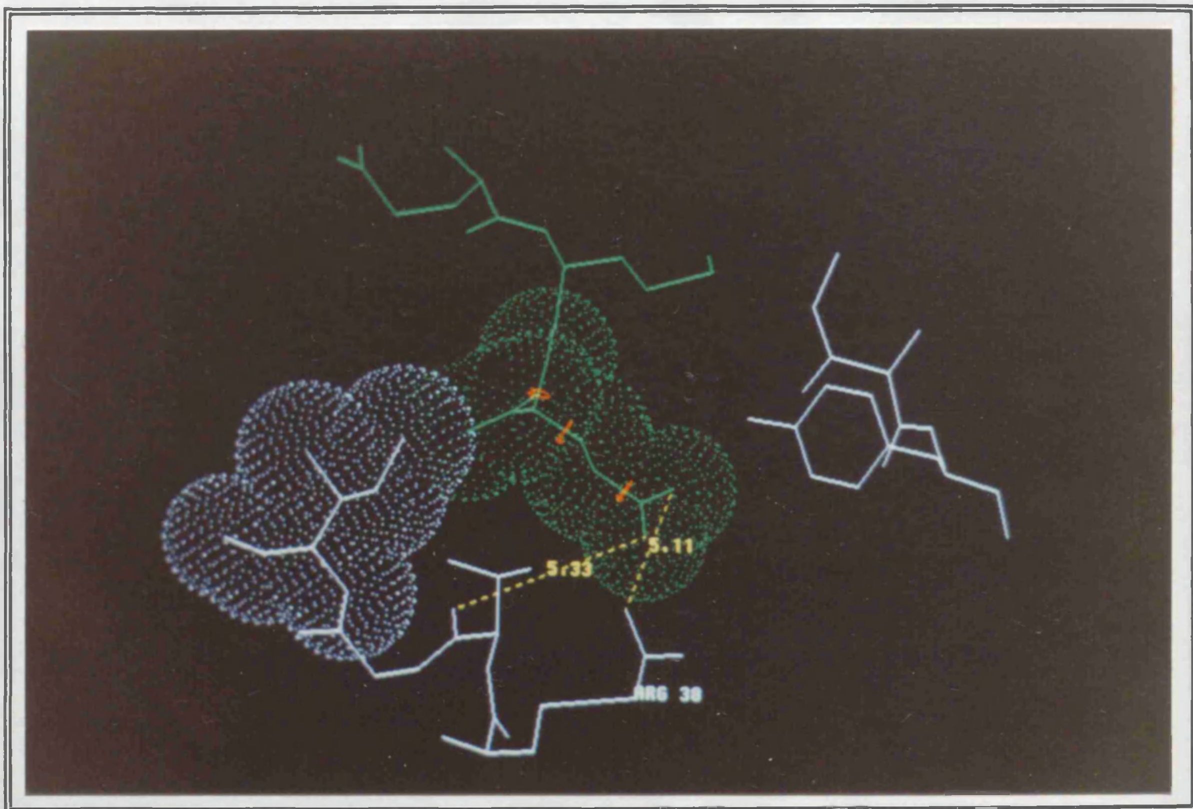



Fig. 3.24. **A300E/L301Δ**
Residues E298 to E300 are coloured in green, with van der Waals radii shown for E300. The red arrows represent rotations about the N-$C_\alpha$, $C_\alpha$-$C_\beta$ and $C_\gamma$-$C_\delta$ bonds of E300 to position the $O_\varepsilon$ atoms of the E300 side chain in an eqivalent position to those of the true carboxylate side chain of L301 in the wild type enzyme (see Fig. 3.22 for comparison). Residues R38 to I41 and I99 and Y100 are shown in white. Van der Waals radii on I41 illustrate the steric clash with the main chain carboxylate group of E300.

Pro residue can adopt. In VZV TS, the Pro residue could re-position the side chain of M299 such that it protrudes into the active site, and may result in steric clashes with the para-aminobenzoic acid ring of the cofactor. This could account for the loss of tritium release activity. This should not disrupt the dUMP binding site, and would therefore account for the wild type levels of debromination activity (see Table 3.16). In TS from other sources where a Pro residue is found at this position, such as HVA, the side chain of M288 (M299 in VZV TS) may be accomodated in the space adjacent to the side chain of N89 (Y100).

An alternative explanation could be that the real structure of VZV TS is slightly different from the model structure. The M299 side chain may lie further out into the solvent, away from the side chain of Y100. When the Glu residue at position 298 is substituted by Pro, this could re-position the Met side chain close to that of Y100 which would cause steric clashes in the active site. In the E298P/L301 mutant, the Pro residue could re-position the side chain of M299 as before. This would explain the similar levels of tritium release activity. In addition, the increased debromination activity could reflect greater access to the dUMP binding site.

#### 3.3.2.7 **TS activity and modelling of the mutant A300E/L301Δ**

This mutant was constructed to see if the side chain carboxylate group of A300E could replace the carboxylate group of the true C-terminus. The low tritium release activity of this mutant (see Table 3.16) suggests that side chain of E300 cannot mimic the C-terminal carboxylate group. In addition, the decreased debromination activity also suggests disruption in this region.

The availability of the VZV TS model has permitted this mutant to be studied in more detail. The side chain of A300E can be positioned such that the its O$\epsilon$ atoms lie in the same position as those of the carboxylate group of L301 in the wild type VZV enzyme (see Fig. 3.24). However, when

placed in this position, a steric clash occurs between the (new) C-terminal carboxylate group of E300 and the side chain of I41. The replacement of I41 with a Gly residue could alleviate the steric clash between these two groups. The construction and characterization of this mutant (I41L/A300E/L301Δ) could be instructive in determining the relative contribution of the hydrophobic and H-bond interactions in the C-terminal region of VZV TS. Similar changes could also be introduced into *L.casei* TS, to further investigate the nature of the interactions that occur between the C-terminus and the rest of the enzyme.

3.3.2.8 **Summary**

These results have shown that the C-terminal interactions of VZV TS are different form those found in *L.casei* TS, and possibly all other known TS enzymes. We have also revealed several examples of covariant accomodation. Many of these observations are directly open to test by several simple mutagenesis experiments on TS from various sources. More importantly, the human TS structure (which has been determined, but not yet published; Bob Stroud and Celia Schiffer, personal communication), will provide an immediate test of the hypothesis that the C-terminal interactions in VZV TS are unique. The ultimate test, the determination of the crystal structure of the VZV ternary complex, and some of the mutant enzymes, is a currently in progress (R. Thompson, personal communication).

## 3.4 Thymidine Kinase

TK catalyses the formation of thymidylate by the phosphorylation of thymidine using ATP as the phosphate donor. The VZV TK gene was of interest for the drug selection protocols used during for the selection of VZV TS variants (see Section 3.1). The gene has been the subject of a mutagenesis experiment, which has lead to the identification of a previously undetected conserved region of the herpesvirus TKs. During the course of these

```
                      SITE 1               SITE 2       SITE 3   SITE 4
                     ********              ******        ***      ***
HSV-1  NH2 -50-  rVYIDGphGmGKTtttqlL -13- EPMtYW -70- lifDRHPiaallcYPaary -38-
HSV-2  NH2 -50-  rVYIDGphGvGKTttsaqL -14- EPMtYW -70- lvfDRHPiasllcYPaary -38-
VZV    NH2 -13-  rIYLDGayGiGKTtaaeef -15- EPLsYW -72- mlsDRHPiasticFPlary -38-
MHV    NH2 -11-  rVYLDGphGvGKSttaeaL -15- EPMaYW -75- lvvDRHavasmvcYPlarf -38-
PRV    NH2 -04-  rIYLDGayGtGKSttarvM -09- EPMaYW -66- vvfDRHPvaatvcFPlarf -38-
FHV    NH2 -21-  rIYIDGayGiGKSltakyL -15- EPMlYW -69- liiDRHPlaslvcFPlary -38-
EHV-1  NH2 -26-  rIYLDGvyGiGKSttgrvM -14- EPMaYW -69- vvfDRHPvasavcFPaary -38-
EHV-4  NH2 -26-  rIYLDGvyGiGKSttgrvM -14- EPMaYW -69- lvfDRHPvastvcFPaary -38-
BHV-1  NH2 -11-  rIYLDGahGlGKTttgraL -14- EPMaYW -79- lvfDRHPvaaclcYPfary -38-
BHV-2  NH2 -05-  rVYVDGphGlGKTtaasrL -11- EPMsYW -70- lifDRHPtasllcYPlary -38-
ILTV   NH2 -30-  lLYVDGpfGvGKTvtaktL -14- EPMqaW -74- flvDRHPlaaclcFPvaqy -38-
MDV    NH2 -20-  rVYLDGsmGiGKTsmlneL -14- EPMkYW -67- lilDRHPisatvcFPiarh -39-
HVT    NH2 -11-  rVYLDGpfGiGKTsilnaM -14- EPMkYW -67- lmlDRHPvaailcFPitry -39-
EBV    NH2-285-  sLFLEGapGvGKTtmlnhL -12- EPMrYW -66- ilhDRHllsasvvFPlmll -37-
HVS    NH2-210-  fIFLEGsiGvGKTtllksM -13- EPIaYW -63- vmfDRHPlsatvvFPymhf -37-
CCV    NH2 -17-  vfcVEGniGcGKStlvkaL -13- EPVdqW -40- dimERsPmsatrvFcavng -43-
• dCK  NH2 -22-  kIsIEGniaaGKStfvniM -11- EPVarW -71- vysDRyifasn.lYesecm -40-

       SITE 5          SITE 6
       *******         ******                            *
HSV-1  RlakRqRpGE -56- phigdTLFtlfrapellapnGdlyn.vfawaldvLak.rLrpMhVfiLdydqsp.agc -40- COOH
HSV-2  RlarRqRpGE -55- priedTLFalfrvpellapnGdlyh.ifawvldvLad.rLlpMhLfvLdydqsp.vgc -39- COOH
VZV    RvskRaRpGE -58- pgiedTLFavlklpelcgefGnilp.lwawgmetLsn.csrsMspfvLsleqtp.qha -36- COOH
MHV    RlraRaRpGE -61- pgladTLFaalkvpefldarGypra.ahawtldiLan.rIraLrVytLdltgpp.eac -52- COOH
PRV    RlraRaRaGE -59- pelqdTLFgaykapelcdrrGrple.vhawamdaLva.kLlpLrVstvdlgpspr.vc -35- COOH
FHV    RlrgRsRtGE -59- sdlcdTLFsvfkarelsdqnGdlld.mhawvldgLme.tLqnLqIftLnlegtp.dec -32- COOH
EHV-1  RlrtRaRiGE -59- pelsdTLFamfktpellddrGvile.vhawaldaLml.kLrnLsVfcadlsgtpr.qc -37- COOH
EHV-4  RlrtRaRiGE -59- pelgdTLFalfktqellddrGvile.vhawaldaLml.kLrnLnVfsadlsgtpr.qc -37- COOH
BHV-1  RlaaRaRpGD -61- palrdTLFaalkcrelypggGaglpavhawaldaLag.rLaaLeVfvLdvsaap.dac -46- COOH
BHV-2  RlvaRgppGE -53- pgghnTLLalihga..gatrGcaam..tswtldlLad.rLrsMnMftwt.tarprslc -23- COOH
ILTV   RiiqRgRpGE -61- psidqTLFailafdq.qnvhGerlktvlsfvvqkLat.vLknLcIfyLpahgltpeac -37- COOH
MDV    RlssRnRtGE -65- vslhhTLLaifkrrelcaedGslst.thawilwgLlm.kLrnInVerfnitglsttkc -37- COOH
HVT    RicsRdRpGE -66- mpldrTLLaifkrkelcsenGellt.qyswilwgLlt.kLhtInVelfdisgmsrrec -44- COOH
EBV    RlkkRgRkhE -61- klyknSIFsvlkeviqpfradavllevclaftrtLaylqfvlVdLsefqddlpgcwte -34- COOH
HVS    RvkkRnRkeE -61- dmlksSIFntwiemtkahrdsctlmeclltfckeLekvqLihVnVspftddipglwas -31- COOH
CCV    RmrrRdRtGE -61- COOH
dCK    RiylRgRneE -63- COOH
```

**Fig. 3.25. Comparison of the herpesvirus TK sequences.** Alignment of the amino acid sequences of six conserved sites in 15 higher vertebrate herpesvirus TKs and comparison with CCV TK and human dCK. **BOLD** UPPER CASE letters are shown for the 15 TKs when all 15 amino acids are identical at that position. UPPER CASE letters are shown for the 15 TKs when at least 13 have the same amino acid at that position, and when at least 14 TKs possess a residue in one of the following groups: (I/L/M/V) (D/E) (S/T) (F/Y) (F/L) (K/R). Upper case letters are shown for CCV TK and dCK when residues are identical to, or in the same group as, upper case residues in the 15 TKs. The six conserved sites proposed by Balasubramanian *et al.*, (1990) are indicated. Dots represent spacing characters introduced to produce the alignment. The abbreviations used are as shown in the abbreviations lists. The references to the sequences can be found in Harrison *et al.* (1991).

studies, the evolutionary source of the herpesvirus TKs was also established (Harrison *et al.*, 1991).

### 3.4.1 Site-directed mutagenesis of VZV TK

The TK gene of the Dumas strain of VZV has previously been cloned and expressed at high levels in the *E.coli* strain C600*tdk*, however, no detectable TK activty was found in crude extracts of these cells (McKee *et al.*, 1990). Our comparison the amino acid sequence of TK from the Dumas strain (Davison & Scott, 1986) with the TK gene of several other VZV strains, revealed a single amino acid substitution (corresponding to a single nucleotide change) with serine (TCA) at position 288 in place of leucine (TTA) (Sawyer *et al.*, 1988; Mittal & Field, 1989). Leucine is highly conserved at this position, being replaced only by isoleucine in MHV (see Fig. 3.25). This suggested that the amino acid difference at this position of the TK gene of the Dumas strain was the cause of the observed lack of TK activity. On this basis, it was decided to replace the serine residue at positon 288 with leucine by site-directed mutagenesis.

#### 3.4.1.1 Mutant construction and enzyme activities

The TK gene of the Dumas strain of VZV was subcloned from the vector pTM9 into pBS+ and pBS- to create the plasmids pAD507 and pAD601 respectively. The yields of ssDNA and ratio of ssDNA to helper phage were high for both vectors. Site-directed mutagenesis was used to create the mutant S288L using oligo 52 and U-ssDNA made from the vector pAD507, with DNA transformed into the *E.coli* strain C600*tdk*.

Transformants were screened by the TK growth complementation assay and approximately 30% grew. Of six complementing clones that were sequenced with oligo 53, all contained a single nucleotide substitution (T863C) corresponding to the replacement of S288 with leucine.

One of these clones, S288LTK+5 was used to re-transform *E.coli* C600*tdk* cells and evaluated for growth complement-

| Source of extract | Specific Activity (pmol dTMP/μg protein | |
|---|---|---|
| | Extract 1 | Extract 2 |
| C600*tdk* | 0.009±0.0 | 0.010±0.0005 |
| C600*tdk*/pAD507 | 0.011±0.0005 | 0.010±0.001 |
| C600*tdk*/pVZVTK+5 | 1.583±0.049 | 0.382±0.012 |

Table 3.17. TK assays.

The TK activities of two independent extracts of a mutant and potential wild type variant of VZV TK. The variants were expressed in *E.coli* C600*tdk*. Activities are given as the mean ± the standard error of the mean.

ation ability. This clone gave complementation, in contrast to pBS+ and pAD507 which repeatedly failed to complement the growth of this strain. Enzyme activities were determined for the S288LTK+5 and pAD507 clones and compared to the cells only activity as shown in Table 3.17.

#### 3.4.1.2 **Summary**

Substantial TK activty is seen in both extracts of the S288LTK+5 clone compared to the cells alone, and the Dumas derived clone. Although the second extract has approximately four-fold less activity than the first, the TK activity is still significantly greater than the background levels.

The active TK gene of S288LTK+5 was subcloned into the vector pCB105 to create the plasmid pCBTK3. This plasmid, was used in the drug screening protocols (see Section 3.1).

The genetic selection described in Section 2.12.1.2, was based on the enrichment procedure used for the isolation of *E.coli thyA* mutants. In the presence of trimethoprim *E.coli* cells fail to grow due to the inhibition of *de novo* thymidylate synthesis (see Section 1.13.2). However, in the presence of exogenous thymidine, *E.coli* cells expressing an active form of TK should be able to catalyse the formation of thymidylate from thymidine, thus overcoming the toxic effects of trimethoprim. Bacteria that fail to express an active form of TK are unable to utilise exogenous thymidine and thus do not grow. The data suggest that this selection scheme was completely successful.

The role of L288 in VZV TK is unknown, but this amino acid lies within a region that contains a number of highly conserved hydrophobic residues. A recently published model of the HSV-1 TK structure places this hydrophobic region in close proximity to the active site of the enzyme (Folkers *et al.*, 1991), although the function of this region remains unknown.

### 3.4.2 **Evolution of the herpesvirus thymidine kinase**

Herpesviruses encode a variety of enzymes that catalyse

| source | Mr (kDa) | subunit composition | phosphate donors | natural substrates | alternative substrates |
|---|---|---|---|---|---|
| herpesvirus TKs | 25-60 | dimer | ATP CTP | TdR, (CdR), (dTMP) | ACV, araA |
| mammalian TKs | 25-27 | tetramer | ATP | TdR | - |
| poxvirus TKs | 20-21 | tetramer | ATP | TdR | - |
| *E.coli* & T4 TK | 22-23 | trimer/ tetramer | ATP | TdR | - |
| mammalian TmpKs | 24-25 | - | ATP | dTMP | - |
| poxvirus TmpKs | 25-26 | - | ATP | dTMP | - |
| dCK | 31 | dimer | ATP GTP | CdR, AdR, GdR | ddC, araC |

Table 3. 18. **Properties of TK.**

The properties of TKs and similar enzymes from various sources (see text for details and references).

reactions involved in nucleotide metabolism, including thymidine kinase (see Section 1.2.3). Many of these enzymes show significant amino acid homology to their cellular counterparts, suggesting they were captured by an ancestral herpesvirus at some point during evolution. However, the herpesvirus TK is clearly not related to its cellular counterpart, and its evolutionary source has so far remained obscure.

### 3.4.2.1 Herpesvirus TKs are genetically and biochemically distinct from their cellular counterparts

The sequences of the TK gene of at least 15 different herpesviruses have been reported to date, making it, in sequence terms at least, the most studied herpesviral gene. Indeed, the wealth of information provided by these sequences has prompted the construction of several evolutionary trees for the Herpesviridae (Gentry *et al.*, 1988) and lead to much speculation on the acquisition of this gene (Gentry *et al.*, 1983; Robertson & Whalley, 1988).

In addition to these phylogenetic studies, several groups have found highly conserved amino acids (Robertson & Whalley, 1988; Honess *et al.*, 1989), which have recently been grouped into six conserved 'sites' (Balasubramaniam *et al.*, 1990). The functions of these sites have been reviewed in detail elsewhere (Balasubramaniam *et al.*, 1990; Harrison *et al.*, 1991). Several comparisons between the sequences of cellular TKs and those of herpesvirus TKs have strongly suggested that they have no relationship, apart from a region common to many different classes of ATP binding proteins (Saraste *et al.*, 1990). This is in marked contrast to the poxviruses which encode TKs that share a clear amino acid similarity with cellular TKs, and have very similar substrate specificities and physical properties. Indeed, these conserved features have led to the proposal that poxvirus TKs evolved from the cell cytosolic TKs (Kwoh & Engler, 1984; Boyle *et al.*, 1987 and references therein).

This lack of amino acid similarity also extends to a

lack of common biochemical features which are summarized in Table 3.18. Unlike cellular TKs, the herpesvirus TKs can utilize a broad range of substrates in addition to thymidine. Many herpesvirus TKs can phosphorylate deoxycytidine and a number of nucleoside analogues including acyclovir (ACV) and (*E*)-5-(2-bromovinyl)-2'-deoxyuridine (BVdU; Fyfe *et al.*, 1978; Cheng *et al.*, 1981; Larder *et al.*, 1983). Moreover, several herpesvirus TKs also possess thymidylate kinase activity (Chen *et al.*, 1979). These combined activities play a crucial role in the activation of certain nucleoside analogues (Fyfe, 1981; reviewed in Larder & Darby, 1984).

In addition to their broad substrate specificity, herpesvirus TKs, like mitochondrial TKs (and in contrast to cytosolic TKs), can use CTP as well as ATP as a phosphate donor (Kit, 1985; Kit *et al.*, 1974). However, the herpes-virus (and cell cytosolic) TKs are not inhibited by dCTP, which differentiates them from the mitochondrial TKs which are inhibited by dCTP (Kit, 1975; Kit *et al.*, 1974). To the best of our knowledge, no mitochondrial TK sequences have been reported to date.

The herpesvirus TKs also differ in their subunit composition, being homodimeric enzymes, whereas the cell cytosolic enzymes consist of four identical subunits (Kit, 1985; Kit *et al.*, 1974). The subunit composition of the mitochondrial enzymes has not been reported (Kit, 1985). Thus, on the basis of these genetic and biochemical differences, the herpesvirus TKs form a distinct group whose evolutionary source remains obscure.

When reporting the sequence of EHV-1 TK, Roberston and Whalley (1988) suggested that herpesvirus TKs may have originated from the cellular thymidylate kinase (TmpK) gene on the basis of local sequence homology at both the ATP-binding and proposed nucleotide-binding regions (Darby *et al.*, 1986). Indeed, the homology at both sites is reasonably high (although homology at the ATP-binding site is not necessarily indicative of homology between TK and TmpK). Moreover, several herpesvirus TKs possess TmpK

activity. However, the homology at other regions of the protein is very low, and the fact that the physical and biochemical properties of TmpK bear little resemblance to those of herpesvirus TKs suggests that **TmpK may** not be the true evolutionary source of the herpesvirus TKs.

### 3.4.2.2 Channel catfish virus TK: the most divergent herpesvirus TK sequence known?

The herpesvirus TK sequences reported to date are encoded by viruses that infect mammals or birds, however information has recently become available on the TK of channel catfish virus (CCV). This enzyme can be distinguished from channel catfish TK and, in contrast to other herpesvirus TKs, cannot use CTP as a phosphate donor. The gene encoding it has been located to within 3kb by marker rescue of a TK deficient CCV mutant (L. A. Hanson, R. L. Thune & K. G. Kousoulas, personal communication). The DNA sequence of the complete CCV genome has recently been determined (Davison, 1992) and shows that the 3kb region including the TK gene contains three complete ORFs. One of these was found to contain the amino acid sequence, GNIGCGKS, which perfectly matches the consensus sequence GXXGXGK(S/T) for the ATP-binding region of a TK enzyme (Saraste *et al.*, 1990). Computer analysis using standard GCG homology search programs (FASTA and GAP), revealed no strong homology with either the herpesvirus, poxvirus or cell cytosolic TKs, suggesting it may represent a novel form of TK (Andrew J. Davison, personal communication).

However, by close inspection of the CCV TK sequence, we have been able to show that it bears significant sequence similarity to five of the six conserved sites recently described for herpesvirus TKs (Balasubramaniam *et al.*, 1990; Harrison *et al.*, 1991; see Fig. 3.25). This analysis was performed using the GCG manual alignment program, LINEUP. Several herpesvirus TK sequences were aligned using this program, based on a previous alignment for herpesvirus TKs (Balasubramaniam *et al.*, 1990). The CCV TK sequence was then aligned using the highly conserved amino

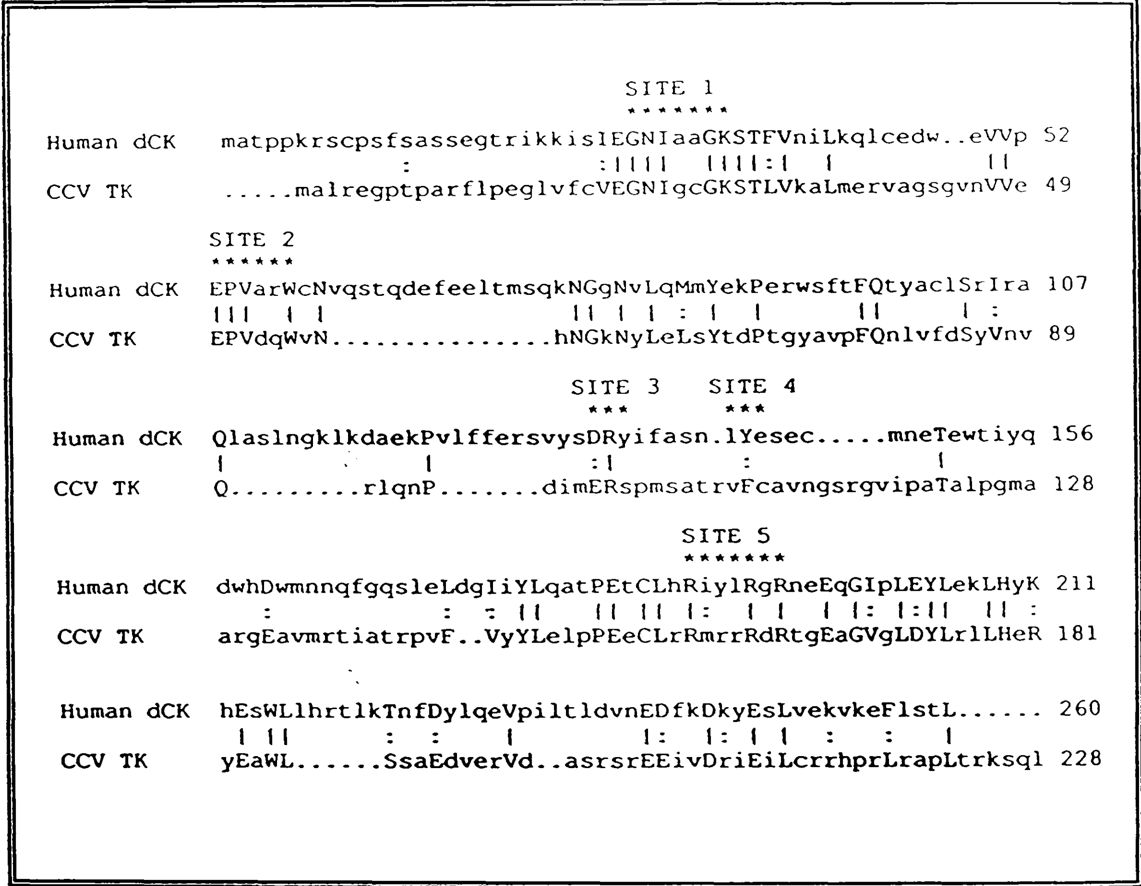



**Fig. 3.26. Comparison of human dCK and CCV TK.**
Alignment of entire human dCK and CCV TK sequences. Vertical bars (|) indicate identical amino acid residues and colons (:) indicate similar residues. Residues are numbered on the right. Dots represent spacing characters introduced to produce the alignment. Upper case letters are used as described in the legend to Fig. 3.25.

acids in the other herpesvirus TK sequences as a reference for the alignment of the residues in the CCV TK sequence. This method of alignment is similar to those described by Argos and co-workers (1991). Similar analysis could not detect any sequence similarity to the cell cytosolic or poxvirus TKs (data not shown). This strongly suggests that the CCV TK is clearly related to herpesvirus TKs. However, the lower degree of sequence similarity, the lack of one of the six conserved sites due to the shorter length (by 73 to 105 amino acids), and the inability to use CTP as an alternate phosphate donor suggest that it is also the most divergent herpesvirus TK studied to date (Harrison *et al.*, 1991).

### 3.4.2.3 Evolutionary source of herpesvirus TK is the cellular deoxycytidine kinase

During the course of our analysis of the proposed CCV TK gene, the sequence of the human deoxycytidine kinase (dCK) was reported by Chottiner and co-workers (1991). An earlier report of the sequence of the human dCK gene (Huang *et al.*, 1989), has since been discounted as the predicted protein product is approximately twice the size of the purified dCK (Chottiner *et al.*, 1991). Moreover, Huang and co-workers have recently concluded that the gene they sequenced represents the human counterpart of the murine ERp72 (luminal endoplasmic reticulum protein; Mazzarella *et al.*, 1989; Huang *et al.*, 1991). As several herpesvirus TKs have been shown to possess dCK activity, we decided to compare the amino acid sequence of herpesvirus TKs with that of human dCK. Somewhat suprisingly, amino acid homology was detected at five of the six conserved sites mentioned, by comparing the human dCK sequence to the herpesvirus TK sequences by the method described above. However, a comparison between the amino acid sequences of dCK and CCV TK, revealed a much higher degree of amino acid similarity. 52 amino acids are identical from a common core of 213 amino acids, and a further 22 amino acids represent conserved changes (see Fig. 3.25). Again many of

these amino acids can be assigned to five of the six sites found in the other herpesvirus TKs. In addition to the homology to these sites, the spacing between sites is also well conserved, strengthening the argument for an evolutionary relationship.

Several of the physical and enzymatic features of dCK are also similar to the herpesvirus TKs. Human dCK, like the herpesvirus TKs is a homodimeric enzyme with a broad substrate specificity which includes purine nucleosides and several cytidine analogues. In addition, both CTP and GTP, as well as ATP, can act as phosphate donors (Chottiner *et al.*, 1991).

Although it is conceivable that the situation has arisen by convergent evolution, on the basis of the strong sequence, physical and enzymatic similarities shown between the human dCK and herpesvirus TKs, it is highly likely that the higher and lower vertebrate TKs have evolved from a captured cellular dCK gene. In developing the ability to phosphorylate thymidine, the enzyme specified by the captured gene in most cases has retained the ability to phosphorylate deoxycytidine, thus the proposed origin of herpesvirus TKs is consistent with their distinctive biochemical properties (Harrison *et al.*, 1991).

# 4. FINAL DISCUSSION

## 4.1 Drug resistance studies

The Discussion outlines the original aims and contrasts them with our findings. It also describes some future avenues for research.

### 4.1.1 Original aims and rationale

The original aim of these studies was to isolate and characterize a variant of VZV TS that could catalyse the normal formation of dTMP in the presence a concentration of the drugs FdUMP or EYdUMP that prevents the wild type enzyme from catalysing this reaction. This was attempted by the mutagenesis of VZV TS *in vitro*, in combination with a suitable screening protocol.

Our initial studies were constrained by the lack of information. Previous work in this field had given no indication as to which amino acid substitutions might give rise to a drug resistant variant (Jastreboff *et al.*, 1983; Heidelberger *et al.*, 1960a,b). In addition, the structural data pertaining to TS was limited to the crystal structure of the native enzyme from *L.casei* at 3.0 Ångströms resolution (Hardy *et al.*, 1987). Although molecular modelling had allowed the identification of a few putative contacts between the protein and ligands, the majority were not known at this point (Hardy *et al.*, 1987).

On this basis, it was decided that a site-directed mutagenesis approach would be too restrictive and a random mutagenesis approach was adopted. Also, as many amino acids, from several different regions of the protein, were predicted to line the active site cavity of TS (Hardy *et al.*, 1987), it was decided to mutagenize the complete gene, rather than to target the mutagenesis to a few selected regions.

### 4.1.2 A drug resistant variant of VZV TS requires more than one nucleotide substitution

In screening a library of all the possible variants of

the VZV TS gene that contain a single nucleotide substitution, we failed to isolate a variant of VZV TS that was resistant to FUdR or EYUdR. This suggests that more than one nucleotide change may be necessary to yield a drug resistant variant.

This observation is in contrast to those relating to the genetic basis of drug resistance in other herpesvirus encoded enzymes. For herpesvirus TK, several variants of both HSV and VZV have been isolated where the enzyme shows wild type levels of TK activity but reduced drug kinase activity (Lacey *et al.*, 1991; Larder *et al.*, 1983). In addition, a single nucleotide change, which results in the same amino acid substitution in the DNA polymerase of both HSV-1 (N815S) and VZV (N779S), has been shown to confer resistance to ACV (Gaillard *et al.*, 1990; Matthews *et al.*, 1989b).

However, not all instances of drug resistance are the result of a single nucleotide change, as is seen in HIV. Some of the most potent inhibitors of HIV reverse transcriptase (RT) include the nucleoside analogues 3'-azidothymidine (AZT) and dideoxyinosine (ddI) (St. Clair *et al.*, 1991). Studies of AZT and ddI resistance by Larder and co-workers (1989; 1991; St. Clair *et al.*, 1991), have shown that more than one nucleotide substitution is necessary before HIV aquires a drug resistant phenotype. In fact a critical step in acquiring an AZT resistant character may be the occurrence of a double nucleotide mutation (**ACC** to **TTC** or **ACC** to **TAC**) that gives rise to the T215F or T215Y changes, thought to be essential for AZT resistance (Larder *et al.*, 1989). This then allows other mutations to increase the levels of drug resistance (Larder *et al.*, 1989; 1991; St. Clair *et al.*, 1991).

### 4.1.3 Possiblities for future studies

We have shown that more than one nucleotide substitution in the VZV TS gene is necessary to give rise to a drug resistant variant. We have also shown that the use of the same protocols to screen a library of *all* the possible

variants of the VZV TS gene that contain two nucleotide substitutions would approach the limits of feasibility, and that screening for variants with three and four nucleotide substitutions would dramatically exceed these limits. This poses two questions. What would be the value of the characterization of a drug resistant TS, and what would be the best approach to isolate such a variant?

The isolation and characterization of a drug resistant variant could increase our understanding of the mechanism of the binding of various inhibitors, and reveal the role that different amino acids play in the recognition and/or binding of drug to TS. This suggests that the isolation of a drug resistant variant should be pursued.

Several recent findings should assist in this process. The determination of the structure of the native enzyme and ternary complex of *E.coli* TS at high resolution, has lead to the identification of the amino acids in contact with the natural substrate, dUMP, and the TS inhibitor, FdUMP, together with a more thorough understanding of the conformational changes associated with catalysis. Such information could be used to identify smaller regions of the enzyme that could be targeted for random mutagenesis. This would dramatically reduce the number of clones that would have to be screened.

Because of the unpredictable effects of site-directed mutagenesis, the structural similarity between dUMP and FdUMP, and the fact that differences exist between VZV and *E.coli* TS, an attempt to design a drug resistant variant *de novo*, using the VZV TS model, would probably be too ambitious at this stage. However, the determination of the crystal structure of VZV TS is in progress (R. Thompson, personal communication). In addition, it may be easier to exploit the structural differences between drug EYdUMP and dUMP. EYdUMP has an ethynyl group at C5 compared to the hydrogen at C5 of dUMP. By molecular modelling, the conformation of the 5-ethynyl group of EYdUMP could be determined. If for example, the 5-ethynyl group lies close to an amino acid with a small side chain, then the sub-

stitution with an amino acid with a large side chain may decrease the affinity of TS for EYdUMP. If this amino acid substitution did not affect enzyme activity, it would confer drug resistance to TS. These ideas may be explored in the future, as the effects of amino acid substitutions become better understood, and protein design skills improve.

### 4.2 VZV TS as a possible target for antiviral chemotherapy: identification of an exploitable difference for rational drug design

During the course of these studies, we have identified a region of VZV TS that differs structurally in TS from other sources, possibly including human TS (see Section 3.3.2.6). The structural difference between VZV TS and TS from other sources resides in the interactions of the C-terminal tail. As this part of the enzyme is involved in catalysis, specifically in interactions with the cofactor, it opens the possibility for the rationale design of differential TS inhibitor. Such an inhibitor of TS is one which inhibits TS from one source with a much greater potency than that from a second source.

If these structural differences between the human and VZV TS are confirmed, the exploitation of them could lead to the development of novel anti-varicella drugs. To this end, the crystal structure of human TS has been determined, although not yet published (Schiffer *et al.*, 1991; Robert Stroud, personal communication). The folate analogue, pteroyl-hexaglutamate, inhibits TS from bacteriophage T2 with a 100-fold greater potency than TS from *E.coli* (Maley *et al*, 1979c) suggesting that it may be possible to achieve similar discrimination between human and VZV TS.

On the basis of the fact that the differential inhibitor of *E.coli* and bacteriophage T2 TS was a folate analogue (Maley *et al.*, 1979c), it could be expected that a suitable differential inhibitor of VZV and human TS would also be a folate analogue. Indeed, several folate analogues are

currently being evaluated for their ability to inhibit the purified VZV and human enzymes, and the growth of VZV in tissue culture (R. Thompson and M. J. Hutchinson, personal communication).

### 4.3 Plastic and covariant accomodation: the ability of proteins to tolerate amino acid substitutions

Plastic and covariant accomodation have been briefly discussed in Section 1.9. This section will outline some of the background that led towards a better understanding of these phenomena and show how our findings are consonant with, and enhance, this understanding.

#### 4.3.1 Globins, *lac* and lambda repressors, and T4 lysozyme

Many of the currently held ideas pertaining to the ability of a protein to tolerate amino acid substitutions emanate from the studies by Perutz and colleagues on the relationship between haemoglobin amino acid sequences and the known 3-D structures (Perutz *et al.*, 1965; Perutz & Lehman, 1968). They noted that, in general, residues which lie at the surface of the protein can have either polar, non-polar or charged amino acids, and can be freely interchanged with a variety of other amino acids without a deleterious affect on the structure and function, except where the amino acids are in direct contact with ligand. However, buried residues, which almost invariably have non-polar side chains, are considerably more sensitive to replacement, with polar or charged residues not being found as naturally occurring substitutions. In addition, they found that substitutions that affect subunit contacts can also be disruptive. Subsequent studies by several groups have reaffirmed these observations for other protein families including the cytochromes, serine proteases and lysozymes (Dickerson, 1971; Hampsey *et al.*, 1988; Bashford *et al.*, 1987; Lesk & Chothia, 1980; Chothia & Lesk, 1987).

With the advent of recombinant DNA technology, the

experimental evaluation of these theories was opened up, led by the pioneering work of Miller and colleagues (Miller *et al.*, 1979; Miller & Schmeissner, 1979; Miller, 1979). They utilized different *E.coli* suppressor strains to incorporate a variety of amino acids at defined nonsense codons of the *E.coli lac* repressor. This technique is very powerful as it allows a single mutation in the gene (i.e. the nonsense codon) to generate many different amino acid substitutions in the protein. The more suppressor strains that become available, the more powerful the technique.

In these early studies, they described the suppression of a large number of nonsense mutations of the *E.coli lac* repressor and found that of 323 amino acid substitutions analysed (at 90 different positions), 187 (58%) were found to be phenotypically "silent", that is, they result in proteins that are not measurably altered in their activity compared to the wild type protein.

A breakdown of these results revealed that 43 out of 53 (81%) polar residues could be freely replaced by any of a selection of the amino acids Ser, Gln, Tyr, Leu, Lys and in some cases Trp, without seriously affecting the activity of the protein. By way of contrast, of 32 non-polar residues tested, 20 (63%) could be replaced only by non-polar residues or not at all without disrupting function, whereas the other 12 (37%) could be freely substituted with the amino acids experimentally available.

In an impressive extension of these studies, Kleina and Miller (1991) have studied close to 1,600 amino acid replacements (nearly one quarter of all the possible single amino acid replacements possible!). They showed that the DNA binding domain of the repressor, involving the N-terminal 59 amino acids is extremely sensitive to amino acid substitutions with 70% of the resultant proteins being inactive, confirming the observation that residues in contact with the ligand will be refractory to substitution (Perutz *et al.*, 1965). However, the remaining 301 amino acids are tolerant to substitution, with only 30% of substitutions leading to an inactive phenotype. The long

awaited determination of the crystal structure of the *E.coli lac* repressor will dramatically improve the interpretation of this data.

An even more remarkable study was performed by Rennell and co-workers (1991), where they described the phenotypic characterisation of 2015 of the 3097 possible single amino acid replacements of phage T4 lysozyme (almost two-thirds of all the possible single amino acid replacements possible!). The availability of the 3-D structure of the enzyme allowed the effects of the amino acid substitutions to be interpreted. They showed that over half (55%) of the residues could be replaced by up to 13 different amino acids with no serious deleterious affect to enzyme activity. When this data was correlated with the solvent accessibility of the amino acids, it was shown that the buried residues were, in general, more sensitive to amino acid replacement than solvent exposed residues. Notable exceptions to this were two residues thought to participate in catalysis (Hardy & Poteete, 1991).

The studies of Sauer and co-workers (Lim & Sauer, 1991, 1989; Bowie *et al.*, 1990), has lead to a quantification of some of the factors determining a proteins tolerance to amino acid replacement. They have used a combination of novel mutagenesis techniques to investigate the molecular basis of the factors that dictate a proteins ability to tolerate amino acid substitutions. Their studies have focused on the hydrophobic core of the bacteriophage lambda repressor. Three interacting residues in the hydrophobic core of the N-terminal domain of the lambda repressor were randomized combinatorially. The randomization was restricted to the five amino acids Val, Leu, Ile, Met and Phe, thereby generating a sterically diverse set of core sequences composed solely of hydrophobic residues. Of 78 of the 125 possible amino acid combinations that were generated during this study, 56 (72%) retained some degree of activity (Lim & Sauer, 1991).

Their results strongly suggest that "repacking" can occur for certain combinations of side-chains, but not all.

Fig. 4.1. **Mutants of VZV TS.**
Representation of the distribution of the mutants made to VZV TS. The figure depicts a monomer of the VZV TS model with the N- and C-terminal residues illustrated. Site-directed mutants are shown in yellow and G48 is shown in green. Mutants introduced by random mutagenesis are shown in red.

Also, they suggest that packing interactions do not appear to play a major role in specifying the basic conformation that the protein will adopt. However, packing interactions do appear to play an important role in specifying the precise structure, stability and ligand binding properties of the protein (Lim & Sauer, 1991), a process of fine tuning. In more general terms, these results suggest that the basic structural information "encoded" by these residues appears to reside largely in their hydrophobic character, that is, the majority of sequences that simply maintain hydrophobicity at core positions are able to adopt the overall lambda repressor fold and maintain moderate stability. These results were interpreted by the authors as implying that the basic structural features of a protein can be predicted from relatively simple, degenerate sequence patterns (Lim & Sauer, 1991).

In summary, the results of these various groups and those described elsewhere (Perry *et al.*, 1990; Bone *et al.*, 1989; see Section 1.9), have lead to the establishment of the concepts of plastic and covariant accomodation, qualities of proteins that allow them to tolerate amino acid substitutions that have occurred during the course of evolution, or have been introduced artificially.

### 4.3.2 Relationship of our studies of VZV TS to previous studies in other proteins

We have discovered novel examples of both plastic and covariant accomodation in VZV TS that have not previously been reported for TS from other sources. We have shown that VZV TS can tolerate amino acid substitutions, (G48A, M49P), insertions (48P49, 58T59) and even a deletion (L301Δ). These mutations are distibuted throughout various regions of the protein (see Fig. 4.1). With the availability of the model of VZV TS, we have been able to predict the manner in which these changes may have been accomodated (see Sections 3.1 and 3.3).

Another important finding has been the identification of several amino acid substitutions that inactivate VZV TS.

Many of the substitutions occur at residues that have hydrophobic side chains and are buried in the protein, and would not be expected to tolerate replacement by residues that are not non-polar. Indeed, this is the case for several of the mutants (V179E; S197T/C198R; see Section 3.2). More interesting are those substitutions that involve the replacement of core hydrophobic residues with other hydrophobic side residues. These include the mutants I41L, I96M and F189L (see Section 3.2). The local environment of each of these residues and the reasons why the substitution cannot be tolerated is described in detail in Section 3.2. In each case, the repacking of adjacent amino acid side chains is not possible for steric reasons, or due to the formation of energetically unfavourable gaps. Indeed, these findings substantiate the findings by other workers. Several of the mutants of the lambda repressor characterized by Sauer and co-workers have also been proposed as incapable of repacking due to local constraints (Lim & Sauer, 1989, 1991; Bowie *et al.*, 1990).

An additional example serves to illustrate the concept of covariant accomodation. The mutant C187G of VZV TS is inactive and cannot be accomodated locally due to the formation of an unfavourable gap in the protein core as described in Section 3.2. However, a Gly residue is found at the equivalent position in TS from *C.fasiculata*. This suggests that a Gly at this position is not disruptive to the enzymes function *per se*, and that the amino acid side chains of residues in the immediate vicinity of the Gly residue can fill the gap in TS from *C.fasciculata*, but not in TS from VZV.

In summary, these observations show that VZV TS, like several other proteins possesses inherent structural plasticity. Our studies, and those of others (Climie *et al.*, 1990; Michaels *et al.*, 1990), have revealed that highly conserved residues are not always resistant to substitution. The limits of plasticity are illustrated by the fact that even less well conserved residues cannot be freely substituted.

### 4.4 Conservation of the TS primary structure

The unusually high degree of primary structure conservation of TS from various sources has been pointed out in several reports in recent years (Hardy *et al.*, 1987; Perry *et al.*, 1990). The aim of this section is to extend these observations, and to offer some reasons for this level of conservation.

#### 4.4.1 The primary structure of TS is highly conserved

TS has one of the most highly conserved primary structures, with 21 different sequences reported from organisms as diverse as bacteriophage through to man (see Section 1.5). From a core of ~250 residues, 27 (11%) are completely conserved (see Fig. 1.6). Such a high degree of conservation is higher than for most other enzymes and similar to that found for cytochrome *c* (Perry *et al.*, 1990; Dickerson, 1971). In contrast to TS, cytochromes are small proteins (104 to 112 amino acids) and their conserved functional component per unit volume is high. In addition, much of the surface of cytochrome *c* is involved in interaction with its oxidase and reductase, both factors contributing to the conservation of the primary structure (Perry *et al.*, 1990; Dickerson, 1971).

It is also instructive to compare the degree of amino acid sequence conservation of TS with that of DHFR, especially considering the intimate relationship of these two enzymes (see Section 1.5) and the fact that they probably evolved at around the same time. DHFR has only 8 residues from a core of ~170 residues (5%) that are completely conserved (Lagosky *et al.*, 1987). The DHFR sequences are from the same range of organisms as those studied for TS. A recently published archaebacterial DHFR sequence shows a high degree of of similarity to the other DHFR sequences and contains all eight of the completely conserved residues (Zusman *et al.*, 1989). Thus, TS has more than twice as many completely conserved residues in the 'core' region when compared to DHFR.

### 4.4.2 Why is the primary structure of TS so highly conserved?

Dickerson (1971) stated that, "The rate of evolution of a protein is the rate of occurrence of mutations in the genome modified by the probability that a random change in the amino acid sequence will be tolerated in a functioning protein". This suggests that the conservation of an amino acid indicates structural and/or functional significance. This therefore implies that the highly conserved residues in TS are essential for, or have a significant role to play in, the correct functioning of the enzyme.

Structural studies of TS have to a certain extent confirmed this hypothesis. Many of the highly conserved residues are in direct contact with the substrate and cofactor molecules. Other highly conserved residues comprise the dimer interface of the protein, whilst some appear to be essential for the correct formation of other portions of the protein. In addition, the complexity of the conformational changes that occur during catalysis, namely the concerted movement of individual elements of secondary structure against each other, may account for the high degree of amino acid conservation.

Although TS does exhibit a very high level of amino acid conservation, two intriguing observations are worthy of comment. Many of the completely conserved residues can individually be freely replaced by a variety of other amino acids without inactivating the enzyme (see Climie *et al.*, 1990; Michaels *et al.*, 1990). Perhaps even more intriguing is the observation of the high degree of divergence between the sequence of TS from *L.lactis* and the TS sequences of other organisms. The *L.lactis* TS sequence can be considered the result of an elegant 'natural' mutagenesis experiment that has shown that a large number of conserved amino acids can be substituted ***en masse*** without inactivating the enzyme.

It would be illuminating to determine the structure of *L.lactis* TS and compare it with the structure of TS from other organisms to see how these amino acid differences

have been accomodated. A complementary approach would be to introduce a combination of the amino acid substitutions found in *L.lactis* TS, into TS from an alternate source. Together, these approaches may reveal much about the ability of TS to tolerate amino acid substitutions.

```
                                                        A                                              i
                                                 |-------------|                          |^^^^^^^^^^^|                   |---
D.discoideum      mgldiqteidkiviekvkpeveyydvmggshrwe.vkVhdhGkval.........vDtmprLapvgqt.adfsicqaarvsygagTKKvtedK
L.lactis                                 mtyadqvFkqniqnILdnGvfsenarpkyk...DgqmanSkyvtG.sfvtYdLqkgeFPI.TTlRpIpiK
Phage phi3T TS                          mtqfdkqYnsiikdIInnGisdeefdvrtkwdsDgTpahTLSvms.kqmRFdnse..vPILTTKKV.awK
Phage T4 TS                               mkqYqdlikdIfenGyetd..........DRTGtGTIalFG.sklRWdLtkg.FPaVTTKKL.awK
Phage T4 g42                                 misdsmtveeirLhlGlalkekdfvv....DKTGvkTIeiiG.asfvadepfifgal.......nde

                  B              C                   D                    G                            H
        -----------|    |------|           |---|                |----|                         |-----------
D.discoideum      gliryLyRhqhtsSfemvefkfhc.vmpvfiarqWirhrtanv........nEysarYsvlpdkF.......................yhps
L.lactis          saikELMWiyqdqTselsvLeeky......gVkyWgewgigd.........gtIGqrYGatvKkYni................igkLLegLak
Phage phi 3T TS   taikELLWiwqlkSndvteLnkmg.......VhIWdqwkqed.........gtIGhaYGfQlgkknrs.......lngekvDQVdyLLhqLKn
Phage T4 TS       aciaELIWFLsGsTn.vndLrliqhdsliqgktVWdenyenqakdlgyhs.gELGPIYGkQWRdF.............ggvDQIieVIdrIKk
Phage T4 g42      yiqrELeWykskslfvkdip........getpkIWqqvassk.........gEInsnYGwaiwse.............dnyaQydmcLaeLgq

                       v              I          iv                iii                   J
                  |  |^^^|      |------|   |^^^^^^^^|        |^^^^^^^^^^^|  |--------------------|
D.discoideum      ieevRKqstsnr........qggeeALePktaqefLdy......................ldkveeNyktYneLlekglsrelgriGlpvsi.
L.lactis          nPwnRRnIinlWq....yedfeetegLlPCafqtmFdVrrekdgqi.yLdatLiQRSnDMlVahhiNamqY vaLqmMIAkhfswkvGk.....
Phage phi 3T TS   nPssRRhItmlWNp....ddld.aMALtPCvyetQWYVkqg......kLhleVraRSnDMaLGnpfNvfqYnvLqrMIAqvtgyelGe.....
Phage T4 TS       lPndRRqIvsaWNp....aelk.yMALPPCHmFyQFnVrng......yLdlqwYQRSvDVfLGlPFNIASYatLvhIVAkmcnLipGD.....
Phage T4 g42      nPdsRRgImiytrpsmqfdynkdgMsdfmCtntvQYlIrdk......kInavVnmRSnDVvfGfrndyawqkyVldkLvsdlnagdstrqyka

                      ii                K
                 |^^^^^^^^^^^|     |------|
D.discoideum      ..YtewywkidlHnl.fHflrlrMDshsqkeiRdyantifaLirpivpvacegiyrlcf
L.lactis          ..Ffyfvnnl..HIYdnq.....fEqanelMkRtasekePrLvlnvp.....dgtnfF......dikpeDfeLvdYepvkpqLKfdLAI
Phage phi 3T TS   ..YIfniGdc..HVYtrH.....IDnlKiQMeReqfea.PeLwinpev....kdfynF........tvdDfkLinYkh.gdkLlfeVAV
Phage T4 TS       ..LIfsgGnt..HIYmnH.....VEqcKeiLrRePkel.ceLvisglpykfrylstkeqlkyvlklrpkDfvLnnYvs.hppIKgkMAV
Phage T4 g42      gsiIwnvGsl..HVYsrHfyl..VDhwwktgethiskkdyvgkya
```

Fig. A.1. **Sequence alignment.**
Alignment of the amino acid sequences of the*thy1* gene product of *D.discoideum*, the TS sequences of *L.lactis*, phage $\phi$3T and phage T4, and the dCMP hydroxymethylase of phage T4. The convention for the use of UPPERCASE and **UPPERCASE BOLD** is as in Fig. 1.6. This convention is extended to the *D.discoideum thy1* and phage T4 dCMP hydroxymethylase sequences where the equivalent amino acids are seen.

# APPENDIX

## A reappraisal of the *thy1* gene product of *Dictyostelium discoideum*

### Introduction

The isolation of the *thy1* gene of the cellular slime mould *Dictyostelium discoideum* has recently been reported (Dynes & Firtel, 1989). The experiments involve the growth complementation of the *Dictyostelium* strain HPS400, a thymidine auxotroph which lacks TS activity, using a partial *Dictyostelium* genomic library. The complementing gene encodes a 1.2 kb mRNA molecule and the authors were unable to show any homology of the derived protein product to TS or to any other protein sequence (Dynes & Firtel, 1989). The isolated gene, *thy1* is either the wild type copy of the gene mutated in HSP400 that confers the requirement for TdR or a suppressor of that mutation. However, since the molecular defect of the original mutation is not known, it is not possible to formally distinguish between these two possibilities. However, the publication of the *L.lactis* TS sequence (Ross *et al.*, 1990a) prompted us to re-examine the sequence of the *D.discoideum thy1* gene product.

### Results

The *D.discoideum thy1* gene product sequence, the *L.lactis*, phage ø3T and T4 TS sequences (Kenny *et al.*, 1985), and the phage T4 dCMP hydroxymethylase sequence were manually aligned using the GCG program LINEUP (see Section 3.4.2). Of the 27 invariant amino acids from a core of ~250 amino acids in TS (see Section 1.5.4), the *D.discoideum thy1* gene product sequence has 15 of these residues at equivalent positions. Of a further 20 highly conserved residues in the core region, the *D.discoideum thy1* gene product has an amino acid of similar properties (Taylor, 1986) in the equivalent position (see Fig. A.1).

The most notable disruptions to elements of predicted

secondary structure are the lack of the N-terminal portion of helices H and I, the loss of ß-strand iii and the shortening of the C-terminus. In addition, the region of the *D.discoideum thy1* gene product that aligns with the sequence of the highly hydrophobic J-helix of TS contains a high proportion of acidic residues. The most significant difference between the TS and *D.discoideum thy1* sequences, is the lack of a cysteine residue at the N-terminus of ß-strand iv that could serve as a catalytic residue.

**Discussion**

The original report by Dynes and Firtel (1989) makes no definitive claims as to the function of the *Thy1* gene product. However, at least two models can be envisaged.

1 **Transactivator Theory**

If *D.discoideum* TS gene expression is regulated by transactivation, then a mutation that results in a defective transactivating protein will result in a Thy$^-$ phenotype (TdR auxotroph; no TS activity). If this is the genetic basis for the Thy$^-$ phenotype of the *Dictyostelium* HSP400 strain, then introduction of the wild type gene that encodes the transactivating protein would complement the Thy$^-$ phenotype. In this model, introduction of the wild type gene for *D.discoideum* TS would not necessarily result in complementation as the gene would still be under the control of the same promoter that requires an active trans-activator for gene expression.

However, the concept of a transactivator protein that regulates only the expression of the TS gene seems unlikely due to precedents set in other eukaryotic organisms. In mouse and human systems, where TS gene expression has been studied, the levels of TS are co-regulated with several other enzymes involved in DNA precursor synthesis (see Section 1.5.). It is reasonable to assume that the expression of these enzymes is therefore regulated by a similar protein (or series of proteins) and that a defect in this protein(s) would affect the expression of many

other genes. Since the lesion in the *Dictyostelium* HSP400 strain seems to affect only TS (Dynes & Firtel, 1989; Podgorski & Deering, 1984), this would seem to discount this theory. Further evidence that the mutation results solely in a defect in TS comes from the report that the mouse TS gene, expressed under the control of the *Dictyostelium* actin promoter, can complement the *thy1* lesion of the *Dictyostelium* HSP400 strain (Chang *et al.*, 1989).

**2 TS Complementation Theory**

The most straightforward model that would explain the observations of Dynes & Firtel (1989), is that the *thy1* gene encodes the enzyme thymidylate synthase. Indeed, from our main argument against the transactivation theory, this seems more likely. However, the stumbling block with this argument is the degree to which the *thy1* gene product sequence has diverged from other TS sequences. As originally reported, the authors could find no sequence homology to TS from a variety of sources (Dynes & Firtel, 1989), however, our sequence comparisons have shown, there is evidence of homology between TS and the *thy1* gene product of *D.discoideum*. Obviously the degree of homology is not as high as seen between TS from most of the sources reported to date (see Fig. A.1). However, the publication of the *L.lactis* TS sequence (Ross *et al.*, 1990a), changed our, and others (Matthews *et al.*, 1989a), dogmatic way of thinking towards what constitutes an active TS sequence.

Prior to the publication of the *L.lactis* sequence, a comparison of 17 diferrent TS sequences, from a very diverse range of organisms, revealed a total of 46 invariant residues from a core of ~250 amino acids (see Section 1.5.3). However, when the *L.lactis* TS sequence was included in this comparison, the number of invariant residues dropped dramatically to 27 invariant residues. Additional amino acids are found in at least 5 sites, including the insertion of an extra residue in the predicted helix I, and the first two amino acids of the

predicted helix H are missing (see Fig. A.1). These observations, coupled with the extensive mutagenesis data available for TS (see Section 1.11) and the inherent structural plasticity of this enzyme (see Section 1.9), suggest that requirement for many of the highly conserved residues is not absolute. Thus the low degree of similarity between the *thy1* gene product of *D.discoideum* and other TSs is not a strong reason to dismiss this gene as encoding a functional TS.

Nevertheless, the lack of a potential active site cysteine residue, the amino acid that forms a covalent adduct with dUMP during the normal course of the TS catalysed reaction (see Section 1.6), is cause for concern. The absence of a serine residue, that can also act as the active site residue in *E.coli* TS (Dev *et al.*, 1988; Michaels *et al.*, 1990), adds to this problem. To explore this concern, we examined the structure/function relationship of several enzymes that have a catalytic mechanism similar to TS. These include dUMP and dCMP hydroxymethylases, and certain RNA and DNA modifying enzymes (Gu & Santi, 1991; Santi & Hardy, 1987; Wu & Santi, 1987).

The first group of enzymes, which includes dUMP and dCMP hydroxymethylase, more closely resemble TS in that they utilize $N^5,N^{10}CH_2FH_4$ as a cofactor (Santi & Danenberg, 1984) and act upon an isolated nucleotide rather than one located in DNA or RNA. Only one sequence of either of these two enzymes from any source has been reported, namely that of the phage T4 dCMP hydroxymethylase (Lamm *et al.*, 1987). Comparison of this sequence with that of T4 TS reveals a significant degree of homology and aligns the conserved Pro-Cys of T4 TS with the amino acid doublet Met-Cys (Lamm *et al.*, 1987). Furthermore, mutagenesis data has shown that the invariant Pro of this doublet can be replaced by Met (and various other amino acids) in *E.coli* TS without a serious affect on enzyme activity (Climie *et al.*, 1990). The amino acid pair Pro-Ser also lies within close proximity to the Met-Cys pair, however, if the Pro-Ser pair is aligned with the Pro-Cys pair of TS, then

there is disruption of several elements of secondary structure.

The DNA[cytosine-5]methyltransferases use S-adenosylmethionine as a methyl donor, in place of $N^5,N^{10}$-methylenetetrahydrofolate (Wu & Santi, 1987), and bear no sequence similarity to TS (even by the method described in Section 3.4.ii), save for the amino acid doublet Pro-Cys (Klimašauskas *et al.*, 1989). The Cys residue of this pair is thought to act as the active site amino acid that forms a covalent adduct with the 5-position of cytosine (Wu & Santi, 1987).

The RNA modifying enzymes, of which the *E.coli* tRNA-($m^5$U54)methyltransferase is the best characterized, also use S-adenosyl methionine as a co-factor (Gu & Santi, 1991; Santi & Hardy, 1987). The same authors have predicted that this enzyme will use a cysteine residue in an analogous manner to that found in TS (Gu & Santi, 1991; Santi & Hardy, 1987), but, somewhat surprisingly, our analysis of the amino acid sequence (Gustafsson *et al.*, 1991) of this enzyme reveals no Pro-Cys amino acid doublet, although the sequence Pro-Ser was found. Several other cysteine and serine residues are found within this protein, but too few sequences of this enzyme have been reported to define a conserved catalytic residue.

**Summary**

In summary, the above considerations suggest that the amino acid sequence of a protein that is capable of catalysing the *de novo* synthesis of dTMP may not need to be as rigidly conserved as was first apparent. The only universal feature of enzymes catalysing a methyltransferase reaction appears to be the presence of an active site cysteine residue, that serves as a nucleophile. That the *Thy1* gene product of *D.discoideum* lacks an appropriately positioned cysteine residue is the major stumbling block to promoting it as an enzyme with TS activity.

It is reasonable to assume that the phage T4 dCMP hydroxymethylase has a similar tertiary fold to that of TS

due to the similarity of the catalytic mechanism and the observed level of amino acid homology (Sander & Schneider, 1991). Indeed, an interesting finding emerged during these alignments which supports this idea. The conserved Asn residue found in the J helix of all TS sequences (including the *D.discoideum* *thy1* gene product; see Fig A.1) is replaced by Asp in dCMP hydroxymethylase. The Asn side chain in TS makes H-bond contacts to the C-4 carbonyl oxygen atom and N-5 hydrogen of the pyrimidine ring of dUMP (Matthews *et al.*, 1990a; Finer-Moore *et al.*, 1990). As dCMP (the substrate for dCMP hydroxymethylase) has an amino group at C-5 of the pyrimidine ring, in contrast to the carbonyl group found at C-5 of dUMP, it is not unreasonable to suppose that the Asp side chain could form H-bonds to dCMP in a similar way to those between Asn and dUMP. These questions will be resolved when structural studies of T4 dCMP hydroxymethylase, which are currently in progress (Thylén & Mathews, 1989), are completed.

How the tertiary fold of the *thy1* gene product of *D.discoideum* compares to that of TS is a matter for debate only after more information is available as to the function of the *thy1* gene product. One approach would be to clone the *thy1* gene under the control of a bacterial promoter, and determine if it has the ability to complement the *E.coli* *thyA* strain 2913, and/or to see if it has detectable TS activity by the conventional TS assays. If such activity is confirmed, then the *thy1* gene represents by far the most divergent TS sequence reported to date, and challenges many of the currently held ideas on the nature of the TS catalysed reaction and the structure/function relationships of the enzyme.

**REFERENCES**


ABELE, G., ERIKSSON, B., HARMENBERG, J. & WARREN, B. (1988). Inhibition of varicella-zoster virus-induced DNA polymerase by a new guanosine analog, 9-[4-hydroxy-2-(hydroxymethyl)butyl]guanine triphosphate. *Antimicrobial Agents and Chemotherapy* **32**, 1137-1142.

ACE, C. I., McKEE, T. A., RYAN, J. M., CAMERON, J. M. & PRESTON, C. M. (1989). Construction and characterization of a herpes simplex virus type 1 mutant unable to transinduce immediate-early gene expression. *Journal of Virology* **63**, 2798-2812.

ACHONG, B. G. & MEURISSE, E. V. (1968). Observation on the fine structure and replication of varicella virus in cultivated human amnion cells. *Journal of General Virology* **3**, 305-308.

ALBANI, D., PARISI, B., CARBONERA, D. & CELLA, R. (1985). Dihydrofolate reductase from *Daucus carota* cell suspension cultures: purification, molecular and kinetic characterization. *Plant Molecular Biology* **5**, 363-372.

ALLEN, J. R., REDDY, G. P. V., LASSER, G. W. & MATHEWS, C. K. (1980). T4 ribonucleotide reductase. *Journal of Biological Chemistry* **255**, 7583-7588.

AMANN, E. & BROSIUS, J. (1985). 'ATG vectors' for regulated high-level expression of cloned genes in *Escherichia coli*. *Gene* **40**, 183-190.

ANDERSON, E. P. & LAW, L. W. (1960). Biochemistry of cancer. *Annual Review of Biochemistry* **29**, 577-608.

ANDERSON, C. M., STENKAMP, R. E., McDONALD, R. C. & STEITZ, T. A. (1978). A refined model of the sugar binding site of yeast hexokinase B. *Journal of Molecular Biology* **123**, 207-219.

ANDREWS, J., CLORE, G. M., DAVIES, R. W., GRONENBORN, A. M., GRONENBORN, B., KALDERON, D., PAPADOPOULOS, P. C., SCHAFER, S., SIMS, P. F. G. & STANCOMBE, R. (1985). Nucleotide sequence of the dihydrofolate reductase gene of methotrexate resistant *Lactobacillus casei*. *Gene* **35**, 217-222.

APPLEYARD, R. K. (1954). Segregation of new lysogenic types during growth of a doubly lysogenic strain of *E.coli* K12. *Genetics* **39**, 440.

ARBETER, A. M., BAKER, L., STARR, S. E., LEVINE, B. L., BOOKS, E. & PLOTKIN, S. A. (1986a). Combination measles, mumps, rubella and varicella vaccine. *Pediatrics* **78**, 742-747.

ARBETER, A. M., STARR, S. E. & PLOTKIN, S. A. (1986b). Varicella vaccine studies in healthy children and adults. *Pediatrics* **78**, 748-756.

ARGOS, P., VINGRON, M. & VOGT, G. (1991). Protein sequence comparisons: methods and significance. *Protein Engineering* **4**, 375-383.

ARONSON, A. I. (1961). The inhibition of bacteriophage protein synthesis by 5-fluorouracil. *Biochimica et Biophysica Acta* **49**, 89-97.

ARVIN, A. M., KOROPCHAK, C. M. & WITTEK, A. E. (1983). Immunological evidence of reinfection with varicella-zoster virus. *Journal of Infectious Diseases* **148**, 200-205.

ARVIN, A. M., KUSHNER, J. H., FELDMAN, S., BAEHNER, R. L., HAMMOND, D. & MERIGAN, T. C. (1982). Human leukocyte interferon for the treatment of varicella in children with cancer. *New England Journal of Medicine* **306**, 761-765.

ASANO, Y., ITAKURA, N., HIOISHI, Y., HIROSE, S., OZAKI, T., YAZAKI, T., YAMANISHI, K. & TAKAHASHI, M. (1985). Viral replication and immunological responses in children naturally infected with varicella-zoster virus and in varicella vaccine recipients. *Journal of Infectious Diseases* **152**, 863-868.

AULL, J. L., LYON, J. A. & DUNLAP, R. B. (1974). Gel electroporesis as a means of detecting ternary complex formation of thymidylate synthetase. *Microchemistry Journal* **19**, 210-218.

AYUSAWA, D., ARAI, H., WATAYA, Y. & SENO, T. (1988). A specialized form of chromosomal DNA degradation induced by thymidylate stress in mouse FM3A cells. *Mutation Research* **200**, 221-230.

AYUSAWA, D., SHIMIZU, K., KOYAMA, H., KANEDA, S., TAKEISHI, K. & SENO, T. (1986). Cell-cycle-directed regulation of thymidylate synthase messenger RNA in human diploid fibroblasts stimulated to proliferate. *Journal of Molecular Biology* **190**, 559-567.

BABA, K., KONNO, K., SHIGETA, S. & DeCLERCQ, E. (1987). *In vitro* activity of (*S*)-9-(3-hydroxy-2-phosphonylmethoxy-propyl)-adenine against newly isolated clinical varicella-zoster virus strains. *European Journal of Clinical Microbiology* **6**, 158-160.

BACHMANN, B. J. (1983). Linkage map of *Escherichia coli* K-12, edition 7. *Microbiological Reviews* **47**, 180-230.

BACHMANN, B. & FOLLMANN, H. (1987). Deoxyribonucleotide biosynthesis in green algae: characterization of thymidylate synthase-dihydrofolate reductase in *Scenedesmus obliquus*. *Archives of Biochemistry and Biophysics* **256**, 244-252.

BAER, R., BANKIER, A. T., BIGGIN, M. D., DEININGER, P. L., FARRELL, P. J., GIBSON, T. J., HATFULL, G., HUDSON, G. S., SATCHWELL, S. C., SEGUIN, C., TUFFNELL, P. S. & BARRELL, B.G. (1984). DNA sequence and expression of the B95-8 Epstein-Barr virus genome. *Nature* **310**, 207-211.

BALASUBRAMANIAN, N. K., VEERISETTY, V. & GENTRY, G. A. (1990). Herpesviral deoxythymidine kinases contain a site analogous to the phosphoryl-binding arginine-rich region of porcine adenylate kinase; comparison of secondary structure predictions and conservation. *Journal of General Virology* **71**, 2979-2987.

BALFOUR, H. H. Jr., GROTH, K. E., McCULLOUGH, J., *et al.* (1977). Prevention or modification of varicella using zoster immune plasma. *American Journal of Diseases of Children* **131**, 693-696.

BALZARINI, J. & DeCLERCQ, E. (1991). 5-phosphoribosyl-1-pyrophospate synthase converts the acyclic nucleoside phosphonates 9-(3-hydroxy-2-phosphonylmethoxypropyl)-adenine and 9-(2-phosphonyl-methoxyethyl) adenine directly to their antivirally active di-phosphate derivatives. *Journal of Biological Chemistry* **260**, 8686-8689.

BALZARINI, J. & DeCLERCQ, E. (1982). 5-Substituted 2'-deoxyuridines: correlation between inhibition of tumor cell growth and inhibition of thymidine kinase and thymidylate synthase. *Biochemical Pharmacology* **31**, 3673-3682.

BAKER, T. S., NEWCOMB, W. W., BOOY, F. P., BROWN, J. C. & STEVEN, A. C. (1990). Three-dimensional structure of maturable and abortive capsids of equine herpesvirus 1 from cryoelectron microscopy. *Journal of General Virology* **64**, 563-573.

BARANY, F. (1985). Two-codon insertion mutagenesis of plasmid genes by using single-stranded hexameric oligonucleotides. *Proceedings of the National Academy of Sciences, U.S.A.* **82**, 4202-4206.

BARANY, F. (1987). A genetic system for isolation and characterization of Taql restriction endonuclease mutants. *Gene* **56**, 13-27.

BARBOUR, K. W., BERGER, S. H. & BERGER, F. G. (1990). Single amino acid substitution defines a naturally occuring genetic variant of human thymidylate synthase. *Molecular Pharmacology* **37**, 515-518.

BARCLAY, B. J., KUNZ, B. A., LITTLE, J. G. & HAYNES, R. H. (1982). Genetic and biochemical consequences of thymidylate stress. *Canadian Journal of Biochemistry* **60**, 172-194.

BARR, P. J., OPPENHEIMER, N. J. & SANTI, D. V. (1983). Thymidylate synthetase-catalysed conversions of *E*-5-(2-bromovinyl)-2'-deoxyuridylate. *Journal of Biological Chemistry* **258**, 13627-13631.

BARR, P. J., NOLAN, P. A., SANTI, D. V. & ROBINS, M. J. (1981). Inhibition of thymidylate synthetase by 5-alkynyl-2'-deoxyuridylates. *Journal of Medicinal Chemistry* **24**, 1385-1388.

BASHFORD, D., CHOTHIA, C. & LESK, A. M. (1987). Determinants of a protein fold. *Journal of Molecular Biology* **196**, 199-216.

BASKIN, F. CARLIN, S. C., KRAUS, P., FRIEDKIN, M. & ROSENBERG, R. N. (1975). Experimental chemotherapy of neuroblastoma II. Increased thymidylate synthetase activity in a 5-fluoro-2'-deoxyuridine resistant variant of mouse neuroblastoma. *Molecular Pharmacology* **11**, 105-117.

BATIK, O. & STEVENS, N. (1989). Varicella in pregnancy. *Journal of Family Practice* **28**, 319-321.

BEARD, P., FABER, S., WILCOX, K. W. & PIZER, L. I. (1986). Herpes simplex virus immediate early infected-cell polypeptide 4 binds to DNA and promotes transcription. *Proceedings of the National Academy of Sciences, U.S.A* **83**, 4016-4020.

BEAUDETTE, N. V., LANGERMAN, N. & KISLIUK, R. L. (1980). A calometric study of the binding of 2'-deoxyuridine-5'-phosphate and its analogues to thymidylate synthetase. *Archives of Biochemistry and Biophysics* **200**, 410-417.

BELFORT, M., MALEY, G. & MALEY, F. (1983a). Characterization of the *Escherichia coli thyA* gene and its amplified thymidylate synthase product. *Proceedings of the National Academy of Sciences, U.S.A.* **80**, 1858-1861.

BELFORT, M., MALEY, G., PEDERSEN-LANE, J. & MALEY, F. (1983b). Primary structure of the *Escherichia coli thyA* gene and its thymidylate synthase product. *Proceedings of the National Academy of Sciences, U.S.A.* **80**, 4914-4918.

BELLISARIO, R. L., MALEY, G. F., GALIVAN, J. H. & MALEY, F. (1976). *Proceedings of the National Academy of Sciences, U.S.A.* **73**, 1848-1852.

BELLISARIO, R. L., MALEY, G. F., GUARINO, D. U. & MALEY, F. (1979). The primary structure of *Lactobacillus casei* thymidylate synthetase: II. The complete amino acid sequence of the active site peptide, CNBr 4. *Journal of Biological Chemistry* **254**, 1296-1300.

BENKOVIC, S. J. (1980). On the mechanism of action of folate- and biopterin-requirin enzymes. *Annual Review of Biochemistry* **49**, 227-251.

BENKOVIC, S. J. & BULLARD, W. P. (1973). On the mechanisms of action of folic acid cofactors. *Progress in Bioorganic Chemistry* **2**, 133-175.

BEN-PORAT, T. & RIXON, F. J. (1979). Replication of herpesvirus DNA. IV. Analysis of concatamers. *Virology* **94**, 61-70.

BERGER, S. H. & BERGER, F. G. (1988). Thymidylate synthase as a determinant of 5-fluoro-2'-deoxyuridine response in human colonic tumor cell lines. *Molecular Pharmacology* **34**, 474-479.

BERGER, S. H., BARBOUR, K. W. & BERGER, F. G. (1988). A naturally occuring variation in thymidylate synthase structure is associated with a reduced response to 5-fluoro-2'-deoxyuridine in a human colonic tumor cell line. *Molecular Pharmacology* **34**, 480-484.

BEVERLEY, S. M., ELLENBERGER, T. E. & CORDINGLEY, J. S. (1986). Primary structure of the gene encoding the bifunctional dihydrofolate reductase-thymidylate synthase of *Leishmania major*. *Proceedings of the National Academy of Sciences, U.S.A.* **83**, 2584-2588.

BIRD, A. D. (1986). CpG-rich islands and the function of DNA methylation. *Nature (London)* **321**, 209-213.

BIRD, O. D., VAITKUS, J. W. & CLARK, J. (1970). 2-amino-4-hydroxy-quinazolines as inhibitors of thymidylate synthetase. *Molecular Pharmacology* **6**, 573-575.

BIRON, K. K., FYFE, J. A., NOBLIN, J. E. & ELION, G. B. (1982). Selection and preliminary characterisation of acyclovir-resistant mutants of varicella-zoster virus. *American Journal of Medicine* **73**, 383-386.

BISSON, L. & THORNER, J. (1977). Thymidine 5'-monophosphate-requiring mutants of *Saccharomyces cerevisiae* are deficient in thymidylate synthetase. *Journal of Bacteriology* **132**, 44-50.

BLACK, L. W. (1989). DNA packaging in dsDNA bacteriophages. *Annual Review of Microbiology* **43**, 267-292.

BLACK, M. E. & HRUBY, D. E. (1991). Nucleotide sequence of the *Escherichia coli* thymidine kinase gene provides evidence for conservation of functional domains and quaternary structure. *Molecular Microbiology* **5**, 373-379.

BLACKLOW, S. C. & KNOWLES, J. C. (1991). In *Directed Mutagenesis A Practical Approach*, pp177-198. M. J. McPherson (Editor). Oxford University Press, New York.

BLAKLEY, R. L. (1963). The biosynthesis of thymidylic acid. IV. Further studies on thymidylate synthetase. *Journal of Biological Chemistry* **238**, 2113-2118.

BOCKAMP, E.-O., BLASCO, R. & VINUELA, E. (1990). *Escherichia coli* thymidine kinase: nucleotide sequence of the gene and relationship to other thymidine kinases. *Gene* **101**, 9-14.

BODEMER, W., NILLER, H. H., NITSCHE, N., SCHOLZ, B. & FLECKENSTEIN, B. (1986). Organization of the thymidylate synthase gene of herpesvirus saimiri. *Journal of Virology* **60**, 114-123.

BONE, R., SILEN, J. L. & AGARD, D. A. (1989). Structural plasticity broadens the specificity of an engineered protease. *Nature* **339**, 191-195.

BONNEY, R. J. & MALEY, F. (1975). Effect of methotrexate on thymidylate synthetase in cultured parenchymal cells isolated from regenerating rat liver. *Cancer Research* **35**, 1950-1956.

BOOY, F. P., NEWCOMB, W. W., TRUS, B. L., BROWN, J. C., BAKER, T. S. & STEVEN, A. C. (1991). Liquid-crystalline, phage-like packaging of encapsidated DNA in herpes simplex virus. *Cell* **64**, 1007-1015.

BORDO, D. & ARGOS, P. (1991). Suggestions for 'safe' residue substitutions in site-directed muytagenesis. *Journal of Molecular Biology* **217**, 721-729.

BOWIE, J. U., REIDHAAR-OLSEN, J. F., LIM, W. A. & SAUER, R. T. (1990). Deciphering the message in protein sequenes: tolerance to amino acid substitutions. *Science* **247**, 1306-1310.

BOYD, C. A. & SHERRATT, D. J. (1986). Polar mobilization of the *E.coli* chromosome by the ColE1 transfer origin. *Molecular and General Genetics* **203**, 496-504.

BOYLE, D. B., COUPAR, B. E. H., GIBBS, A. J., SEIGMAN, L. J. & BOTH, G. W. (1987). Fowlpox virus thymidine kinase: nucleotide sequence and relationship to other thymidine kinases. *Virology* **156**, 355-365.

BRADFORD, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protien utilizing the principle of protein dye-binding. *Analytical Biochemistry* **72**, 248-252.

BRENDEL, M. & LANGJHAR, U. G. (1974). Thymine-less death in a strain of *Saccharomyces cerevisiae* auxotrophic for deoxythymidine 5'-monophosphate. *Molecular and General Genetics* **131**, 351-358.

BRENT, B & PTASHNE, M. (1981). Mechanism of action of the *lexA* gene product. *Proceedings of the National Academy of Sciences, U.S.A.* **78**, 4204-4208

BRIGHT, R. (1831). *Report of Medical Cases* **2**, 383.

BROCKMAN, R. W. & ANDERSON, E. P. (1963). Pyrimidine analogues. In R. M. Hochster & J. H. Quastel (Editors), *Metabolic Inhibitors*. Academic Press Inc., New York.

BROCKMAN, R. W. & LAW, L. W. (1960). Metabolism of pyrimidines and 5-fluoropyrimidines by sensitive and resistant neoplasms. *Proceedings of the American Association of Cancer Research* **3**, 98.

BROCKMAN, R. W., DAVIS, J. M. & STUTTS, P. (1960). Metabolism of uracil and 5-fluorouracil by drug-sensitive and drug-resistant bacteria. *Biochimie et Biophysica Acta* **40**, 22-32.

BROCKMAN, R. W., SHADDIX, S. C., ROSE, L. M., ELLIOTT, R. D. & MONTGOMERY, J. A. (1984). Activity of 5'-[(halo-acetyl)amido]-5'-deoxy derivatives of pyrimidine nucleosides in tumour cells in culture and *in vivo*. *Proceedings of the American Association for Cancer Research* **25**, 1426.

BROSIUS, J., DULL, T. J., SLEETER, D. D. & NOLLER, H. F. (1981). Gene organization and primary structure of a ribosomal RNA operon from *Escherichia coli*. *Journal of Molecular Biology* **148**, 107-127.

BROUGHTON, C. R. (1966). Varicella-zoster in Sydney. I. Varicella and its complications. *Medical Journal of Australia* **2**, 392-397.

BROWN, G. R. (1976). Correlation of age, sex distribution, neuralgia, and associated disorders. *Southern Medical Journal* **69**, 576-578.

BRUNELL, P. A. (1983). Fetal and neonatal varicella-zoster infections. *Seminars in Perinatology* **7**, 47-56.

BRUNELL, P. A. (1977). Protection against varicella. *Pediatrics* **59**, 1-2.

BRUNELL, P. A. (1981). Varicella-zoster infections. In *Textbook of pediatric infectious diseases*. Eds. R. D. Feigen & J. D. Cherry. Philidephelia: W. B. Saunders.

BRUNELL, P. A., ROSS, A., MILLER, L. H. & KUO, B. (1969). Prevention of varicella by zoster immune globulin. *New England Journal of Medicine* **280**, 1191-1194.

BRNGER, A., KURIYAN, J. & KARPLUS, M. (1987). Crystallographic *R* factor refinement by molecular dynamics. *Science* **235**, 458-460.

BUCHMAN, A. R. & BERG, P. (1988). Comparison of intron-dependent and intron-independent gene expression. *Molecular and Cellular Biology* **8**, 4395-4405.

BUELL, M. V. & HANSEN, R. E. (1960). Reaction of pyridoxal-5-phosphate with aminothiols. *Journal of the American Chemical Society* **82**, 6042-6049.

BULLOCK, W. O., FERNANDEZ, J. M. & SHORT, J. M. (1987). *Biotechniques* **5**, 376-379.

BULLOWA, J. G. M. & WISHIK, S. M. (1935). Complications of varicella, I. Their occurence among 2,534 patients. *American Journal of Diseases of Children* **49**, 923-926.

BUNNI, M., DOIG, M. T., DONATO, H., KESAVAN, V. & PRIEST, D. G. (1988). Role of methylenetetrahydrofolate depletion in methotrexate-mediated intracellular thymidylate synthesis inhibition in cultured L1210 cells. *Cancer Research* **48**, 3398-3404.

BURCHALL, J. J. & HITCHINGS, G. H. (1965). Inhibitor binding analysis of DHFR from various species. *Molecular Pharmacology* **1**, 126-136.

BURDESKA, A., OTT, M., BANNWARTH, W. & THEN, R. L. (1990). Identical genes from trimethoprim-resistant dihydrofolate reductase from *Staphylococcus aureus* in Australia and Central Europa. *FEBS Letters* **266**, 159-162.

BURES, A. K., DARON, H. H. & AULL, J. L. (1991). Inhibition of thymidylate synthase by glceraldehyde 3-phosphate. *International Journal of Biochemistry* **23**, 733-736.

BURGESS, A. W., PONNUSWAMY, P. K. & SCHERAGA, H. A. (1974). Analysis of conformation of amino acid residues and prediction of backbone topography in proteins. *Israeli Journal of Chemistry* **12**, 239-286.

BURGOON, C. F., BURGOON, J. S. & BALDRIDGE, G. D. (1957). The natural history of herpes zoster. *Journal of the American Medical Association* **164**, 265-269.

BURTON, K. (1977). Transport of adenine, hypoxanthine and uracil into *Escherichia coli*. *Biochemistry Journal* **168**, 195-204.

BURTON, K. (1983). Transport of nucleic acid bases into *Escherichia coli*. *Journal of General Microbiology* **129**, 3505-3513.

BUSSARD, A., NAONA, S., GROS, F. & MONOD, J. (1960). Effects d'un analogue de l'uracile sur les propits d'une protine enzymatique synthtise en sa prsence. *Comptes Rendus de l'Academie des Sciences* **250**, 4049-4051.

BYRD, R. A., DAWSON, W. H., ELLIS, P. D. & DUNLAP, R. B. (1978). Elucidation of the detailed structures of the native and denatured ternary complexes of thymidylate synthetase via $F^{19}$NMR. *Journal of the American Chemical Society* **100**, 7478-7486.

BZIK, D. J., LI, W.-B., HORII, T. & INSELBURG, J. (1987). Molecular cloning and sequence analysis of the *Plasmodium falciparum* dihydrofolate reductase-thymidylate synthase gene. *Proceedings of the National Academy of Sciences, U.S.A.* **84**, 8360-8364.

CABRIAC, G. F., MAHALINGHAM, R., WELLISH, M. & GILDEN, D. H. (1990). Trans-activation of viral tk promoters by proteins encoded by varicella-zoster virus opoen reading frames 61 and 62. *Virus Research* **15**, 57-68.

CAMERON, J. M., McDOUGALL, I., MARSDEN, H. S., PRESTON, V. G., RYAN, D. M. & SUBAK-SHARPE, J. H. (1988). Ribonucleotide reductase encoded by herpes simplex virus is a determinant of the pathogenicity of the virus in mice and a valid antiviral target. *Journal of General Virology* **69**, 2607-2612.

CAPCO, G. R., KRUPP, T. L. & MATHEWS, C. K. (1973). Bacteriophage-coded thymidylate synthetase: characteristics of the T4 and T5 enzymes. *Archives of Biochemistry and Biophysics* **158**, 726-775.

CHALLBERG, M. D. & KELLY, T. J. (1989). Animal virus DNA replication. *Annual Review of Biochemistry* **58**, 671-717.

CHANG, A. C. M., WILLIAMS, K. L., WILLIAMS, J. G. C. & CECCARELLI, A. (1989). Complementation of a *Dictyostelium discoideum* thymidylate synthase mutation with the mouse gene provides a new selectable marker for transformation. *Nucleic Acids Research* **17**, 3655-3661.

CHEE, M. S. & BARRELL, B. G. (1990). Herpesviruses: a study of parts. *Trends in Genetics* **6**, 86-91.

CHEE, M. S.,BANKIER, A. T., BECK,, S., BOHNI, R., BROWN, C. M., CERNY, R., HORSNELL, T., HUTCHINSON, C. A., III, KOUZARIDES, T., MARTIGETTI, J. A., PREDDIE, E., SATCHWELL, S.C., TOMLINSON, P., WESTON, K. M. & BARRELL, B. G. (1990). Analysis of the protein-coding content of the sequence of the human cytomegalovirus strain AD169. *Current Topics in Microbiology and Immunology* **154**, 125-169.

CHEN, M. S., SUMMERS, W. P., WALKER, J., SUMMERS, W. C. & PRUSSOFF, W. H. (1979). Charaterization of pyrimidine deoxyribonucleotide kinase (thymidine kinase) and thymidylate kinase as a multifunctional enzyme in cells transformed by herpes simplex type 1 and in cells infected with mutant strains of herpes simplex virus. *Journal of Virology* **30**, 942-945.

CHEN, S., DARON, H. H. & AULL, J. L. (1988). Inhibition of thymidylate synthase by pyridoxal phosphate. *FASEB Journal* **2**, A1546

CHEN, S., DARON, H. H. & AULL, J. L. (1989). Inhibition of thymidylate synthase by pyridoxal phosphate. *International Journal of Biochemistry* **21**, 1217-1221.

CHENG, Y.-C., DUTSCHMAN, G., DeCLERCQ, E., JONES, A. S., VERHELST, G. & WALKER, R. T. (1981). Differential affinities of 5-(2-halogenvinyl)-2'-deoxyuridines for deoxythymidine kinases of various origins. *Molecular Pharmacology* **20**, 230-233.

CHENG, Y.-C., TSOU, T. Y., HACKSTADT, T. & MALLAVIA, L. P. (1979). Induction of thymidine kinase and DNase in varicella-zoster virus-infected cells and kinetic properties of the virus induced thymidine kinase. *Journal of Virology* **31**, 172-177.

CHEUNG, A. K. (1991). Cloning the latency gene and the early protein 0 gene of pseudorabies virus. *Journal of Virology* **65**, 5260-5271.

CHIU, C.-S., COOK, K. S. & GREENBERG, G. R. (1982). Characteristics of a bacteriophage T4-induced complex synthesizing deoxyribonucleotides. *Journal of Biological Chemistry* **257**, 15087-15097.

CHOTHIA, C. & JANIN, J. (1981). Relative orientation of close-packed ß-pleated sheets in proteins. *Proceedings of the National Academy of Sciences, U.S.A.* **78**, 4146-4150.

CHOTHIA, C. & LESK, A. R. (1986). The relation between divergence of sequence and structure in proteins. *EMBO Journal* **5**, 823-826.

CHOTHIA, C., LEVITT, M. & RICHARDSON, D. (1977). Structures of proteins: packing of '-helices and pleated sheets. *Proceedings of the National Academy of Sciences, U.S.A.* **74**, 4130-4134.

CHOTTINER, E. G., SHEWACH, D. S., DATTA, N. S., ASHCRAFT, E., GRIBBIN, D., GINSBURG, D. FOX, I. H. & MITCHELL, B. S. (1991). *Proceedings of the National Academy of Sciences, U.S.A.* **88**, 1531-1535.

CHOU, P. Y. & FASMAN, G. (1978). Prediction of the secondary structure of proteins from their amino acid sequence. *Advances in Enzymology* **47**, 45-148.

CHU, E., DRAKE, J. C., BOARMAN, D., BARAM, J. & ALLEGRA, C. J. (1990). Mechanism of thymidylate synthase inhibition by methotrexate in human neoplastic cell lines and normal human myeloid progenitor cells. *Journal of Biological Chemistry* **265**, 8470-8478.

CHU, F. K., MALEY, G. F., MALEY, F. & BELFORT, M. (1984). Intervening sequence in the thymidylate synthase gene of bacteriophage T4. *Proceedings of the National Academy of Sciences, U.S.A.* **81**, 3049-3053.

CHUNG, C. T. & MILLER, R. H. (1988). A rapid and convenient method for the preparation and storage of competent bacterial cells. *Nucleic Acids Research* **16**,

CHUNG, S. & PERRY, R. P. (1989). Importance of introns for expression of mouse ribosomal protein gene rpL32. *Molecular and Cellular Biology* **9**, 2075-2082.

CIPOLLO, K. L. & DUNLAP. R. B. (1978). Essential arginyl residues in thymidylate synthetase. *Biochemical and Biophysical Research Communications* **81**, 1139-1144.

CIPOLLO, K. L. & DUNLAP, R. B. (1979). Essential arginyl residues in thymidylate symthetase from aminopterin resistant *Lactobacillus casei*. *Biochemistry* **18**, 5573-5541.

CLARKE, L. & CARBON, J. (1976). A colony bank containing synthetic Col E1 hybrid plasmids representative of the entire *E.coli* genome. *Cell* **9**, 91-99.

CLEMENTS, J. B. (1990). Herpesviridae: Properties of the viruses. In M. T. Parker & L. H. Collier (Editors), *Principles of Bacteriology, Virology and Immunity.* Vol. 4. *Virology* 414-429. Edward Arnold, London.

CLEMENTS, J. B., WATSON, R. J. & WILKIE, N. M. (1977). Temporal regulation of herpes simplex virus type 1 transcription: location of transcripts on the viral genome. *Cell* **12**, 275-285.

CLIMIE, S. & SANTI, D. V. (1990). Chemical synthesis of the thymidylate synthase gene. *Proceedings of the National Academy of Sciences, U.S.A.* **87**, 633-637.

CLIMIE, S., RUIZ-PEREZ, L., GONZALEZ-PACANOWSKA, D., PRAPUNWATTANA, P., CHO, S.-W., STROUD, R. & SANTI, D. V. (1990). Saturation site-directed mutagenesis of thymidylate synthase. *Journal of Biological Chemistry* **265**, 18776-18779.

COHEN, F. E., STERNBERG, M. J. E. & TAYLOR, W. R. (1981). Analysis of the tertiary structure of protein ß-sheet structures. *Journal of Molecular Biology* **148**, 253-272.

COOK, M. L., BASTONE, V. B. & STEVENS, J. G. (1974). Evidence that neurons harbor latent herpes simplex virus. *Infections and Immunity* **9**, 946-951.

CONRAD, A. H. (1971). Thymidylate synthetase activity in cultured mammalian cells. *The Journal of Biologial Chemistry* **246**, 1318-1323.

CONRAD, A. H. & RUDDLE, F. H. (1972). Regulation of thymidylate synthetase activity in cultured mammalian cells. *Journal of Cell Science* **10**, 471-486.

CORNELISSEN, A. W. C. A., VERSPIEREN, M. P., TOULME, J.-J., SWINKELS, B. W. & BORST, P. (1986). The common 5' terminal sequence on trypanosome mRNAs: a target for anti-messenger oligodeoxynucleotides. *Nucleic Acids Research* **14**, 5605-5614.

CORY, J. G., CRUMLEY, J. & WILKINSON, D. S. (1977). Evidence for role of purine nucleoside phosphorylase in sensitivity of Novikoff hepatoma cells to 5-fluorouracil. In G. Weber (Editor), *Advances in Enzyme Regulation*, vol. 15 pp153-166. Pergamon Press, New York.

COWMAN, A. F. & LEW, A. M. (1989). Anti-folate selection results in duplication and rearrangement of chromosome 7 in *Plasmodium chabaudi*. *Molecular and Cellular Biology* **9**, 5182-5188.

COX, E. C. & HORNER, D. L. (1982). Dominant mutators in *Escherichia coli*. *Genetics* **100**, 7-18.

CROEN, K. D. & STRAUS, S. E. (1991). Varicella-zoster virus latency. *Annual Review of Microbiology* **45**, 265-282.

CROEN, K. D., OSTROVE, J. M., DRAGOVIC, L. J. & STRAUS, S. E. (1988). Patterns of gene expression and sites of latency in human nerve ganglia are different for varicella-zoster and herpes simplex viruses. *Proceedings of the National Academy of Sciences, U.S.A.* **85**, 9773-9777.

CRUSBERG, T. C., LEARY, R. & KISLUIK, R. L. (1970). Properties of thymidylate synthetase from dichloro-methatrexate-resistant *Lactobacillus casei*. *Journal of Biological Chemistry* **245**, 5292-5296.

DALRYMPLE, M. A., McGEOCH, D. J., DAVISON, A. J. & PRESTON, C. M. (1985). DNA sequence of the herpes simplex virus type 1 gene whose promoter is responsible for transcriptional activation of immediate early promoters. *Nucleic Acids Research* **13**, 7865-7879.

DANENBERG, P. V. (1977). Thymidylate synthetase - a target enzyme in cancer chemotherapy. *Biochimica et Biophysica Acta* **473**, 73-92.

DANENBERG, K. D. & DANENBERG, P. V. (1979). Evidence for a sequential interaction of the subunits of thymidylate synthase. *Journal of Biological Chemistry* **254**, 4345-4348.

DANENBERG, K. D. & DANENBERG, P. V. (1989). Activity of thymidylate synthetase and its inhibition by 5-fluorouracil in highly enzyme-overproducing cells resistant to 10-propargyl-5,8-dideazafolate. *Molecular Pharmacology* **36**, 219-223.

DANENBERG, P. V. & LOCKSHIN, A. (1982). Tight-binding complexes of thymidylate synthetase, folate analogues and deoxyribonucleotides. *Advances in Enzyme Regulation* **20**, 99-110.

DANENBERG, P. V., BHATT, R. S., KUNDU, N. G., DANENBERG, K. & HEIDELBERGER, C. (1981). Interaction of 5-ethynyl-2'-deoxyuridylate with thymidylate synthetase. *Journal of Medicinal Chemistry* **24**, 1537-1540.

DANENBERG, P. V., LANGENBACH, R. J. & HEIDELBERGER, C. (1974). Structures of reversible and irreversible complexes of thymidylate synthetase and fluorinated pyrimidine nucleotides. *Biochemistry* **13**, 926-933.

DARBY, G., LARDER, B. A. & INGLIS, M. M. (1986). Evidence that the 'active centre' of the herpes simplex virus thymidine kinase invloves an interaction between three distinct regions of the polypeptide chain. *Journal of General Virology* **67**, 753-758.

DARGAN, D. J. (1986). The structure and assembly of herpesviruses. In *Electron Microscopy of Proteins: Viral Structure.* Eds. J. R. Harris & R. W. Horne. **5**, 359-437. Orlando/Florida: Academic Press.

DARON, H. & AULL, J. L. (1978). A kinetic study of thymidylate synthetase from *L.casei*. *Journal of Biological Chemistry* **253**, 940-945.

DAVISON, A. J. (1983). DNA sequence of the Us component of the varicella-zoster virus genome. *EMBO Journal* **2**, 2203-2209.

DAVISON, A. J. (1984). Structure of the genome terminii of varicella-zoster virus. *Journal of General Virology* **65**, 1969-1977.

DAVISON, A. J. (1991). Varicella-zoster virus. *Journal of General Virology* **72**, 475-486.

DAVISON, A. J. (1992). Channel catfish virus: a new type of herpesvirus. *Virology* **186**, 9-14.

DAVISON, A. J. & WILKIE, N. M. (1983). Location and orientation of homologous sequences in the genomes of five herpesviruses. *Journal of General Virology* **64**, 1927-1942.

DAVISON, A. J. & McGEOCH, D. J. (1986). Evolutionary comparisons of the S segments in the genomes of herpes simplex virus type 1 and varicella-zoster virus. *Journal of General Virology* **67**, 597-611.

DAVISON, A. J. & RIXON, F. J. (1984). Cloning of DNA of alphaherpesvirinae. In *Developments in Molecular Virology*, vol. 5. Edited by Y. Becker. The Hague: Martinus Nijhoff.

DAVISON, A. J. & SCOTT, J. E. (1985). DNA sequence of the major inverted repeat in the varicella-zoster virus genome. *Journal of General Virology* **66**, 207-220.

DAVISON, A. J. & SCOTT, J. E. (1986). The complete DNA sequence of varicella-zoster virus. *Journal of General Virology* **67**, 1759-1816.

DAVISON, A. J., WATERS, D. J. & EDSON, C. M. (1985). Identification of the products of a varicella-zoster virus glycoprotein gene. *Journal of General Virology* **66**, 2237-2242.

DAVISON, A. J., EDSON, C. M., ELLIS, R. W., FORGHANI, B., GILDEN, D., GROSE, C., KELLER, P. M., VAFAI, A., WROBLEWSKA, Z. & YAMANISHI, K. (1986). New common nomenclature for glycoprotein genes of varicella-zoster virus and their glycosylated products. *Journal of Virology* **57**, 1195-1197.

DAYRINGER, H. E., TRAMONTANO, A., SPRANG, S. R. & FLETTERICK, R. J. (1986). Interactive program for visualization and modelling of proteins, nucleic acids and small molecules. *Journal of Molecular Graphics* **4**, 82-87.

DeCLERCQ, E., HOLY, A., ROSENBERG, I., SAKUMA, T., BALZARINI, J. & MAUDGAL, P. C. (1986). Anovel selective broad-spectrum anti-DNA virus agent. *Nature* **323**, 464-467.

DeCLERCQ, E., BALZARINI, J., TORRENCE, P. F., MERTES, M. P., SCHMIDT, C. L., SHUGAR, D., BARR, P.J., JONES, A. S., VERHELST, G. & WALKER, R. T. (1981). Thymidylate synthetase as target enzyme for the inhibitory activity of 5-substituted-2'-deoxyuridines on mouse leukemia L1210 cell growth. *Molecular Pharmacology* **19**, 321-330.

DEISS, L. P., CHOU, J. & ROIZMAN, B. (1986). Functional domains within the *a* sequence involved in the cleavage-packaging of herpes simplex virus DNA. *Journal of Virology* **59**, 605-618.

DeLUCA, N.A., McCARTHY, A. M. & SCHAFFER, P. A. (1985). Isolation and characterization of deletion mutants of herpes simplex virus type 1 in the gene encoding immediate-early regulatory protein ICP4. *Journal of Virology* **56**, 558-570.

deMORAGAS, J. M. & KIERLAND, R. R. (1957). The outcome of patients with herpes zoster. *Archives of Dermatology* **75**, 193-196.

DENG, T., LI, D., JENH, C.-H. & JOHNSON, L. F. (1986). Structure of the gene for mouse thymidylate synthase: locations of introns and multiple transcriptional start sites. *The Journal of Biological Chemistry* **261**, 16000-16005.

DENG, T., LI, D., JOLIFF, K. & JOHNSON, L. F. (1989). The thymidylate synthase promoter: essential elements are in close proximity to the transcriptional initiation sites. *Molecular and Cellular Biology* **9**, 4079-4082.

DENNY-BROWN, D., ADAMS, R. D. & FITZGERALD, P. J. (1944). Pathological features of herpes zoster. A note on "geniculate herpes". *Archives of Neurology and Psychiatry* **51**, 216-231.

DEV, I. K., YATES, B. B., ATASHI, J. & DALLAS, W. S. (1989). Catalytic role of histidine 147 in *Escherichia coli* thymidylate synthase. *Journal of Biological Chemistry* **264**, 19132-19137.

DEV, I. K., YATES, B. B., LEONG, J. & DALLAS, W. S. (1988). Functional role of cysteine-146 in *Escherichia coli* thymidylate synthase. *Proceedings of the National Academy of Sciences, U.S.A.* **85**, 1472-1476.

DeWILLE, J. W., HARENDZA, C. J., JENH, C.-H. & JOHNSON, L. F. (1989). Analysis of the multiple 5' and 3' termini of poly(A)+ and poly(A)- deficient thymidylate synthase mRNA in growth-stimulated mouse fibroblasts. *Journal of Cellular Physiology* **138**, 358-366.

DICKERSON, R. E. (1971). The structure of cytochrome c and the rates of molecular evolution. *Journal of Molecular Evolution* **1**, 26-45.

DISNEY, G. H. & EVERETT, R. D. (1990). A herpes simples virus type 1 recombinant with both copies of $V_{MW}$175 coding sequence replaced by the homologous varicella-zoster virus open reading frame. *Journal of General Virology* **71**, 2681-2689.

DISNEY, G. H., McKEE, T. A., PRESTON, C. M. & EVERETT, R. D. (1991). The product of varicella-zoster virus gene 62 autoregulates its own promoter. *Journal of General Virology* **72**, 2999-3003.

DIXON, R. A. F. & SCHAFFER, P. A. (1980). Fine-structure mapping and functional analysis of temperature-sensititve mutants in the gene encoding the herpes simplex virus type 1 immediate-early protein VP175. *Journal of Virology* **36**, 189-203.

DOLIN, R., REICHMAN, R. C., MAZUR, M. H. & WHITLEY, R. J. (1978). Herpes zoster-varicella infection in immunosuppressed patients. *Annals of Internal Medicine* **89**, 375-388.

DOLNICK, B. J. & CHENG, Y. C. (1978). Human thymidylate synthetase II. Derivatives of pteroylmono- and-polyglutamates as substrates and inhibitors. *Journal of Biological Chemistry* **253**, 3563-3567.

DONATO, H. J., Jr., AULL, J. L., LYON, J. A., REINSCH, J. W. & DUNLAP, R. B. (1976). Formation of ternary complexes of thynidylate synthetase as followed by absorbance, fluoresence, and circular dichroism spectra and gel electrophoresis. *Journal of Biological Biochemistry* **251**, 1303-1310.

DOTTO, G. P., HORIUCHI, K. & ZINDER, N. D. (1984). The functional origin of bacteriophage f1 DNA replication. *Journal of Molecular Biology* **172**, 507-521.

DRETZEN, G., BELLARD, M., SASSONE-CORSI, P. & CHAMBON, P. (1981). A reliable method for the recovery of DNA fragments from agarose and acrylamide gels. *Analytical Biochemistry* **112**, 295-298.

DUSCHINSKY, R., PLEVEN, E. & HEIDELBERGER, C. (1957). The synthesis of 5-fluoropyrimidines. *Journal of the American Chemical Society* **79**, 4559-4560.

DUMAS, A. M., GEELEN, J. L. M. C., WESTSTRATE, M. W., WERTHEIM, P. & VAN DER NOORDAA, J. (1981). *Xba*I, *Pst*I and *Bgl*II restriction endonuclease maps of the two orientations of the varicella-zoster virus genome. *Journal of Virology* **39**, 390-400.

DYNES, J. L. & FIRTEL, R. A. (1989). Molecular complementation of a genetic marker in *Dictyostelium* using a genomic DNA library. *Proceedings of the National Academy of Sciences, U.S.A.* **86**, 7966-7970.

DZIK, J. M., BRETNER, M., KULIKWSKI, T., CIESLA, J., CEISLA, J. M., RODE, W. & SHUGAR, D. (1988). Interaction of the 5'-phosphates of the anti-HIV agents, 3'azido-3'-deoxythymidine and 3'-azido-2',3'-dideoxyuridine, with thymidylate synthase. *Biochemical and Biophysical Research Communications* **155**, 1418-1423.

EASTON, H. G. (1970). Zoster sine herpete causing trigeminal neuralgia. *Lancet* **2**, 1065-1066.

ECKER, J. R., KUDLER, L. & HYMAN, R. W. (1984). Variation in the structure of varicella-zoster virus DNA. *Intervirology* **21**, 25-37.

EDMAN, U., EDMAN, J. C., LUNDGREN, B. & SANTI, D. V. (1989). Isolation and expression of the *Pneumocystis carinii* thymidylate synthase gene. *Proceedings of the National Academy of Sciences, U.S.A.* **86**, 6503-6507.

ELION, G. B., FURMAN, P. A., FYFE, J. A., MIRANDA, P. D., BEAUCHAMP, L. & SCHAEFFER, H. J. (1977). Selectivity of action of an antiherpetic agent, 9-(2-hydroxyethoxymethyl)guanine. *Proceedings of the National Academy of Sciences, U.S.A.* **74**, 5716-5720.

ELLIOTT, R. D., PRUETT, P. S., BROCKMAN, R. W. & MONTGOMERY, J. A. (1987). Reactive 5'-substituted 2',5'-dideoxyuridine derivatives as potential inhibitors of nucleotide biosynthesis. *Journal of Medicinal Chemistry* **30**, 927-930.

EPSTEIN, M. A. (1962). Observations on the fine structure of mature herpes simplex virus and on the composition of its nucleoid. *Journal of Experimental Medicine* **115**, 1-9.

ESIRI, M. M. & TOMLINSON, A. H. (1972). Herpes zoster: demonstration of virus in trigeminal nerve and ganglion by immunofluoresence and electron microscopy. *Journal of Neurological Science* **15**, 35-48.

ETO, I. & KRUMDIECK, C. L. (1981). Determination of three different pools of reduced one-carbon-substituted folates. II. Quantitation and chain-length determination of the pteroylpolyglutamtes of rat liver. *Analytical Biochemistry* **115**, 138-146.

EVANS, R. M., LASKIN, J. D. & HAKALA, M. T. (1981). Effect of excess folates and deoxyinosine on the activity and site of action of 5-fluorouracil. *Cancer Research* **41**, 3288-3295.

EVERETT, R. D. (1984a). *Trans* activation of transcription by herpes-virus products: requirement for two HSV-1 immediate-early poly-peptides for maximum activity. *EMBO Journal* **3**, 3135-3141.

EVERETT, R. D. (1984b). A detailed analysis of an HSV-1 early promoter: sequences involved in *trans*-induction by viral immediate-early gene products are not early-gene specific. *Nucleic Acids Research* **12**, 3039-3056.

EVERETT, R. D. (1986). The products of herpes simplex virus type 1 (HSV-1) immediate early genes 1, 2 and 3 can activate HSV-1 gene expression in *trans*. *Journal of General Virology* **67**, 2507-2513.

FELDMAN, S. & EPP, E. (1979). Detection of viremia during incubation with varicella. *Journal of Pediatrics* **94**, 746-748.

FELDMAN, S., HUGHES, W. T. & DANIEL, C. B. (1975). Varicella in children with cancer: seventy-five cases. *Pediatrics* **56**, 388-397.

FELSER, J. M., KINCHINGTON, P. R., INCHAUSPE, G., STRAUSS, S. E. & OSTROVE, J. (1988). Cell lines containing varicella-zoster virus open reading frames 62 and expressing the "IE" 175 protein complement ICP4 mutants of herpes simpplex virus type 1. *Journal of Virology* **62**, 2076-2082.

FERONE, R., BURCHALL, J. J. & HITCHINGS, G. H. (1969). *Plasmodium berghei* dihydrofolate reductase. *Molecular Pharmacology* **5**, 49-59.

FIFE, & PELLINO, (1981). *Journal of the American Cnemical Society* **103**, 1201-1207.

FINER-MOORE, J. S., MONTFORT, W. R. & STROUD, R. M. (1990). Pairwise specificity and sequential binding in enzyme catalysis: thymidylate synthase. *Biochemistry* **29**, 6977-6986.

FLECKENSTEIN, B. & DESROSIERS, R. C. (1982). Herpesvirus saimiri and herpesvirus ateles, p 253-332. In B. Roizman (Editor), *The Herpesviruses*, vol. 1. Plenum Publishing Corp., New York.

FLECKENSTEIN, B., BORNKAMM, G. W., MULDER, C., WERNER, F.-J., DANIEL, M. D., FALK, L. A. & DELIUS, H. (1978). Herpesvirus ateles DNA and its homology with herpesvirus saimiri nucleic acid. *Journal of Virology* **25**, 361-373.

FOLKERS, G., TRUMPP-KALLMEYER, S., GUTBROD, O., KRICKL, S., FETZER, J. & KEIL, G. M. (1991). Computer-aided active-site-directed modelling of the herpes simplex virus 1 and human thymidine kinase. *Journal of Computer Aided Design* **5**, 385-404.

FOO, S. K. & SHANE, B. (1982). Regulation of folypoly-gamma-glutamate synthesis in mammalian cells. *In vivo* and *in vitro* synthesis of pteroylpoly-gamma-glutamates by Chinese hamster ovary cells. *Journal of Biological Chemistry* **257**, 13587-13592.

FOOTE, S. J., GALATIS, D. G. & COWMAN, A. F. (1990). Amino acids in the dihydrofolate reductase-thymidylate synthase gene of *Plasmodium falciparum* involved in cycloguanil resistance differ from those involved in pyrimethamine resistance. *Proceedings of the National Academy of Sciences USA* **87**, 3014-3017.

FRASCA, V., LaPAT-POLASKA, L., MALEY, G. F. & MALEY, F. (1988). Site-directed mutagenesis of the T4-phage thymidylate synthase gene. In *Proceedings of the 1988 Miami Bio/Technology Winter Symposium*, p149 (K. Brew *et al.*, Editors). IRL Press, Washingon, D.C.

FREEMONT, P. S., HANSON, I. M. & TROWSDALE, J. (1991). A novel cysteine-rich sequence motif. *Cell* **64**, 483-484.

FRENKEL, N., SCHIRMER, E. C., WYATT, L. S., KATSAFANAS, G., ROFFMAN, E., DANOVICH, R. M. & JUNE, C. H. (1990). Isolation of a new herpesvirus from human $CD4^+$ T cells. *Proceedings of the National Academy of Sciences, U.S.A.* **87**, 748-752.

FRIEDKIN, M. & KORNBERG, A. (1957). In W. D. McElroy & B. Glass (Editors), *The Chemical Basis of Heredity*, pp. 609-614. The John Hopkins Press, Baltimore.

FRIEDKIN, M. (1959). The transfer of $H^3$ from tritiated tetrahydrofolic acid to thymidylic acid. *Federal Proceedings* **18**, 909.

FRIEDKIN, M., CRAWFORD, E., HUMPHREYS, S. R. & GOLDIN, A. (1962). The association of increased dihydrofolate reductase with amethopterin resistance in mouse leukemia. *Cancer Research* **22**, 600-606.

FRIEDMAN, D., COWARD. J. E., ROSENKRANZ, H. S. & MORGAN, C. (1975). Electron microscopic studies on the assembly of herpes simplex virus and assembly on removal of hydroxyurea block. *Journal of General Virology* **26**, 171-181.

FURLONG, D., SWIFT, H. & ROIZMAN, B. (1972). Arrangement of herpes-virus DNA in the core. *Journal of Virology* **10**, 1071-1074.

FYFE, J. A. (1981). Differential phosphorylation of (*E*)-5-(2-bromovinyl)-2'-deoxyuridine monophosphate by thymidine kinases from herpes simplex viruses types 1 and 2 and varicella-zoster virus. *Molecular Pharmacology* **21**, 432-437.

FYFE, J. A., KELLAR, P. M., FURMAN, P. A., MILLER, R. L. & ELION, G. B. (1978). Thymidine kinase from herpes simplex virus phosphorylates the new antiviral compound 9-(2-hydroxyethoxymethyl)guanine. *Journal of Biological Chemistry* **253**, 8721-8727.

GAILLARD, R., SHORT, S. A., GAILLARD, M., STANAT, S. C. & BIRON, K. K. (1990). Sequence alterations in the DNA polymerase genes of VZV strains selected for *in vitro* resistance to acyclovir and phosphonoacetic acid. *15th International Herpesvirus Workshop*, Georgetown University, Washington D. C.

GALIVAN, J. H., MALEY, G. F. & MALEY, F. (1975). The effect of substrate analogues on the circular dichroic spectra of thymidylate synthetase from *L.casei*. *Biochemistry* **14**, 3338-3344.

GALIVAN, J. H., MALEY, F. & BAUGH, C. M. (1976a). Demonstration of separate binding sites for the folate coenzymes and deoxynucleotides with inactivated *L.casei* thymidylate synthetase. *Biochemical and Biophysical Research Communications* **71**, 527-534.

GALIVAN, J. H., MALEY, G. F. & MALEY, F. (1976b). The effect of substrate analogs on the circular dichroism spectra of thymidylate synthetase from *Lactobacillus casei*. *Biochemistry* **15**, 356-362.

GALIVAN, J., RHEE, M. S., JOHNSON, T. B., DILWITH, R., NAIR, M. G., BUNNI, M. & PRIEST, D. G. (1989). The role of cellular folates in the enhancement of the thymidylate synthase inhibitor 10-propargyl-5,8-dideazafolate against hepatoma cells *in vitro* by inhibitors of dihydrofolate reductase. *Journal of Biological Chemistry* **264**, 10685-10692.

GALLAGHER, J. G. & MERIGAN, T. C. (1979). Prolonged herpes-zoster infection associated with immunosuppressive therapy. *Annals of Internal Medicine* **91**, 842-846.

GARNIER, J. H., OSGUTHORPE, D. J. & ROBSON, B. (1978). Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *Journal of Molecular Biology* **120**, 97-120.

GARRETT, C., WAYATA, Y. & SANTI, D. V. (1979). Thymidylate synthetase. Catalysis of dehalogenation of 5-bromo- and 5-iodo-2'-deoxyuridylate. *Biochemistry* **18**, 2798-2804.

GARVEY, E. P. & SANTI, D. V. (1985). Limited proteolysis of the bifunctional thymidylate synthase-dihydrofolate reductase from *Leishmania tropica*. *Proceedings of the National Academy of Sciences USA* **82**, 7188-7192.

GARVEY, E. P. & SANTI, D. V. (1987). A stable binary complex between *Leishmania major* thymidylate synthase and the substrate deoxyuridylate. *Journal of Biological Chemistry* **262**, 9068-9074.

GEIDER, K. (1986). DNA cloning vectors utilizing replication functions of the filamentous phages of *Escherichia coli*. *Journal of General Virology* **67**, 2287-2303.

GEISER, C. F., BISHOP, Y., MYERS, M., JAFFE, N. & YANKEE, R. (1975). Prophylaxis of varicella in children with neoplastic disease: comparative results with zoster immune plasma and gamma globulin. *Cancer* **35**, 1027-1030.

GELB, L. D., DOHNER, D. E., GERSHON, A. A. *et al.*, (1987). Molecular epidemiology of live, attenuated varicella virus vaccine in children with leukemia and in normal adults. *Journal of Infectious Diseases* **155**, 633-640.

GELB, L. D. (1990). Varicella zoster virus. In *Virology*, 2011-2054. Edited by B. N. Fields, D. M. Knipe, R. M. Chanock, M. S. Hirsch, J. L. Melnick, T. P. Monath & B. Roizman. New York: Raven Press Ltd.

GENTRY, G. A., ALLEN, G. P., HOLTON, R., NEVINS, R. B., McGOWAN, J. J. & VEERISETTY, V. (1983). Thymine salvage, mitochondria, and the evolution of herpesviruses. *Intervirology* **19**, 67-76.

GENTRY, G. A., LOWE, M., ALFORD, G. & NEVINS, R. (1988). Sequence analyses of herpesviral enzymes suggest an ancient origin for human sexual behaviour. *Proceedings of the National Academy of Sciences* **85**, 2658-2661.

GENTZ, R., LANGNER, A., CHANG, A. C. Y., COHEN, S. N. & BUJARD, H. (1981). Cloning and analysis of strong promoters is made possible by the downstream placement of a RNA termination signal. *Proceedings of the National Academy of Sciences, U.S.A.* **78**, 4936-4940.

GERSHON, A. A., STEINBERG, S. & BRUNELL, P. A. (1974). Zoster immune globulin: a further assessment. *New England Journal of Medicine* **13**, 1549-1553.

GERSHON, A. A., STEINBERG, S. P., BORKOWSKY, W., LENNETTE, D. & LENNETTE, E. (1982). IgM to varicella-zoster virus: demonstration in patients with and without clinical zoster. *Pediatrics Infectious Diseases Journal* **1**, 164-167.

GERSHON, A. A., STEINBERG, S. P. & GELB, L. D. (1986). Live attenuated varicella vaccine use in immunocompromised children and adults. *Pediatrics* **78**, 757-762.

GIBSON, T., STOCKWELL, P., GINSBURG, M. & BARRELL, B. (1984). Homology between two EBV early genes and HSV ribonucleotide reductase and 38K genes. *Nucleic Acids Research* **12**, 5087-5099.

GILDEN, D. H., ROZENMAN, Y., MURRAY, R., DEVLIN, M. & VAFAI, A. (1986). Detection of varicella-zoster virus nucleic acid in neurons of normal human thoracic ganglia. *Annals of Neurology* **22**, 377-380.

GILDEN, D. H., VAFAI, A., SHTRAM, Y., BECKER, Y., DEVLIN, M. & WELLISH, M. (1983). Varicella-zoster virus DNA in human sensory ganglia. *Nature* **306**, 478-480.

GILLI, R., LOPEZ, C., SARI, J. C. & BRIAND, C. (1990). Thermodynamic study of the interaction of methotrexate, its metabolites, and new antifolates with thymidylate synthase: influence of FdUMP. *Biochemical Pharmacology* **40**, 2241-2246.

GLUSKER, J. P. & TRUEBLOOD, K. N. (1985). *Crystal Structure Analysis.* New York: Oxford University Press.

GOLDSTEIN, D. J. & WELLER, S. K. (1988). Factor(s) present in herpes simplex virus type 1-infected cells can compensate for the loss of the large subunit of the viral ribonucleotide reductase: characterization of an ICP6 deletion mutant. *Virology* **166**, 41-51.

GRAY, W. L., BAUMAN, R. P., ROBERTSON, A. T., CAUGHMAN, G. B., O'CALLAGHAN, D. J. & STACZEK, J. (1987). Regulation of equine herpesvirus type 1 gene expression: characterization of immediate early, early and late transcription. *Virology* **158**, 79-87.

GREEN, M. T., COURTNEY, R. J. & DUNKEL, E. C. (1981). Detction of an immediate early herpes simplex type 1 polypeptide in trigeminal ganglia from latently infected animals. *Infections and Immunity* **34**, 987-992.

GREENHALGH, D. A. & PARISH, J. H. (1989). Effects of 5-fluorouracil on cytotoxicity and RNA metabloism in human colonic carcinoma cells. *Cancer Chemotherapy and Pharmacology* **25**, 37.

GREENHALGH, D. A. & PARISH, J. H. (1990). Effect of 5-fluorouracil combination therapy on RNA processing in human colonic carcinoma cells. *British Journal of Cancer* **61**, 415-419.

GRIFFIN, A. M. & BOURSNELL, M. E. G. (1990). Analysis of the nucleotide sequence of DNA from the region of the thymidine kinase gene of infectious laryngotracheitis virus; potential evolutionary relationship between the herpesvirus subfamilies. *Journal of General Virology* **71**, 841-850.

GROS, F. (1960). Dynamik des Eisweisses. *Colloquium der Gesellschaft fr physiologische chemie* **10**, p82. Springer, Berlin.

GRUMONT, R., WASHTEIN, W. L., CAPUT, D. & SANTI, D. V. (1986). Bifunctional thymidylate synthase-dihyrofolate reductase from *Leishmania tropica*: sequence homology with the corresponding monofunctional proteins. *Proceedings of the National Academy of Sciences USA* **83**, 5387-5391.

GU, X. & SANTI, D. V. (1991). The T-arm of tRNA is a substrate for tRNA($m^5$U54)-methyltransferase. *Biochemistry* **30**, 2999-3002.

GUESS, H. A. (1986). Population-based studies of varicella complications. *Pediatrics* **78**, 723-727.

GUSTAFSON, D. P. (1981). Herpesvirus diseases of mammals and birds: Comparitive aspects and diagnosis. In *Comparative Diagnosis of Viral Disease*, vol. III, pp205-263. Editors E. Kurstak & C. Kurstak. New York: Academic Press.

GUSTAFSSON, C., LINDSTROM, P. H., HAGERVALL, T. G., ESBERG, K. B. & BJOERK, G. R. (1991). The *trmA* promoter has regulatory features and sequence elements in common with the rRNA P1 promoter family of *Escherichia coli*. *Journal of Bacteriology* **173**, 1757-1764.

HAMER, D. H. & LEDER, P. (1979). Splicing and the stable formation of RNA. *Cell* **18**, 1299-1302.

HAMMERSCHMIDT, W., LUDWIG, H. & BUHK, H.-J. (1988). Specificity of cleavage in replicative-form DNA of bovine herpesvirus 1. *Journal of Virology* **62**, 1355-1363.

HAMPSEY, M. D., DAS, G. & SHERMAN, F. (1988). *FEBS Letters* **231**, 275.

HANAHAN, D. (1983). Studies on transformation of *E.coli* with plasmids. *Journal of Molecular Biology* **166**, 557-580.

HANTKE, K. (1976). Phage T6-colicin K receptor and nucleoside transport in *Escherichia coli*. *FEBS Letters* **70**, 109-112.

HARDY, L. W. & POTEETE, A. R. (1991). Re-examination of the role of Asp20 in catalysis by bacteriophage T4 lysozyme. *Biochemistry* In Press,

HARDY, L. W., FINER-MOORE, J. S., MONTFORT, W. R., JONES, M. O., SANTI, D. V. & STROUD, R. M. (1987). Atomic structure of thymidylate synthase: target for rational drug design. *Science* **235**, 448-455.

HARNDEN, M. R., WYATT, P. G., BOYD, M. R. & SUTTON, D. (1990). Synthesis and antiviral activity of 9-alkoxypurines. 1. 9-(3-hydroxypropoxy)- and 9-[3-hydroxy-2-(hydroxymehtyl)propoxy]-purines. *Journal of Medicinal Chemistry* **33**, 187-196.

HARRAP, K. R., JACKMAN, A. L., NEWELL, D. R., TAYLOR, G. A., HUGHES, L. R. & CALVERT, A. H. (1989). Thymidylate synthase: a target for anticancer drug design. *Advances in Enzyme Regulation* **29**, 161-179.

HARRISON, P. T., THOMPSON, R. & DAVISON, A. J. (1991). Evolution of the herpesvirus thymidine kinase from a cellular deoxycytidine kinase. *Journal of General Virology* **72**, 2583-2586.

HARTMANN, K.-U. & HEIDELBERGER, C. (1961). Studies on fluorinated pyrimidines XIII. Inhibition of thymidylate synthetase. *Journal of Biological Chemistry* **236**, 3006-3013.

HARTWELL, L. H. (1973). Three additional genes required for deoxyribonucleic acid synthesis in *Saccharomyces cerevisiae*. *Journal of Bacteriology* **115**, 423-428.

HASTIE, N. D. & HELD, W. A. (1978). Analysis of mRNA populations by cDNA.mRNA hybrid-mediated inhibition of cell-free protein synthesis. *Proceedings of the National Academy of Sciences USA* **75**, 1217-1221.

HATTORI, M. & SAKAKI, Y. (1986). Dideoxy sequencing method using denatured plasmid templates. *Analytical Biochemistry* **152**, 232-238.

HAYAKAWA, Y., TORIGOE, S., SHIRAKI, K., YAMANISHI, K. & TAKAHASHI, M. (1984). Biologic and biophysical markers of a live varicella vaccine strain (Oka): identification of clinical isolates from vaccine recipients. *Journal of Infectious Diseases* **149**, 956-963.

HEAD, H. & CAMPBELL, A. W. (1900). The pathology of herpes zoster and its bearing on sensory localization. *Brain* **23**, 353-532.

HELENE, C. & TOULME, J.-J. (1990). Specific regulation of gene expression by antisense, sense and antigene nucleic acids. *Biochimie et Biophysica Acta* **1049**, 99-125.

HEIDELBERGER, C., CHAUDHURI, N. K., DANENBERG, P. V., MOOREN, D., GRIESBACH, L., DUSCHINSKY, R., SCHNITZER, R. J., PLEVEN, E. & SCHNEIDER, T. (1957). A new class of fluorinated pyrimidines, tumour inhibitory compounds. *Nature* **179**, 663-666.

HEIDELBERGER, C., GHOBAR, A., BAKER, R. K. & MUKHERJEE, K. L. (1960a). Studies on fluorinated pyrimidines X. *In vivo* studies on tumor resistance. *Cancer Research* **20**, 897-902.

HEIDELBERGER, C., KALDOR, G., MUKHERJEE, K. L., & DANENBERG, P. B. (1960b). Studies on fluorinated pyrimidines XI. *In vitro* studies on tumor resistance. *Cancer Research* **20**, 903-909.

HEIDELBERGER, C., DANENBERG, P. V. & MORAN, R. G. (1983). Fluorinated pyrimidines and their nucleosides. *Advances in Enzymology* **54**, 57-119.

HERMES, J. D., BLACKLOW, S. C. & KNOWLES, J. R. (1990). Searching sequence space by definably random mutagenesis: improving the catalytic potency of an enzyme. *Proceedings of the National Academy of Sciences, U.S.A.* **87**, 696-700.

HERMES, J. D., PAREKH, S. M., BLACKLOW, S. C., KOSTER, H. & KNOWLES, J. R. (1989). A reliable method for random mutagenesis: the generation of mutant libraries using spiked oligodeoxyribonucleotide primers. *Gene* **84**, 143-151.

HILGEN, S. E. & PIELAK, G. J. (1991). The function of the *Saccharomyces cerevisiae* iso-1-cytochrome *c* gene is independent of the codon at invariant Phe82 when the gene is expressed on a low copy number plasmid. *Protein Engineering* **4**, 575-578.

HOLM, L., KOIVULA, A. K., LEHTOVAARA, P. M., HEMMINKI, A. & KNOWLES, J. K. C. (1990). Random mutagenesis used to probe the structure and function of *Bacillus stearothermophilus* alpha-amylase. *Protein Engineering* **3**, 181-191.

HOLMES, S. M. & QUIGLEY, M. (1981). A rapid boiling method for the preparation of bacterial plasmids. *Analytical Biochemistry* **114**, 193-197.

HONDO, R. & YOGO, Y. (1988). Strain variation of R5 direct repeats in the right-hand portion of the long unique segment of varicella-zoster virus DNA. *Journal of Virology* **62**, 2916-2921.

HONDO, R., YOGO, Y., KURATA. T. & AOYAMA, Y. (1987). Genome variation among varicella-zoster virus isolates derived from different individuals and from the same individual. *Archives of Virology* **93**, 1-12.

HONESS, R. W. (1984). Herpes simplex and 'the herpes complex': diverse observations and a unifying hypothesis. *Journal of General Virology* **65**, 2077-2107.

HONESS, R. W. & ROIZMAN, B. (1974). Regulation of herpes-virus macromolecular synthesis. I. Cascade regulation of the synthesis of three groups of viral proteins. *Journal of Virology* **14**, 8-19.

HONESS, R. W. & WATSON, D. H. (1977). Unity and diversity in the herpesviruses. *Journal of General Virology* **37**, 15-37.

HONESS, R. W., BODERMER, W., CAMERON, K. R., NILLER, H.-H., FLECKENSTEIN, B. & RANDALL, R. E. (1986). The A+T-rich genome of herpesvirus saimiri contains a highly conserved gene for thymidylate synthase. *Proceedings of the National Academy of Sciences, U.S.A.* **83**, 3604-3608.

HONESS, R. W. CRAXTON, M. A., WILLIAMS, L. & GOMPELS, U. A. (1977). A comparative analysis of the sequence of the thymidine kinase gene of a gammaherpesvirus, herpes-virus samiri. *Journal of General Virology* **37**, 15-37.

HONG, G.-F. (1987). The use of DNaseI, buffer gradient gel, and $^{35}S$ label for DNA sequencing. *Methods in Enzymology* **155**, 93-110.

HOPE-SIMPSON. R. E. (1965). Infectiousness of communicable diseases in the household (measles, chickenpox and mumps). *Lancet* **2**, 549-554.

HORI, T., AYUSAWA, D., SHIMUZU, K., KOYAMA, H. & SENO, T. (1984a). Chromosome breakage induced by thymidylate stress in thymidylate synthase-negative mutants of mouse FM3A cells. *Cancer Research* **44**, 703-709.

HORI, T., AYUSAWA, D. & SENO, T. (1984b). Thymidylate stress and sister chromatid exchanges. In R. R. Rice & A. Hollander (Editors). *Sister Chromatid Exchanges*, Plenum, New York.

HOROWITZ, J. & CHARGAFF, E. (1959). Massive incorporation of 5-FU into a bacterial ribonucleic acid. *Nature* **184**, 1213-1215.

HOROWITZ, J., SAUKKONEN, J. J. & CHARGAFF, E. (1960). Effects of fluoropyrimidines on the synthesis of bacterial proteins and nucleic acids. *Journal of Biological Chemistry* **235**, 3266-3272.

HORTEN, B., PRICE, R. W. & JIMENEZ, D. (1981). Multifocal varicella-zoster virus leukoencephalitis temporally remote from herpes zoster. *Ann. Neurol.* **9**, 251-266.

HOUGHTON, J. A., MARODA, S. J., PHILIPS, J. O. & HOUGHTON, P. J. (1981). Biochemical determinants of responsiveness to 5-fluorouracil and its derivatives in human colorectal adenocarcinoma xenografts. *Cancer Research* **41**, 144-149.

HOUGHTON, P. J., GERMAIN, G. S., HAZELTON, B. J., PENNINGTON, J. W. & HOUGHTON, J. A. (1989). Mutants of human colon adenocarcinoma, selected for thymidylate synthase deficiency. **86**, 1377-1381.

HUANG, S., TOMICH, J. M., WU, H., JONG, A. & HOLCENBERG, J. (1989). Human deoxycytidine kinase. Sequence of cDNA clones and analysis of expression in cell lines with and without enzyme activity. *Journal of Biological Chemistry* **264**, 14762-14768.

HUANG, S., TOMICH, J. M., WU, H., JONG, A. & HOLCENBERG, J. (1991). Human deoxycytidine kinase. Sequence of cDNA clones and analysis of expression in cell lines with and without enzyme activity. (Correction). *Journal of Biological Chemistry* **265**, 5353.

HUGHES, D. E., OPEOLU, A., SHONEKAN, E. A. & SIMPSON, L. (1989). Structure, genomic organization and transcription of the bifunctional dihydrofolate reductase-thymidylate synthase gene from *Crithidia fasciculata. Molecular Parasitology* **34**, 155-166.

HUGHES, L. R., JACKMAN, A. L., OLDFIELD, J., SMITH, R. C., BURROWS, K. D., MARSHAM, P. R., BISHOP, J. A. M., JONES, T. R., O'CONNOR, B. M., CALVERT, A. H. (1990). Quinazoline antifolate thymidylate synthase inhibitors: alkyl, substituted alkyl, and aryl substituents in the C2 position. *Journal of Medicinal Chemistry* **33**, 3060-3067.

HUMPRHEYS, G. K. & GREENBERG, D. M. (1958). Studies on the conversion of deoxyuridylic acid to thymidylic acid by a soluble extract from rat thymus. *Archives of Biochemistry and Biophysics* **78**, 275.

HUMPREYS, G. O., WESTON, A., BROWN, M. G. M. & SAUNDERS, J. R. (1979). Plasmid transformation in *E.coli*. In *Transformation*. GLOVER, S. W. & BORTER, L. O. (Eds.), pp254-279. Oxford Cotswold Press.

HURWITZ, E. S., NELSON, D. B., DAVIS, C., MORENS, D. & SCHONBERGER, L. B. (1982). National surveillance for Reye syndrome: a five year review. *Pediatrics* **70**, 895-900.

HUTCHINSON, C. A. III., NORDEEN, S. K., VOGT, K. & EDGELL, M. H. (1986). A complete library of point substitution mutations in the glucocorticoid response element of mouse mammary tumor virus. *Proceedings of the National Academy of Sciences, U.S.A.* **83**, 710-714.

HYMAN, R. W., ECKER, J. R. & TENSER, R. B. (1983). Varicella-zoster virus RNA in human trigeminal ganglia. *Lancet* **2**, 814-816.

IIGO, M., NISHIKATA, K.-I., NAKAJIMA, Y., SZINAI, I., VERES, Z., SZABOLCS, A. & DE CLERCQ, E. (1990). Differential effects of 2,2'-anhydro-5-ethyluridine, a uridine phosphorylase inhibitor, on the antitumour activity of 5-fluorouridine and 5-fluoro-2'-deoxyuridine. *Biochemical Pharmacology* **39**, 1247-1253.

INCHAUSPE, G. & OSTROVE, J. M. (1989). Differential regulation by varicella-zoster virus and herpes simplex virus type-1 trans-activating genes. *Virology* **173**, 710-714.

INCHAUSPE, G., NAGPAL, S. & OSTROVE, J. M. (1989). Mapping of two varicella virus-encoded genes that activate the expression of viral early and late genes. *Virology* **173**, 700-709.

INGLIS, S. C., GUILLEMETTE, J. G., JOHNSON, J. A. & SMITH, M. (1991). Analysis of the invariant Phe82 residue of yeast iso-1-cytochrome *c* by site-directed mutagenesis using a phagemid yeast shuttle vector. *Protein Engineering* **4**, 569-575.

IRMIERE, A. & GIBSON, W. (1983). Isolation and characterization of a noninfectious virion-like particle released from cells infected with human strains of cytomegalovirus. *Virology* **130**, 118-133.

IVANETICH, K. M. & SANTI, D. V. (1990). Thymidylate synthase-dihydrofolate reductase in protozoa. *Experimental Parasitology* **70**, 367-371.

IWAKURA, M., KAWATA, M., TSUDA, K. & TANAKA, T. (1988). Nucleotide sequence of the thymidylate synthase B and dihydrofolate reductase genes contained in one *Bacillus subtilis* operon. *Gene* **64**, 9-20.

JACKMAN, A. L., MARSHAM, P. R., THORNTON, T. J., BISHOP, J. A. M., O'CONNOR, B. M., HUGHES, L. R., CALVERT, A. H., JONES, T. R. (1990). Quinazoline antifolate thymidylate synthase inhibitors: 2'-fluoro-$N^{10}$-propargyl-5,8-dideazafolic acid and derivatives with modifications in the C2 position. *Journal of Medicinal Chemistry* **33**, 3067-3071.

JACKSON, R. C., JACKMAN, A. L. & CALVERT, A. H. (1983). Biochemical effects of a quinazoline inhibitor of thymidylate synthetase, *N*-(4-(*N*-((2-amino-4-hydroxy-6-quinazolinyl)methyl)prop-2-ynylamino)benzoyl)L-glutamic acid (CB3717) on human lymphoblastoid cells. *Biochemical Pharmocology* **32**, 3783-3790.

JACOB, R. J., MORSE, L. S. & ROIZMAN, B. (1979). Anatomy of herpes simplex virus DNA. XII. Accumulation of head-to-tail concatemers in nuclei of infected cells and their role in the generation of the four isomeric arrangements of viral DNA. *Journal of Virology* **29**, 448-457.

JACOBSON, M. A., BERGER, T. G., FIKRIG, S., BECHERER, P., MOOHR,J. W., STANAT, S. C. & BIRON, K. K. (1990). Acyclovir-resistant varicella-zoster virus infection after chronic oral acyclovir therapy in patients with the aquired immune deficiency syndrome (AIDS). *Annals of Internal Medicine* **112**, 187-191.

JASTREBOFF, M., KEDZIERSKA, B. & RODE, W. (1982). Properties of thymidylate synthetase from Ehrlich ascites carcinoma cells. *Biochemical Pharmacology* **31**, 217-223.

JASTREBOFF, M., KEDZIERSKA, B. & RODE, W. (1983). Altered thymidylate synthetase in 5-fluorodeoxyuridine-resistant Erhlich ascites carcinoma cells. *Biochemical Pharmacology* **32**, 2259-2267.

JAZWINSKI, S. M. & EDELMAN, G. M. (1984). Evidence for participation of a multiprotein complex in yeast DNA replication *in vitro*. *Journal of Biological Chemistry* **259**, 6852-6857.

JENH, C.-H., DENG, T., LI, D., DeWILLE, J. W. & JOTNSON, L. F. (1986). Mouse thymidylate synthase messenger RNA lacks a 3' untranslated region. *Proceedings of the National Academy of Sciences USA* **83**, 8482-8486.

JENH, C.-H., GEYER, P. K. & JOHNSON, L. F. (1985). Control of thymidylate synthase messenger RNA content and gene transcription in an overproducing mouse cell line. *Molecular and Cellular Biology* **5**, 2527-2532.

JOHNSON, M. T., LUBY, J. P., BUCHANAN, R. A. & MIKULEC, D. (1975). Treatment of varicella-zoster virus infections with adenine arabinoside. *Journal of Infectious Diseases* **131**, 225-229.

JOLLIFF, K., LI, Y. & JOHNSON, L. F. (1991). Multiple protein-DNA interactions in the TATAA-less mouse thymidylate synthase promoter. *Nucleic Acids Research* **19**, 2267-2273.

JONES, P. C. & ROIZMAN, B. (1979). Regulation of herpesvirus macromolecular synthesis. VIII. The transcription programme consists of three phases during which both extent of transcription and accumulation of RNA in the cytoplasm are regulated. *Journal of Virology* **31**, 299-314.

JONES, T. R., BETTERIDGE, NEIDLE, S., JACKMAN, A. L. & CALVERT, A. H. (1989). Quinazoline antifolates inhibiting thymidylate synthase: computer modelling of the $N^{10}$ substituent. *Anti-Cancer Drug Design* **3**, 243-248.

JONES, T. R., CALVERT, A. H., JACKMAN, A. L., BROWN, S. J., JONES, M. & HARRAP, K. R. (1981). A potent antitumour quinazoline inhibitor of thymidylate synthetase: synthesis, biological properties an therapeutic results in mice. *European Journal of Cancer* **17**, 11-19.

JOYCE, C. M. & GRINDLEY, N. D. F. (1984). *Journal of Bacteriology* **158**, 636.

JUEL-JENSON, B. W. & MacCALLUM, F. O. (1972). In *Herpes simplex, varicella and zoster.* Philadelphia: Lippincott.

JURA, E., CHADWICK, E. G., JOSEPHS, S. H., STEINBERG, S. P., YOGEV, R., GERSHON, A. A., KRANSINSKI, K. M. & BORKOWSKY, W. (1989). Varicella-zoster virus infections in children infected with human immunodeficiency virus. *Pediatrics and Infectious Diseases Journal* **8**, 586-590.

KALLANDER, C. F. R., GRONOWITZ, J. S. & OLDING-STENKVIST, E. (1989). Varicella-zoster virus deoxythymidine kinase is present in serum before the onset of varicella. *Scandanavian Journal Infectious Diseases* **21**, 255-257.

KALLEN, R. G. & JENCKS, W. P. (1966). The mechanism of condenstion of formaldehyde with tetrahydrofolic acid. *Journal of Biological Chemistry* **241**, 5821

KAEHLER, R., STRAUSS, M. & KIESSLING, U. (1984). A well transformable *E.coli tdk* strain suitable for direct rescue of tk gene plasmids from mammalian cells. *Biomedica et Biochemica Acta* **43**, K25-K29.

KANEDA, S., NALBANTOGLU, J., TAKEISHI, K., SHIMIZU, K., GOTOH, O., SENO, T. & AYUSAWA, D. (1990). Structural and functional analysis of the human thymidylate synthase gene. *The Journal of Biological Chemistry* **265**, 20277-20284.

KENNEDY, P. G. E., AL-SAADI, S. A. & GLEMENTS, G. B. (1983). Reactivation of herpes simplex virus from dissociated identified dorsal root ganglion cells in culture. *Journal of General Virology* **64**, 1629-1635.

KENNY, E., ATKINSON, T. & HARTLEY, B. S. (1985). Nucleotide sequence of the thymidylate synthetase gene (ThyP[3]) from the *Bacillus subtilis* phage m3T. *Gene* **34**, 335-342.

KESSEL, D. & WODINSKY, I. (1970). Thymidine kinase as a determinant of the response to FdUMP in transplantable murine leukemias. *Molecular Pharmacology* **6**, 251-254.

KEYOMARSI, K. & MORAN, R. G. (1988). Mechanism of the cytotoxic synergism of fluoropyrimidines and folinic acid in mouse leukemic cells. *Journal of Biological Chemistry* **263**, 14402-14409.

KEYOMARSI, K. & MORAN, R. G. (1990). Quinazoline folate analogs inhibit the catalytic activity of thymidylate synthase but allow binding of 5-fluorodeoxyuridylate. *Journal of Biological Chemistry* **265**, 19163-19169.

KINCHINGTON, P. R., REMENICK, J., OSTROVE, J. M., STRAUSS, S. E., RUYECHAN, W. T. & HAY, J. (1986). Putative glycoprotein gene of varicella-zoster virus with variable copy numbers of a 42-base-pair repeat sequence has homology to herpes simplex virus glycoprotein C. *Journal of Virology* **59**, 660-668.

KINOSHITA, H., HONDO, R., TAGUCHI, F. & YOGO, Y. (1988). Variation of R1 repeated sequence present in open reading frame 11 of varicella-zoster virus strains. *Journal of Virology* **62**, 1097-1100.

KIRKPATRICK, C. M., KANITZ, C. L. & McCROCKLIN, S. M. (1980). Possible role of wild animals in transmission of pseudorabiesvirus to swine. *Journal Wildl. Dis.* **16**, 601-614.

KISLIUK, R. L., GAUMONT, Y. & BAUGH, C. M. (1974). Polyglutamyl derivatives of folate as substrates and inhibitors of thymidylate synthetase. *Journal of Biological Chemistry* **249**, 4100-4103.

KIT, S. (1985). Thymidine kinase. *Microbiological Sciences* **2**, 369-375.

KIT, S., LEUNG, W.-C., JORGENSEN, G. N., TRKULA, D. & DUBBS, D. R. (1974). Thymidine kinase isozymes of normal and virus infected cells. *Cold Spring Harbour Symposia on Quantitative Biology* **XXXIX**, 703-715.

KLEINA, L. G. & MILLER, J. H. (1990). Genetic studies of the *lac* repressor. *Journal of Molecular Biology* **212**, 295-318.

KLEINA, L. G., MASSON, J.-M., NORMANLY, J., ABELESON, J. & MILLER, J. H. (1990). Construction of *Escherichia coli* amber suppressor tRNA genes. *Journal of Molecular Biology* **213**, 705-717.

KLIMASAUSKAS, S., TIMINSKAS, A., MENKEVICIUS, S., BUTKIENE, D., BUTKUS, V. & JANULAITIS, A. (1989). Sequence motifs characteristic of DNA[cytosine-N4]methyltransferases: similarity to adenine and cytosine-C5 DNA-methylases. *Nucleic Acids Research* **17**, 9823-9832.

KOSZ-VNENCHAK, M., COEN, D. M. & KNIPE, D. M. (1990). Restricted expression of herpes simplex virus lytic genes during establishment of latent infection by thymidine kinase-negative mutant viruses. *Journal of Virology* **64**, 5396-5402.

KOYAMA, H., AYUSAWA, D., TSUJI, M. & SENO, T. (1982). Thymineless death and mutation induction in cultured mouse FM3A cell mutants deficient in thymidylate synthase. *Mutation Research* **105**, 433-438.

KRAMER, W. & FRITZ, H.-J. (1987). Oligonucleotide-directed construction of mutations via gapped duplex DNA. *Methods in Enzymology* **154**, 350-367.

KRUGMAN, S., WARD, R. & KATZ, R. L. (1977). Varicella-zoster infections. In *Infectious diseases of children,* 6th edition, 451-471. St. Louis: Mosby.

KUNKEL, T. A. (1985). Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proceedings of the National Academy of Sciences U.S.A.* **82**, 488-492.

KUNKEL, T. A., ROBERTS, J. D. & ZAKOUR, R. A. (1987). Rapid and efficient site-specific mutagenesis without phenotypic selection. *Methods in Enzymology* **54**, 367-382.

KUNZ, B. A. & HAYNES, R. H. (1982). DNA repair and the genetic effects of thymidylate stress in yeast. *Mutation Research* **93**, 353-375.

KWOH, T. J. & ENGLER, J. A. (1984). The nucleotide sequence of the chicken thymidine kinase gene and the relationship of its predicted polypeptide to that of the vaccinia virus thymidine kinase. *Nucleic Acids Research* **12**, 3959-3971.

KYPRIANOU, N. & ISAACS, J. T. (1989). "Thymineless" death in androgen-independent prostatic cancer cells. *Biochemical and Biophysical Research Communications* **165**, 73-81.

LACEY, S. F., SUZUTANI, T., POWELL, K. L., PURIFOY, D. J. M. & HONESS (1991). Analysis of mutations in the thymidine kinase genes of drug-resistant varicella-zoster virus populations using the polymerase chain reaction. *Journal of General Virology* **72**,

LAGOSKY, P. A., TAYLOR, G. R. & HAYNES, R. H. (1987). Molecular characterization of the *Saccharomyces cerevisiae* dihydrofolate reductase gene (*DFR1I*). *Nucleic Acids Research* **15**, 10355-10371.

LAKE-BAKAAR, D. M., LINDBORG, B. & DATEMA, R. (1989). Improvement of the absorbtion of oral (*R,S*)-9-[4-hydroxy-2-(hydroxymethyl)butyl]guanine, an anti-varicella-zoster virus drug, in rats and monkeys. *Antimicrobial Agents and Chemotherapy* **33**, 110-112.

LAMM, N., TOMASCHEWSKI, J. & RUGER, W. (1987). Nucleotide sequence of the deoxycytidylate hydroxymethylase gene of bacteriophage T4 (g42) and the homology of its gene product with thymidylate synthase of *E.coli*. *Nucleic Acids Research* **15**, 3920.

LANGENBACH, R. (1976). Thymidylate synthetase of *in vitro* chemically and virally transformed rat cells. *Biochimie et Biophysica Acta* **422**, 295-301.

LAPAT-POLASKO, L., MALEY, G. F. & MALEY, F. (1990). Properties of T4-phage thymidylate synthase following mutagenic changes in the active site region and folate binding region. *Biochemistry* **29**, 9561-9572.

LARDER, B. A. & DARBY, G. (1984). Virus drug resistance: mechanisms and consequences. *Antiviral Research* **4**, 1-42.

LARDER, B. A. & KEMP, S. D. (1989). Multiple mutations in HIV reverse transcriptase confer high-level resistance to zidovudine (AZT). *Science* **246**, 1155-1158.

LARDER, B. A., CHENG, Y.-C. & DARBY, G. (1983). Characterization of abnormal thymidine kinases induced by drug-resistant strains of herpes simplex virus type 1. *Journal of General Virology* **64**, 523-532.

LARDER, B. A., COATES, K. C. & KEMP, S. D. (1991). Zidovudine-resistant human immunodeficiency virus selected by passage in cell culture. *Journal of Virology* **65**, 5233-5236.

LAWRENCE, G. L., CHEE, M. S., CRAXTON, M. A., GOMPELS, U. A., HONESS, R. W. & BARRELL, B. G. (1990). Human herpesvirus 6 is closely related to human cytomegalovirus. *Journal of Virology* **64**, 287-299.

LAZAR, G., TOTH, I., HAAKONSEN, L. & GOODMAN, H. M. (1989). Coregulation of dihydrofolate reductase and thymidylate synthase in overproducer cell lines of wild carrot. *Plant Physiology* **91**, 1168-1173.

LEARY, R. P., BEAUDETTE, N. & KISLUIK, R. L. (1975). Interaction of deoxyuridylate with thymidylate synthetase. *Journal of Biological Chemistry* **250**, 4869-4878.

LEHTOVAARA, P. M., KOIVULA, A. K., BAMFORD, J. & KNOWLES, J. K. C. (1988). A new method for the random mutagenesis of complete genes: enzymatic generation of mutant libraries *in vitro*. *Protein Engineering* **2**, 63-68

LEPAULT, J., DUBOCHET, J., BASCHONG, W. & KELLENBERG, E. (1987). Organization of double-stranded DNA in bacteriophage: a study by cryo-electron microscopy of vitrified samples. *EMBO Journal* **6**, 1507-1512.

LESK, A. M. & CHOTHIA, C. (1980). How amino acid sequences determine similar protein structures. *Journal of Molecular Biology* **136**, 225-270.

LEWIS, C. A. & DUNLAP, R. B. (1981). Thymidylate synthase and its interaction with 5-fluoro-2'-deoxyuridylate. In *Topics in Molecular Pharmacology*, 170-219. Edited by A. S. V. Burgen & G. C. K. Roberts. Elsevier/North Holland Biomedical Press.

LEWIS, G. W. (1958). Zoster sine herpete. *British Medical Journal* **2**, 418-421.

LIM, W. A. & SAUER, R. T. (1989). Alternative packing arrangements in the hydrophobic core of lambda repressor. *Nature* **339**, 31-36.

LIM, W. A. & SAUER, R. T. (1991). The role of internal packing interactions in determining the structure and stability of a protein. *Journal of Molecular Biology* **219**, 359-376.

LIM, V. I. (1974). Algorithms for prediction of '-helical and ß-structural regions in globular proteins. *Journal of Molecular Biology* **88**, 873-894,

LING, P., KINCHINGTON, P. R., RUYECHAN, W. T. & HAY, J. (1991). A detailed analysis of transcripts mapping to varicella zoster virus gene 14 (glycoprotein V). *Virology* **184**, 625-635.

LINNEMANN, C. C. Jr., BIRON, K. K., HOPPENJANS, W. G. & SOLINGER, A. M. (1990). Emergence of acyclovir-resistant varicella-zoster virus in an AIDS patient on prolonged acyclovir treatment. *AIDS* **4**, 577-579.

LOCKSHIN, A. & DANENBERG, P. V. (1980). Hydrodynamic behaviour of human and bacterial thymidylate synthetase and TS-5-fluorodeoxyuridylate-methylenetetrahydrofolate complexes. Evidence for large conformational changes during catalysis. *Biochemistry* **19**, 4244-4251.

LOCKSHIN, A. & DANENBERG, P. V. (1981). Biochemical factors affecting the tightness of 5-fluorodeoxy-uridylate binding of thymidylate synthase. *Biochemical Pharmacology* **30**, 247-257.

LOPETEGUI, P., CAMPO-VERA, H. & YAMANISHI, K. (1985). Varicella-zoster virus (VZV)-specific polypeptides detected in cells treated with metabolic inhibitors. *Microbiology and Immunology* **29**, 569-575.

LOPEZ, C., PELLET, P., STEWART, J., GOLDSMITH, C., SANDERLIN, K., BLACK, J., WARFIELD, D. & FEORINO, P. (1988). Characteristics of human herpesvirus 6. *Journal of Infectious Diseases* **157**, 1271-1273.

LU, Y.-Z., AIELLO, P. D. & MATTHEWS, R. G. (1984). Studies on the polyglutamate specificity of thymidylate synthetase from fetal pig liver. *Biochemistry* **23**, 6870-6876.

LUBY, J. P., RAMIREZ-RONDA, C., RINNER, S., HULL, A., VERGNE-MARINI, P. (1977). A longitudinal study of varicella zoster infections in renal transplant patients. *Journal of Infectious Diseases* **135**, 659-663.

McCOMBS, R., BRUNSCHWIG, J. P., MIRKOVIC, R. & BENYESH-MELNICK, M. (1971). Electrom microscopic characterization of a herpes-like virus isolated from tree shrews. *Virology* **45**, 816-820.

McGEOCH, D. J. (1989). The genomes of the human herpesviruses: contents, relationships and evolution. *Annual Review of Microbiology* **43**, 235-265.

McGEOCH, D. J., DOLAN, A., DONALD, S. & BRAUER, D. H. K. (1986). Complete DNA sequence of the short repeat region in the genome of herpes simplex virus type 1. *Nucleic Acids Research* **14**, 1727-1745.

McGEOCH, D. J., DALRYMPLE, M. A., DAVISON, A. J., DOLAN, A., FRAME, M. C., McNAB, D., PERRY, L. J., SCOTT, J. E. & TAYLOR, P. (1988). The complete DNA sequence of the long unique region in the genome of herpes simplexvirus type 1. *Journal of General Virology* **69**, 1531-1574.

McKEE, T. A. (1990). The control of varicella-zoster virus immediate early gene expression. *Ph.D. Thesis*, University of Glasgow.

McKEE, T. A., DISNEY, G. H., EVERETT, R. D. & PRESTON, C. M. (1990). Control of expression of the varicella-zoster virus major immediate early gene. *Journal of General Virology* **71**, 897-906.

McKEOWN, M., KAHN, M. & HANAWALT, P. (1976). Thymidine uptake and utilization in *Escherichia coli*: a new gene controlling nucleoside transport. *Journal of Bacteriology* **126**, 814-822.

McLAUCHLAN, J., GAFFNEY, D., WHITTON, J. L. & CLEMENTS, J. B. (1985). The consensus sequence YGTGTTYY located downstream from the AATAAA signal is required for efficient formation of mRNA 3'terminii. *Nucleic Acids Research* **13**, 1347-1368.

McLENNON, J. L. & DARBY, G. (1980). Herpes simplex virus latency: the cellular location of virus in dorsal root ganglia and the fate of the infected cells following virus activation. *Journal of General Virology* **51**, 233-243.

McNEIL, J. B. & SMITH, M. (1985). *Saccharomyces cerevisiae* CYC1 mRNA 5'-end positioning by *in vitro* mutagenesis, using synthetic duplexes with random mismatch base pairs. *Molecular and Cellular Biology* **5**, 3545-3551.

McPHERSON (1991). *Directed Mutagenesis.* Oxford University Press, New York.

MACHIDA, H., KUNINAKA, A. & YOSHINO, H. (1982). Inhibitory effects of antiherpesviral thymidine analogs against varicell-zoster virus. *Antimicrobial Agents and Chemotherapy* **21**, 358-361.

MACHIDA, H. (1986). Comparison of susceptibilities of varicella-zoster virus and herpes simplex viruses to nucleoside analogues. *Antimicrobial Agents and Chemotherapy* **29**, 524-526.

MACKEM, S. & ROIZMAN, B. (1982). Regulation of alpha genes of herpes simplex virus 1: the functional domains and sequence of a movable alpha regulator. *Proceedings of the National Academy of Sciences, U.S.A.* **79**, 4917-4921

MAEHARA, Y., MORIGUCHI, S., EMI, Y., WATANABE, A., KOHNOE, S., TSUJITANI, S. & SUGIMACHI, K. (1990). Comparison of pyrimidine nucleotide synthetic enzymes involved in 5-fluorouracil metabolism between human adenocarcinomas and squamous cell carcinomas. *Cancer* **66**, 156-161.

MAGUIRE, H. F. & HYMAN, R. W. (1986). Polyadenylated cytoplasmic transcripts of varicella-zoster virus. *Intervirology* **26**, 181-191.

MALEY, G. F. & MALEY, F. (1988). Properties of a defined mutant of *Escherichia coli* thymidylate synthase. *Journal of Biological Chemistry* **263**, 7620-7627.

MALEY, G. F., BELLISARIO, R. L., GUARINO, D. U., & MALEY, F. (1979a). The primary structure of *Lactobacillus casei* thymidylate synthetase I. The isolation of cyanogen bromide peptides 1 through 5 and complete amino acid sequence of CNBr 1, 2, 3 and 5. *Journal of Biological Chemistry* **254**, 1288-1295.

MALEY, G. F., BELLISARIO, R. L., GUARINO, D. U., & MALEY, F. (1979b). The primary structure of *Lactobacillus casei* thymidylate synthetase III. The use of 2-(2-nitrophenylsulfenyl)-3-methyl-3-bromoindolenine and limited tryptic peptides to estasblish the complete amino acid sequence of the enzyme. *Journal of Biological Chemistry* **254**, 1301-1304.

MALEY, G. F., MALEY, F. & BAUGH, C. M. (1979c). Differential inhibition of host and viral thymidylate synthase by folyl-polyglutamates. *Journal of Biological Chemistry* **254**, 7485-7487.

MANAVALAN, P., MITTELSTAEDT, D. M., SCHIMERLIK, M. I. & JOHNSON, W. C. (1986). Conformational analysis of thymidylate synthase from amino acid sequence by circular dichroism. *Biochemistry* **25**, 6650-6655.

MANIATIS, T., FRITSCH, E. F. & SAMBROOK, J. (1982). *Molecular Cloning. A Laboratory Manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

MARSHAM, P. R., JACKMAN, A. L., OLDFIELD, J., HUGHES, L. R., THORNTON, T. J., BISSET, G. M. F., O'CONNOR, B. M., BISHOP, J. A. M. & CALVERT, A. H. (1990). Quinazoline antifolate thymidylate synthase inhibitors: benzoyl ring modifications in the C2-methyl series. *Journal of Medicinal Chemistry* **33**, 3067-3071.

MARQUEZ, V. E., Mu-IllLIM, TREANOR, S. P., PLOWMAN, J., PRIEST, M. A., MARKOVAC, A., KHAN, M. S., KASKAR, B. & DRISCOLL, J. S. (1990). Cyclopentenyl-cytosine. A carbocyclic nucleoside with antitumour and antiviral properties. *Journal of Medicinal Chemistry* **31**, 1687-1694.

MARTIN, S. L., APARISIO, D. I. & BANDYOPADHYAY, P. K. (1989). Genetic and biochemical characterization of the thymidine kinase gene from herpesvirus of turkeys. *Journal of Virology* **63**, 2847-2852

MATHERLY, L. H., SCHUETZ, J. D., WESTIN, E. & GOLDMAN, I. D. (1989). A method for the synchronization of cultured cells with amphidicolin: application to the large-scale synchronization of L1210 cells and the study of the cell cycle regulation of thymidylate synthase and dihydrofolate reductase. *Analytical Biochemistry* **182**, 338-345.

MATHEWS, C. K., MOEN, L. K., WANG, Y. & SARGENT, R. G. (1988). Intracellular organization of precursor biosynthetic enzymes. *Trends in Biochemical Sciences* **13**, 394-397.

MATHEWS, C. K., NORTH, T. W. & REDDY, G. P. V. (1978). Multienzyme complexes in DNA precursor biosynthesis. *Advances in Enzyme Regulation* **17**, 133-156.

MATSUO, Y. (1957). Formation of schiff bases of pyridoxal phosphate. Reactions with metal ions. *Journal of the American Chemical Society* **79**, 2011-2015.

MATTHEWS, D. A., APPELT, K. & OATLEY, S. J. (1989a). Stacked beta-bulges in thymidylate synthase account for a novel right-handed rotation between opposing beta-sheets. *Journal of Molecular Biology* **205**, 449-454.

MATTHEWS, J. T., CARROLL, R. D., STEVENS, J. T. & HAFFEY, M. L. (1989b). *In vitro* mutagenesis of the herpes simplex virus type 1 DNA polymerase gene results in altered drug sensitivity of the enzyme. *Journal of Virology* **63**, 4913-4918.

MATTHEWS, D. A., APPELT, K., OATLEY, S. J. & XUONG, N. H. (1990a). Crystal structure of *Eschericia coli* thymidylate synthase containing bound 5-fluoro-2'-deoxyuridylate and 10-propararargyl-5,8-dideazafolate. *Journal of Molecular Biology* **214**, 923-936.

MATTHEWS, D. A., VILLAFRANC, J. E., JANSON, C. A., SMITH, W. W., WELSH, K. & FREER, S. (1990b). Stereochemical mechanism of action for thymidylate synthase based on the X-ray structure of the covalent inhibitory ternary comlex with 5-fluoro-2'-deoxyuridylate and 5,10-methylene tetrahydrofolate. *Journal of Molecular Biology* **214**, 937-948.

MATTHEWS, R. E. F., (1982). Classification and nomenclature of viruses. *Intervirology* **17**, 1-199.

MAZZARELLA, R. A., SRINIVASAN, M., HAIGEJORDEN, S. M. & GREEN, M. (1989). ERp72, an abundant luminal endoplasmic reticulum protein, contains three copies of the active site sequences of protein disulfide isomerase. *Journal of Biological Chemistry* **265**, 1094-1101.

MEAD, D. A., SZCZESNA-SKORUPA, E. & KEMPER, B. (1986). Single-stranded DNA 'blue' T7 promoter plasmids: a versatile tandem promoter system for cloning and protein engineering. *Protein Engineering* **1**, 67-74.

MEEK, T. D., GARVEY, E. P. & SANTI, D. V. (1985). Purification and characterization of the bifunctional thymidylate synthase-dihydrofolate reductase from methotrexate-resistant *Leishmania tropica*. *Biochemistry* **24**, 678-686.

MEISSNER, P. S., SISK, W. P. & BERMAN, M. L. (1987). Bacteriophage lambda cloning system for the construction of directional cDNA libraries. *Proceedings of the National Academy of Sciences* **84**, 4171-4175.

MESSING, J. (1983). New M13 vectors for cloning. *Methods in Enzymology* **101**, 20-78.

MICHAELS, M. L., MATTHEWS, D. A. & MILLER, J. H. (1990). *Escherichia coli* thymidylate synthase: amino acid substitutions by suppression of amber nonsense mutations. *Proceedings of the National Academy of Sciences, U.S.A.* **87**, 3957-3961.

MILLER, J. H. (1979). Genetic studies of the *lac* repressor XI. *Journal of Molecular Biology* **131**, 249-258.

MILLER, J. H. & SCHMEISSER, U. (1979). Genetic studies of the *lac* repressor X. *Journal of Molecular Biology* **131**, 223-248.

MILLER, J. H., COULONDRE, C., HOFER, M., SCHMEISSER, U., SOMMER, H., SCHMITZ, A. & LU, P. (1979). Genetic studies of the *lac* repressor IX. *Journal of Molecular Biology* **131**, 191-222.

MITTAL, S. K. & FIELD, H. J. (1989). Analysis of the bovine herpesvirus type 1 thymidine kinase (TK) gene from wild-type virus and TK-deficient mutants. *Journal of General Virology* **70**, 901-908.

MOCARSKI, E. S. & ROIZMAN, B. (1982a). Herpesvirus dependent amplification and inversion of cell-associated viral thymidine kinase gene flanked by viral sequences and linked to an origin of viral DNA replication. *Proceedings of the Nationalk Academy of Sciences, U.S.A.* **79**, 5626-5630.

MOCARSKI, E. S. & ROIZMAN, B. (1982b). Structure and role of herpes simplex virus DNA terminii in inversion, circularization and generation of virion DNA. *Cell* **22**, 243-255.

MOEN, L. K., HOWELL, M. L., LASSER, G. W. & MATHEWS, C. K. (1988). *Journal of Molecular Recognition* **1**, 48-57.

MONTFORT, W. R., PERRY, K. M., FAUMAN, E. B., FINER-MOORE, J. S., MALEY, G. F., HARDY, L., MALEY, F. & STROUD, R. M. (1990). Structure, multiple site binding, and segmental accomodation in thymidylate synthase on binding dUMP and an anti-folate. *Biochemistry* **29**, 6964-6977.

MOORE, M. A., AHMED, F. & DUNLAP, R. B. (1986). Evidence for the existance of covalent nucleotide-thymidylate synthase complexes, identification of sites of attachment, and enhancement by folates. *Journal of Biological Chemistry* **261**, 12745-12749.

MORAN, R. G., SPEARS, C. P. & HEIDELBERGER, C. (1979). Biochemical determinants of tumor sensitivity to 5-fluorouracil: ultrasensitive methods for the determination of 5-fluoro-2'-deoxyuridylate, 2'-deoxyuridylate, and thymidylate synthetase. *Proceedings of the National Academy of Sciences, U.S.A.* **76**, 1456-1460.

MORGAN, C., ROSE, H. M. & MENDIS, B. (1968). Electron microscopy of herpes simplex virus. I. Entry. *Journal of Virology* **2**, 507-516.

MORRISON, P. F. & ALLEGRA, C. J. (1989). Folate cycle kinetics in human breast cancer cells. *Journal of Biological Chemistry* **264**, 10552-10566.

MULKINS, M. A. & HEIDELBERGER, C. (1982). Biochemical characterization of fluoropyrimidine-resistant murine leukemic cell lines. *Cancer Research* **42**, 965-973.

MUNCH-PETERSON, A. & MYGIND, B. (1983). Transport of nucleic acid precursors, p. 259-305. In A. Munch-Peterson (Editor), *Metabolism of nucleotides, nucleosides and nucleobases in microorganisms*. Academic Press, Inc., London.

MYERS, C. E., YOUNG, R. C., JOHNS, D. G. & CHABNER, B. A. (1974). Assay of 5-fluorodeoxyuridylate and deoxyuridylate pools following 5-fluorouracil. *Cancer Research* **34**, 2682-2688.

NAGASHIMA, K., NAKAZAWA, M. & ENDO, H. (1975). Pathology of the human spinal ganglia in varicella-zoster virus infection. *Acta Neuropathology (Berlin)* **33**, 105-117.

NAGPAL, S. & OSTROVE, J. M. (1991). Characterization of a potent varicella-zoster virus encoded trans-repressor. *Journal of Virology* **65**, 5289-5296.

NAONA, S. & GROS, F. (1960a). Effets d'un analogue de base nucleique sur la biosynthese de proteines bacteriennes. *Comptes rendus de l'Academie des sciences* **250**, 3527-3529.

NAONA, S. & GROS, F. (1960b). Synthese par *E.coli* d'une phosphate modifee en presence d'un analogue pyrimidine. *Comptes rendus de l'Academie des sciences* **250**, 3889-3891.

NAVALGUND, L. G., ROSSANA, C., MUENCH, A. J. & JOHNSON, L. F. (1980). Cell cycle regulation of thymidylate synthetase gene expression in cultured mouse fibroblasts. *Journal of Biological Chemistry* **255**, 7386-7390.

NAZERIAN, K. (1974). DNA configuration in the core of Marek's disease virus. *Journal of Virology* **13**, 1148-1150.

NEIPEL, F., ELLINGER, K. & FLECKENSTEIN, B. (1991). The unique region of the human herpesvirus 6 genome is essentially colinear with the $U_L$ segment of human cytomegalovirus. *Journal of General Virology* **72**, 2293-2297.

NELSON, K., ALONSO, G., LANGER, P. J. & BEVERLEY, S. M. (1990). Sequence of the dihydrofolate reductase-thymidylate synthase (DHFR-TS) gene of *Leishmania amazonensis*. *Nucleic Acids Research* **18**, 2819.

NER, S. S., GOODIN, D. B. & SMITH, M. (1988). A simple and efficient procedure for generating random point mutations and for codon replacements using mixed oligodeoxynucleotides. *DNA* **7**, 127-134.

NEUHARD, J. & NYGAARD, P. (1987). Purines and pyrimidines. In F. C. NEIDHARDT (Editor). Esherichia coli and Salmonella typhimurium - Cellular and Molecular Biology. American Society for Microbiology, Washington, D.C.

NEUHARD, J., PRICE, A. R., SCHACK, L. & THOMASSEN, E. (1978). Two thymidylate synthases in *Bacillus subtilis*. *Proceedings of the National Academy of Sciences, U.S.A.* **75**, 1194-1198.

NEWLON, C. S. & FANGMAN, W. L. (1975). Mitochondrial DNA synthesis in cell cycle mutants of *Saccharomyces cerevisiae*. *Cell* **5**, 423-428.

NICHOLSON, L., CULLINANE, A.A. & ONIONS, D. E. (1990). The nucleotide sequence of the equine herpesvirus 4 thymidine kinase gene. *Journal of General Virology* **71**, 1801-1805.

NIELSEN, E. & CELLA, R. (1988). Thymidylate synthase in plant cells: kinetic and molecular properties of the enzyme from *Daucus carota* L. cell cultures. *Plant and Cell Physiology* **29**, 503-508.

NII, S & YASUDA, I. (1975). Detection of viral cores having toroid structures in eight herpesviruses. *Biken Journal* **18**, 41-46.

O'DONOVAN, G. A. & NEUHARD, J. (1970). Pyrimidine metabolism in microorganisms. *Bacteriological Reviews* **34**, 278-343.

O'HARE, P. & HAYWARD, G. S. (1985). Three trans-activating regulatory proteins of herpes simplex virus modulate immediate-early gene expression in a pathway involving positive and negative feedback regulation. *Journal of Virology* **56**, 723-733.

OHAYAMA, K. (1976). A basis for bromodeoxyuridine resistance in plants cells. *Env. Exp. Botany* **16**, 209-216.

OSTROVE, J. M., REINHOLD, W., FAN, C.-M., ZORN, S., HAY, J. & STRAUSS, S. E. (1985). Transcription mapping of the varicella-zoster virus genome. *Journal of Virology* **56**, 600-606.

OXMAN, M. N. (1981). Varicella and herpes zoster. In *Medical Microbiology and Infectious Diseases*, pp 1652-1671. Edited by A. I. Braude, C. E. Davis & J. Fierer. Philadelphia: W. B. Saunders.

PARK, J. S., CHENG, C. T.-C. & MERTES, M. P. (1979). 5-[(4-Methyl-1,2,3,4-tetrahydroquinoxalyl)methyl]-2'-deoxyuridine 5'-phosphate: an analogue of a proposed intermediate in thymidylate synthetase catalysis. *Journal of Medicinal Chemistry* **22**, 1134-1137.

PARYANI, S. G. & ARVIN, A. M. (1986). Intrauterine infection with varicella-zoster virus after maternal varicella. *New England Journal of Medicine* **314**, 1542-1546.

PATERSON, B. M., ROBERTS, B.E. & KUFF, E. L. (1977). Structural gene identification and mapping by DNA.mRNA hybrid-arrested cell-free translation. *Proceedings of the National Academy of Sciences, U.S.A.* **74**, 4370-4374.

PAULING, L. (1944). *The Nature of the Chemical Bond.* Ithaca, Cornell University Press.

PEDERSDEN, M., TALLEY-BROWN, S. & MILLETTE, R. L. (1981). Gene expression of herpes simplex virus. III. Effect of arabinosyladenine on viral polypeptide synthesis. *Journal of Virology* **38**, 712-719.

PELLINO, A. M. & DANENBERG, P. V. (1985). Evidence from chemical degradation studies for a covalent bond from 5-fluoro-2'-deoxyuridylate to N5 of tetrahydrofolate in the ternary complex of thymidylate synthetase-5-fluoro-2'-deoxyuridylate-5,10-methylenetetrahydrofolate. *Journal of Biological Chemistry* **260**, 10996-11000.

PERRY, K. M., FAUMAN, E. B., FINER-MOORE, J. S., MONTFORT, W. R., MALEY, G. F., MALEY, F. & STROUD, R. M. (1990). Plastic adaptation toward mutations in proteins: structural comparison of thymidylate synthase. *Proteins: Structure, Function and Genetics* **8**, 315-333.

PERRY, L. J., RIXON, F. J., EVERETT, R. D., FRAME, M. C. & McGEOCH, D. J. (1986). Characterization of the IE110 gene of herpes simplex virus type 1, strain MP. *Journal of General Virology* **67**, 2365-2380.

PERRYMAN, S. M., ROSSANA, C., DENG, T., VANIN, E. F. & JOHNSON, L. F. (1986). Sequence of a cDNA for mouse thymidylate synthase reveals striking similarity with the prokaryotic enzyme. *Molecular Biology and Evolution* **3**, 313-321.

PERUTZ, M. F., KENDREW, J. C. & WATSON, H. C. (1965). Structure and function of haemoglobin. *Journal of Molecular Biology* **13**, 669-678.

PETERSON, D. S., MILHOUS, W. K. & WELLEMS, T. E. (1990). Molecular basis of differential resistance to cycloguanil and pyrimethamine in *Plasmodium falciparum* malaria. *Proceedings of the National Academy of Sciences, U.S.A.* **87**, 3018-3022.

PIGGOT, P. J. & HOCH, J. A. (1985). Revised linkage map of *Bacillus subtilis*. *Microbiological Review* **49**, 158-179.

PINE, R. & HUANG, P. C. (1987). An improved method to obtain a large number of mutants in adefined region of DNA. *Methods in Enzymology* **154**, 415-430.

PINTER, K., DAVISSON, V. J. & SANTI, D.V. (1988). Cloning, sequencing, and expression of the *Lactobacillus casei* thymidylate synthase gene. *DNA* **7**, 235-241.

PLESE, P. C. & DUNLAP, R. B. (1977). Sulphydryl group modifications of thymidylate synthetase and its effects on activity and ternary complex formation. *Journal of Biological Chemistry* **252**, 6139-6144.

PODGORSKI, G. & DEERING, R. A. (1984). Thymidine-requiring mutants of *Dictyostelium discoideum*. *Molecular and Cellular Biology* **4**, 2784-2791.

POGOLOTTI, & SANTI, D. V. (1977). *Bioorganic Chemistry* **I**, 277-311.

POGOLOTTI, A. L. Jr., DANENBERG, P. V. & SANTI, D. V. (1986). Kinetics and mechanism of interaction of 10-propargyl-5,8-dideazafolate with thymidylate synthase. *Journal of Medicinal Chemistry* **29**, 478-482.

POST, L. E. & ROIZMAN, B. (1981). A generalized technique for deletion of specific genes in large genomes: gene 22 of herpes simplex virus 1 is not essential for growth. *Cell* **25**, 227-232.

PREBLUD, S. R., BREGMAN, D. J. & VERNON, L. L. (1985). Deaths from varicella in infants. *Pediatrics and Infectious Diseases Journal* **4**, 503-507

PREBLUD, S. R., ARBETER, A. M., PROCTOR, E. A., STARR, S. E. & PLOTKIN, S. A. (1984a). Susceptibility of vaccine strains of varicella-zoster virus to antiviral compounds. *Antimicrobial Agents and Chemotherapy* **25**, 417-421.

PREBLUD, S. R., ORENSTEIN, W. A. & BART, K. J. (1984b). Varicella: clinical manifestations, epidemiology, and health impact in children. *Pediatrics and Infectious Diseases* **3**, 505-509.

PRESTON, C. M. (1979a). Control of herpes simplex virus type 1 mRNA synthesis in cells infected with wild type virus or the temperature sensitive mutant *ts*K. *Journal of Virology* **29**, 275-285.

PRESTON, C. M. (1979b). Abnormal properties of an immediate early polypeptide in cells infected with the herpes simplex virus type 1 mutant *tsK*. *Journal of Virology* **32**, 357-369.

PUGA, A. & NOTKINS, A. L. (1987). Continued expression of a poly(A)+ transcript of herpes simplex virus type 1 in trigeminal ganglia of latently infected mice. *Journal of Virology* **61**, 1700-1703.

PUROHIT, S. & MATHEWS, C. K. (1984). Nucleotide sequence reveals overlap between T4 genes encoding dihydrofolate reductase and thymidylate synthase. *Journal of Biological Chemistry* **259**, 6261-6266.

RADPARVAR, S., HOUGHTON. P. J., GERMAIN. G., PENNINGTON, J. RAHMAN, A. & HOUGHTON, J. A. (1989a). Cellular pharmacology of 5-fluorouracil in a human colon adenocarcinoma cell line selected for thymidine kinase deficiency. *Biochemical Pharmacology* **39**, 1759-1765.

RADPARVAR, S., HOUGHTON. P. J. & HOUGHTON, J. A. (1989b). Effect of polyglutamation of 5,10-methylenetetrahydrofolate on the binding of 5-fluoro-2'-deoxyuridylate to thymidylate synthase purified from a human colon adenocarcinoma xenograft. *Biochemical Pharmacology* **38**, 335-342.

RAU, D. C., LEE, B. K. & PARSEGIAN, V. A. (1984). Measurement of the repulsive forces between polyelectrolyte molecules in solution: hydration forces between parallel DNA double helices. *Proceedings of the National Academy of Sciences, U.S.A.* **81**, 2621-2625.

REICHARD, P., SKOLD, O. & KLEIN, G. (1959). Possible enzymic mechanism for the development of resistance against fluorouracil in ascites tumours. *Nature* **183**, 939-941.

REINHOLD, W. C., STRAUSS, S. E. & OSTROVE, J. M. (1988). Directionality and further mapping of varicella zoster virus transcripts. *Virus Research* **9**, 249-261.

REMINGTON, P. L., ROWLEY, D., McGEE, H., HALL, W. N. & MONTO, A. S. (1986). Decreasing trends in Reye syndrome and aspirin use in Michigan, 1979-1984. *Pediatrics* **77**, 93-98.

REMINGTON, S., WIEGAND, G. & HUBER, R. (1982). Crystallographic refinement and atomic refinement of two different forms of citrate synthase at 2.7 and 1.7 Angstroms resolution. *Journal of Molecular Biology* **158**, 111-152.

RENNELL, D., BOUVIER, S. E., HARDY, L. W. & POTEETE, A. R. (1991). Systematic mutation of bacteriophage T4 lysozyme. *Journal of Molecular Biology* **222**, 67-87.

REYES, P. & HEIDELBERGER, C. (1965). Fluorinated pyrimidines. *Molecular Pharmacology* **1**, 14-30.

RICE, A. R. & KNIPE, D. M. (1988). Gene-specific transactivation by herpes simplex virus type 1 alpha protein ICP27. *Journal of Virology* **62**, 3814-3823.

RICHARDSON, J. S. (1981). The anatomy and taxonomy of protein structure. *Advances in Protein Chemistry* **34**, 167-339.

RICHARDSON, J. S., GETZOFF, E. & RICHARDSON, D. (1981). The ß-bulge: a common small unit of nonrepetitive protein structure. *Proceedings of the National Academy of Sciences* **75**, 2574-2578.

RICHTER, J., PUCHTLER, I. & FLECKENSTEIN, B. (1988). Thymidylate synthase gene of herpesvirus ateles. *Journal of Virology* **62**, 3530-3535.

RICHTER, K. & DUBOCHET, J. (1990). High resolution study of DNA in vitrified sections. *Proceedings of the XIIth International Congress on Electron Microscopy.*, vol. 1, pp 488-489. San Francisco: San Francisco Press.

ROBERTSON, G. R. & WHALLEY, J. M. (1988). Evolution of the herpesvirus thymidine kinase: identification and comparison of the equine herpesvirus 1 thymidine kinase gene reveals similarity to a cell encoded thymidylate kinase. *Nucleic Acids Research* **16**, 11303-11317.

RODE, W., CIESLA, J., ZIELINSKI, Z. & KEDZIERSKA, B. (1986). Purification and properties of mouse thymus thymidylate synthase. Comparison of the enzyme from mammalian normal and tumour tissues. *International Journal of Biochemistry* **18**, 361-368.

RODE, W., SCANLON, K. J., MOROSON, B. A. & BERNITO, J. R. (1980). Regulation of thymidylate synthetase in mouse leukemic cells (L1210). *Journal of Biological Chemistry* **255**, 1305-1311.

RODE, W., ZIELINSKI, Z., DZIK, J. M., KULIKOWSKI, T., BRETNER, M., KIERDASZUK, B., CIESLA, J. & SHUGAR, S. (1990). Mechanism of inhibition of mammalian tumor and other thymidylate synthases by $N^4$-hydroxy-dCMP, $N^4$-hydroxy-5-fluoro-dCMP, and related analogues. *Biochemistry* **29**, 10835-10842.

ROGERS, H. J. & PERKINS, H. R. (1960). 5-fluorouracil and mucopeptide biosynthesis. *Biochemical Journal* **77**, 448-459.

ROIZMAN, B. (1982). The family Herpesviridae: general description, taxonomy, and classification. In *The Herpesviruses*, vol. 1, pp 1-23. Edited by B. Roizman. New York & London: Plenum Press.

ROIZMAN, B. (1990). Herpesviridae: a brief introduction. In *Virology*, pp 1787-1794. Edited by B. N. Fields & D. M. Knipe. Raven Press: New York.

ROIZMAN, B. & SEARS, A. E. (1990). Herpes simplex viruses and their replication. In *Virology*, pp 1795-1841. Edited by B. N. Fields & D. M. Knipe. Raven Press: New York.

ROIZMAN, B. (1991). In *The Classification and Nomenclature of Viruses*, pp103-110 (*Archeives of Virology* Supplementum 2). Edited by R. I. B. Francki, C. M. Fauquet, D. L. Knudson and F. Brown. Springer-Verlag: Wein.

ROIZMAN, B. & FURLONG, D. (1974). The replication of herpesviruses. In: *Comprehensive Virology*, vol 3, pp229-403. Edited by H. Fraenkel-Conrat & R. R. Wagner. New York: Plenum Press.

ROIZMAN, B., CARMICHAEL, L. E. & DEINHARDT, F., de-THE, G., NAHMIAS, A. J., PLOWRIGHT, W., RAPP, F., SHELDRICK, P., TAKAHASHI, M. & WOLF, K. (1981). Herpesviridae: definition, provisional nomenclature and taxonomy. *Intervirology* **16**, 201-217.

ROSS, P., O'GARA, F. & CONDON, S. (1990a). Cloning and character-ization of the thymidylate synthase gene from *Lactococcus lactis* subsp. *lactis*. *Applied and Environmental Microbiology* **56**, 2156-2163.

ROSS, P., O'GARA, F. & CONDON, S. (1990b). Thymidylate synthase gene from *Lactococcus lactis* as a genetic marker: an alternative to anti-biotic resistance genes. *Applied and Environmental Microbiology* **56**, 2164-2169.

ROSSON, D., OTWELL, H. B. & DUNLAP, R. B. (1980). Essential tyrosyl residues in *L.casei* thymidylate synthetase. *Biochemical and Biophysical Research Communications* **97**, 500-505.

ROUCH, D. A., MESSEROTTI, L. J., LOO, L. S. L., JACKSON, C. A. & SKURRAY, R. A. (1989). Trimethoprim resistance transposon Tn*4003* from *Staphylococcus aureus* encodes genes for a dihydrofolate reductase and thymidylate synthetase flanked by three copies of IS*257*. *Molecular Microbiology* **3**, 161-175.

RUBIN, E. M., WILSON, G. A. & YOUNG, F. E. (1978). Transformation of *Bacillus subtilis* with cloned thymidylate synthase genes, p. 363-376. In S. W. Glover and L. O. Butler (Editors), *Proceedings of the 4th European Meeting on Bacterial Transformation and Transduction*. Cotswold Press, York.

RGER, R., DANIEL, M. D. & FLECKENSTEIN, B. (1980). Herpesvirus aotus type 2, a virus related to the oncogenic herpesviruses of New World primates. *Zentralblatt fr Bakteriologie, Parasitenkunde, Infectionskrankheiten und Hygeine Abt. I Originale Reiche A* **246**, 441.

RUSSEL, M., KIDD, S. & KELLEY, M. R. (1986). An imporved filamentous helper phage for generating single-stranded plasmid DNA. *Gene* **45**, 333-338.

RUSSELL, J. & PRESTON, C. M. (1986). An *in vitro* latency system for herpes simplex virus type 2. *Journal of General Virology* **67**, 397-403.

RUSSELL, J., STOW, N. D., STOW, E. C. & PRESTON, C. M. (1987). Herpes simplex virus genes involved in latency *in vitro*. *Journal of General Virology* **68**, 3009-3018.

RYDER, J. W., CROEN, K., KLEINSCHMIDT-DeMASTERS, B. K., OSTROVE, J. M., STRAUSS, S. E. & COHN, D. L. (1986). Progressive encephalitis three months after resolution of cutaneous zoster in a patient with AIDS. *Annals of Neurology* **19**, 182-188.

SACKS, W. R. & SCHAFFER, P. A. (1987). Deletion mutants in the gene encoding the herpes simplex virus type 1 immediate-early protein ICP0 exhibit impared growth in cell culture. *Journal of Virology* **61**, 826-839.

SACKS, W. R., GREENE, C. C., ASCHMAN, D. P. & SCHAFFER, P. A. (1985). Herpes simplex virus type 1 ICP27 is an essential regulatory protein. *Journal of Virology* **55**, 796-805.

SAKUMA, T. (1984). Strains of varicella-zoster virus resistant to 1-**B**-D-arabinofuranosyl-*E*-5-(2-bromovinyl)-uracil. *Antimicrobial Agents and Chemotherapy* **25**, 742-746.

SAKUMA, T., SUZUTANI, T., KAWAI, H. & MASANOBU, A. (1985). Strains of varicella-zoster virus (VZV) resistant to 1-**B**-D-arabinofuranosyl-*E*-5-(2-bromovinyl)uracil (BV-araU). In R. Kono (Editor), *Herpes Viruses and Virus Chemotherapy*, p. 217-220. Elsevier Science Publications, B. V.

SAMBROOK, J., FRITSCH, E. F. & MANIATIS, T. (1989). Molecular Cloning. *Cold Spring Harbor Laboratory Press.*

ST. CLAIR, M. H., MARTIN, J. L., TUDOR-WILLIAMS, G., BACH, M. C., VAVRO, C. L., KING, D. M., KELLAM, P., KEMP, S. D. & LARDER, B. A. (1991). Resistance to ddI and sensitivity to AZT induced by a mutation in HIV-1 reverse transcriptase. *Science* **253**, 1557-1559.

SANDER, C. & SCHNEIDER, R. (1991). Database of homology-derived protein structures and the structural meaning of sequence alignment. *Proteins: Structure, Function, and Genetics* **9**, 56-68.

SANGER, F., NICKLEN, S. & COULSON, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences U.S.A.* **74**, 5463-5467.

SANI, B. P., VAID, A., CORY, J. G., BROCKMAN, R. W., ELLIOTT, R. D. & MONTGOMERY, J. A. (1986). 5'-haloacetamido-5'-deoxythymidines: novel inhibitors of thymidlate synthase. *Biochimica et Biophysica Acta* **881**, 175-184.

SANTI, D. V. & DANENBERG, P. V. (1984). Folates in pyrimidine and nucleotide biosynthesis. In *Folates and Pterins*, vol. 1, pp345-398. Edited by R. L. Blakley & S. J. Benkovic. New York: Wiley.

SANTI, D. V. & HARDY, L. W. (1987). Catalytic mechanism and inhibition of tRNA (uracil-5) methyltransferase: evidence for covalent catalysis. *Biochemistry* **26**, 8599-8606.

SANTI, D. V., McHENRY, C. S. & SOMMER, H. (1974). Mechanism of interaction of thymidylate synthetase with FdUMP. *Biochemistry* **13**, 471-481.

SANTI, D. V., PINTER, K., KEALEY, J. & DAVISSON, V. J. (1990). Site-directed mutagenesis of arginine 179 of thymidylate synthase. *Journal of Biological Chemistry* **265**, 6770-6775.

SARASTE, M., SIBBALD, P. R. & WHITTINGHOFER, A. (1990). The P-loop - a common motif in ATP- and GTP-binding proteins. *Trends in Biochemical Sciences* **15**, 430-434.

SARTORIUS, C. & FRANKLIN, R. M. (1991a). Hybridization arrest of the cell-free translation of the malarial dihydrofolate reductase/thymidylate synthse mRNA by anti-sense oligodeoxyribo-nucleotides. *Nucleic Acids Research* **19**, 1613-1618.

SARTORIUS, C. & FRANKLIN, R. M. (1991b). The use of anti-sense oligonucleotides as chemotherapeutic agents for parasites. *Parasitology Today* **7**, 90-93.

SAWYER, M. H., INCHAUSPE, G., BIRON, K. K., WATERS, D. J., STRAUS, S. E. & OSTROVE, J. M. (1988). Molecular analysis of the pyrimidine deoxyribonucleoside kinase gene of wild-type and acyclovir-resistnat strains of varicella zoster virus. *Journal of General Virology* **69**, 2585-2593.

SCHAEFFER, H. J., BEAUCHAMP, L. deMIRANDA, P., ELION, G. B., BANFR, D. J. & COLLINS, P. (1978). 9-(2-Hydroxy-ethoxymethyl)guanine activity against viruses of the herpes group. *Nature* **272**, 583-585.

SCHIFFER, C. A., DAVISSON, V. J., SANTI, D. V. & STROUD, R. M. (1991). Crystallization of human thymidylate synthase. *Journal of Molecular Biology* **219**, 161-163.

SCHRAG, J. D., VENKATARAM PRASAD, B. V., RIXON, F. J. & CHIU, W. (1989). Three-dimensional structure of the HSV-1 nucleocapsid. *Cell* **56**, 651-660.

SEARS, A. E., HALLIBURTON, I. W., MEIGNIER, B., SILVER, S. & ROIZMAN, B. (1985). Herpes simplex virus 1 mutant deleted in the alpha22 gene: growth and gene expression in permissive and restrictive cells and establishment of latency in mice. *Journal of Virology* **55**, 338-346.

SEITHER, R. L., TRENT, D. F., MIKULECKY, D. C., RAPE, T. J. & GOLDMAN, I. D. (1991). Effect of direct suppresion of thymidylate synthase at the 5,10-methylenetetra-hydrofolate binding site on the interconversion of tetrahydrofolate cofactors to dihydrofolate by anti-folates: influence of degree of dihydrofolate reductase inhibition. *Journal of Biological Chemistry* **266**, 4112-4118.

SEKULOVICH, R. E., LEARY, K. & SANDRI-GOLDIN, R. M. (1988). The herpes simplex virus type 1 alpha protein ICP27 can act as a *trans*-repressor and *trans*-activator in combination with ICP4 and ICP0. *Journal of Virology* **62**, 4510-4522.

SENO, T., AYUSAWA, D., SHIMUZU, K., KOYAMA, H., TAKEISHI, K. & HORI, T. (1985). Thymineless death and genetic events in mammalian cells. In F. J. de Serres (Editor), *Genetic Consequences of Nucleotide Pool Imbalance*. Plenum, New York.

SHEPP, D. H., DANDLIKER, P. S. & MEYERS, J. D. (1986). Treatment of varicella-zoster virus infection in severely immunocompromised patients: a randomized comparison of acyclovir and vidarabine. *New England Journal of Medicine* **314**, 208-212.

SHEPPARD, M. & MAY, J. T. (1989). Location and characterization of the bovine herpesvirus type 2 thymidine kinase gene. *Journal of General Virology* **70**, 3067-3071.

SHIGETA, S., MORI, S., YOKATA, T., KONNO, K. & DeCLERCQ, E. (1986). Characterization of a varicella-zoster virus variant with altered thymidine kinase activity. *Antimicrobial Agents and Chemotherapy* **29**, 1053-1058.

SHIRAISHI, H. & SHIMURA, Y. (1988). A rapid and efficient method for targeted random mutagenesis. *Gene* **64**, 313-319.

SHIRAKI, K., HORIUCHI, K., ASANO, Y., YAMANISHI, K. & TAKAHASHI, M. (1991). Differentiation of Oka varicella vaccine strain from wild varicella-zoster virus strains isolated from vaccinees and household contact. *Journal of Medical Virology* **33**,128-132.

SHIRAKI, K. & HYMAN, R. W. (1987). The immediate early proteins of varicella-zoster virus. *Virology* **156**, 423-426.

SHLOMAI, J. & KORNBERG, A. (1978). dUTPase of *E.coli*. *Journal of Biological Chemistry* **253**, 3305

SINGER, S. C., RICHARDS, C. A., FERONE, R., BENEDICT, D. & RAY, P. (1989). Cloning, purification, and properties of *Candida albicans* thymidylate synthase. *Journal of Bacteriology* **171**, 1372-1378.

SINGH, U. C., WEINER, P. K., CALDWELL, J. W. & KOLLMAN, P. A. (1986). AMBER (UCSF Version 3.0), Department of Pharmaceutical Chemistry, University of California, San Francisco, CA.

SIRAWARAPORN, W., SIRAWARAPORN, R., COWMAN, A. F., YUTHAVONG, Y. & SANTI, D. V. (1990). Heterologous expression of active thymidylate synthase-dihydrofolate reductase from *Plasmodium falciparum*. *Biochemistry* **29**, 10779-10785.

SLEIKER, L. J. & BENKOVIC, S. J. (1984). Synthesis of (6*R*,11*S*)- and (6*R*,11*R*)-5,10-methylene[11-$^{1}$H,$^{2}$H]tetrahydrofolate. Stereochemical paths of serine hydroxymethyl transferase, 5,10-methylenetetrahydrofolate dehyrogenase, and thymidylate synthetase catalysis. *Journal of the American Chemical Society* **106**, 1833-1840.

SMITH, G. A., YOUNG, P. L. & MATTICK, J. S. (1990). The location and nucleotide sequence of the thymidine kinase gene of the bovine herpesvirus type 1.2. *Journal of General Virology* **71**, 2417-2424.

SMITH, G. L., De CARLOS, A. & CHAN, Y. S. (1989). Vaccinia virus encodes a thymidylate kinase gene: sequence and transcriptional mapping. *Nucleic Acids Research* **17**, 7581-7590.

SONDEK, P. & SHORTLE, D. (1990). *Proteins, Structure, Function, and Genetics* **8**.

SPECTOR, T., HARRINGTON, J. A., MORRISON, R. W. Jr., LAMBE, C. U., NELSON, D. J., AVERETT, D. R., BIRON, K. & FURMAN, P. A. (1989). 2-acetylpyridine 5-[(dimethylamino)thiocarbonyl]-thiocarbonohydrazone (A1110U), a potent inactivator of ribonucleotide reductase of herpes simplex viruses and varicella-zoster virus and a potentiator of acyclovir. *Proceedings of the National Academy of Sciences, U.S.A.* **86**, 1051-1055.

SRINIVASAN, A., AMARNATH, V., BROOM, A. D., FOU, F. C. & CHENG, Y.-C. (1984). A potent multisubstrate inhibitor of human TS. *Journal of Medicinal Chemistry* **27**, 1710-1717.

STACEY, K. A. & SIMPSON, E. (1965). Improved method for the isolation of thymine-requiring mutants of *Escherichia coli*. *Journal of Bacteriology* **90**, 554-555.

STANNARD, L. M., FULLER, A. O. & SPEAR, P. G. (1987). Herpes simplex virus glycoproteins associated with different morphological entities projecting from the virion envelope. *Journal of General Virology* **68**, 715-725.

STERN, E. S. (1937). The mechanism of herpes zoster and its relation to chickenpox. *British Journal of Dermatology and Syphilis* **49**, 264-271.

STEVENS, J. G., WAGNER, E. K., DEVI-RAO, G. B., COOK, M. L. & FELDMAN, L. (1987). RNA complimentary to a herpesvirus alpha gene mRNA is prominent in latently infected neurons. *Science* **235**, 1056-1059.

STEVENS, J. G. (1989). Human herpesviruses: a consideration of the latent state. *Microbiological Reviews* **53**, 318-332.

STEVENSON, D., COLEMAN, K. L. & DAVISON, A. J. (1992). Characterization of the varicella-zoster virus gene 61 protein. *Journal of General Virology* **73**, In Press.

STORMS, R. K., ORD, R. W., GREENWOOD, M. T., MIRDAMADI, B., CHU, F. K. & BELFORT, M. (1984). Cell cycle-dependent expression of thymidylate synthase in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology* **4**, 2858-2864.

STOW, N. D. (1982). Localization of an origin of DNA replication within the $TR_S/IR_S$ repeated region of the herpes simplex virus type 1 genome. *EMBO Journal* **1**, 863-867.

STOW, N. D. & DAVISON, A. J. (1986). Identification of a varicella-zoster virus origin of DNA replication and its activation by herpes simplex virus type 1 gene products. *Journal of General Virology* **67**, 1613-1623.

STOW, N. D. & McMONAGLE, E. C. (1983). Characterisation of the $TR_S/IR_S$ origin of DNA replication of herpes simplex type 1. *Virology* **130**, 427-438.

STOW, N. D. & STOW, E. C. (1986). Isolation and characterization of a herpes simplex virus type 1 mutant containing a deletion within the gene encoding the immediate early polypeptide $V_{MW}110$. *Journal of General Virology* **67**, 2571-2585.

STOW, N. D., WEIR, H. M. & STOW, E. C. (1990). Analysis of the binding sites for the varicella-zoster virus gene 51 product within the viral origin of DNA replication. *Virology* **177**, 570-577.

STRAUS, S. E. (1987). The management of varicella and zoster infections. *Infectious Diseases Clinics North America* **1**, 367-382.

STRAUS, S. E. (1989). Clinical and biological differences between recurrent herpes simplex virus and varicella-zoster virus infections. *Journal of the American Medical Association* **262**, 3455-3458.

STRAUS, S. E., AULAKH, H. S., RUYECHAN, W. T., HAY, J., CASEY, T. A., VANDE WOUDE, G. F., OWENS, J. & SMITH, H. A. (1981). Structure of varicella-zoster virus DNA. *Journal of Virology* **40**, 516-525.

STRAUS, S. E., HAY, J., SMITH, H. & OWENS, J. (1983). Genome differences among varicella-zoster virus isolates. *Journal of General Virology* **64**, 1031-1041.

STRAUS, S. E., OSTROVE, J. M., INCHAUSPE, G., *et al.*, (1988). Varicella-zoster virus infections: biology, natural history, treatment, and prevention. *Annals of Internal Medicine* **108**, 221-237.

STRAUS, S. E., REINHOLD, W., SMITH, H. A., *et al.*, (1984). Endonuclease analysis of viral DNA from varicella and subsequent zoster infections in the same patient. *New England Journal of Medicine* **311**, 1362-1364.

SU, J.-Y. & SCLAFANI, R. A. (1991). Molecular cloning and expression of the human deoxythymidylate kinase gene in yeast. *Nucleic Acids Research* **19**, 832-837.

SUCK, D., SAENGER, W. & ZECHMEISTER, K. (1972). Molecular and crystal structure of the tRNA minor constituent dihydrouridine. *Acta Crystallography section B* **28**, 596-605.

SUOLINNA, E., SLAVIK, K. & HAKALA, M. T. (1967). Content of thymidylate synthetase and inhibition by FUdR in sarcoma 180 cells. *Federal Proceedings* **26**, 813.

TAKAHASHI, M., OKUNO, Y., OTSUKA, T., *et al.* (1975). Development of a live attenuated varicella vaccine. *Biken Journal* **18**, 25-33.

TAKAHASHI, M., OTSUKA, T., OKUNA, T., ASANO, Y., YAZAKI, T. & ISOMURA, S. (1974). Live vaccine used to prevent the spread of varicella in children in hospital. *Lancet* **2**, 1288-1290.

TAKEISHI, K., KANEDA, S., AYUSAWA, D., SHIMIZU, K., GOTOH, O. & SENO, T. (1985). Nucleotide sequence of a functional cDNA for thymidylate synthase. *Nucleic Acids Research* **13**, 2035-2043.

TANESE, N. & GOFF, S. P. (1988). Domain structure of the Moloney murine leukemia virus reverse transcriptase: mutational analysis and separate expression of the DNA polymerase and RNase H activities. *Proceedings of the National Academy of Sciences, U.S.A.* **85**, 1777-1781.

TAYLOR, G. R., LAGOSKY, P. A., STORMS, R. K. & HAYNES, R. H. (1987). Molecular characterization of the cell cycle-regulated thymidylate synthase gene of *Sacchasromyces cerevisiae*. *The Journal of Biological Chemistry* **262**, 5298-5307.

TAYLOR, W. R. (1986). The classification of amino acid conservation. *Journal of Theoretical Biology* **119**, 205-218.

TAYLOR-ROBINSON, D. (1959). Chickenpox and herpes zoster. III. Tissue culture studies. *British Journal of Experimental Pathology* **40**, 521-532.

THANASSI, J. W., MEISLER, N. T. & KITTLER, J. M. (1985). Vitamin B-6 metabolism and cancer. *Current Topics in Nutrition and Disease*. Vol. 13, pp319-336. Edited by R. D. Reynolds and J. E. Leklem. Liss, New York.

THOMPSON, R., HONESS, R. W., TAYLOR, L., MORRAN, J. & DAVISON, A. J. (1987). Varicella-zoster virus specifies a thymidylate synthase. *Journal of General Virology* **68**, 1449-1455.

THYLEN, C. & MATHEWS, C. K. (1989). Essential role of T4 phage deoxycytidylate hydroxymethylase in a multienzyme complex for deoxyribonucleotide synthesis. *Journal of Biological Chemistry* **264**, 15169-15172.

TOLLESON, W. H., CISNEROS, R. J., ALIBHAI, M. & DUNLAP, R. B. (1990). Comparison of enzyme immunoassay and active site titration methods for the determination of thymidylate synthase. *FASEB Journal* **4**, A2301.

TOMASZ, A. & BOREK, E. (1959). An early phase in bacteriocidal action of 5-fluorouracil on *E.coli* K12: osmotic imbalance. *Proceedings of the National Academy of Sciences, U.S.A.* **45**, 929-932.

TOMASZ, A. & BOREK, E. (1962). The mechanism of an osmotic instability induced in *E.coli* K-12 by 5-fluorouracil. *Biochemistry* **1**, 543-552.

TOTH, I., LAZAR, G. & GOODMAN, H. M. (1987). Purification and immunochemical characterization of a dihydrofolate reductase-thymidylate synthase enzyme complex from wild-carrot cells. *The EMBO Journal* **6**, 1853-1858.

TRAINER, D. O. & KARSTAD, L. (1963). Experimental pseudorabies in some wild North American mammals. *Zoonoses Research* **2**, 135-150.

UMEDA, M. & HEIDELBERGER, C. (1968). Comparative studies of fluorinated pyrimidines with various cell lines. *Cancer Research* **28**, 2529-2538.

VAFAI, A., MURRAY, R. S., WELLISH, M., DEVLIN, M. & GILDEN, D. H. (1988). Expression of varicella-zoster virus and herpes simplex virus in normal human trigeminal ganglia. *Proceedings of the National Acadamey of Sciences, U.S.A.* **85**, 2362-2366.

VALENTIN-HANSEN, P., *et al.*, (1978). *Molecular and General Genetics* **159**, 191.

VEDA, T., DUTSCHMAN, G. E., BROOM, A. D. & CHENG, Y.-C. (1986). Interactions of human thymidylate synthetase (TS) with bisubstrate analogues. *Proceedings of the American Association of Cancer Research* **27**, 998.

VERES, Z., SZABOLCS, A., SZINAI, I., DNES, G., KAJTAR-PEREDY, M. & TVS, L. (1985). 5-substituted-2,2'-anhydrouridines, potent inhibitors of uridine phosphorylase. *Biochemical Pharmacology* **34**, 1737-1740.

VERES, Z., SZINAI, I., SZABOLCS, A., UJSZASKY, K. & DNES, G. (1987). Biological activity of the potent uridine phosphorylase inhibitor 5-ethyl-2,2'-anhydrouridine. *Drugs and Experimental Clinical Research* **13**, 615-621.

VIEIRA, J. & MESSING, J. (1987). Production of single-stranded plasmid DNA. *Methods in Enzymology* **153**, 3-11.

WAGNER, E. K. (1985). Individual HSV transcripts. Characcterization of specific genes. In B. Roizman (editor), *The Herpesviruses*, p.45-104, vol.3. Plenum Press, New York & London.

WANG, Y. & MATHEWS, C. K. (1989). Analysis of T4 bacteriophage deletion mutants that lack *td* and *frd* genes. *Journal of Virology* **63**, 4736-4743.

WATAYA, Y. & SANTI, D. V. (1977). Inhibition of *Lactobacillus casei* thymidylate synthase by 5-substituted 2'-deoxyuridylates. Preliminary quantítative structure-activity relationship. *Journal of Medicinal Chemistry* **20**, 1469-1473.

WATAYA, Y, MATSUDA, A. & SANTI, D. V. (1980). Interaction of thymidylate synthetase with 5-nitro-2'-deoxyuridylate. *Journal of Biological Chemistry* **255**, 5538-5544.

WATSON, R. J. & CLEMENTS, J. B. (1980). Identification of a herpes simplex virus type 1 function continuously required for synthesis of early and late virus mRNAs. *Nature* **285**, 329-330.

WATSON, P. N. & EVANS, R. J. (1986). Postherpetic neuralgia: a review. *Archives of Neurology* **43**, 836-840.

WEIBEL, R. E., NEFF, B. J., KUTER, B. J., *et al.*, (1984). Live attenuated varicella virus vaccine: efficacy trial in healthy children. *New England Journal of Medicine* **148**, 630-638.

WEIGLE, K. A. & GROSE, C. (1984). Molecular dissection of the humoural immune response to individual varicella-zoster viral proteins during chickenpox, quiescence, reinfection, and reactivation. *Journal of Infectious Diseases* **149**, 741-749.

WEINER, S. J, KOLLMAN, P. A., CASE, D. A., SINGH, U. C., GHIO, C., ALAGONA, G., PROFETA, S. Jr. & WEINER, P. (1984). A new force field for molecular mechanical simulationof nucleic acids and proteins. *Journal of the American Chemical Society* **106**, 765-784.

WEINER, S. J, KOLLMAN, P. A., NGUYEN, D. T. & CASE, D. A. (1986). An all atom force field for simulates of proteins and nucleic acids. *Journal of Computational Chemistry* **7**, 230-252.

WELLER, T. H. (1953). Serial propagation *in vitro* of agents producing inclusion bodies derived from varicella and herpes zoster. *Proceedings of the Society for Experimental Biology and Medicine* **83**, 340-346.

WELLS, J. A., VASSER, M. & POWERS, D. B. (1985). Cassette mutagenesis: an efficient method for generation of multiple mutations at defined sites. *Gene* **34**, 315-323.

WHITLEY, R., HILTY, M., HAYNES, R., *et al.* and the NIAID Collaborative Antiviral Study Group. (1982a). Vidarabine therapy of varicella in immunosuppressed patients. *Journal of Pediatrics* **101**, 125-131.

WHITLEY, R. J., SOONG, S. J., DOLIN, R., BETTS, R., LINNEMANN, C. Jr., ALFORD, C. A. Jr. & the NIAID Collaorative Antiviral Study Group (1982b). Early vidarabine therapy to control the compications of herpes zoster in immunosuppressed patients. *New England Journal of Medicine* **307**, 971-975.

WHITLEY, R. J. & ALFORD, C. A. (1981). Parenteral antiviral chemotherapy of human herpesviruses. In *The Human Herpesviruses: An Interdisciplinary Perspective.* Edited by A. J. Nahmias, W. R. Dowdle & R. F. Schinazi. New York: Elsevier. 478-490.

WICKNER, R. B. (1975). Mutants of *Saccharomyces cerevisiae* that incorporate deoxythymidine 5'-phosphate into DNA *in vivo.* In *Methods in Cell Biology* vol XI. Edited by D. M. Prescott. Academic Press Incorporated. New York.

WILDY, P., RUSSELL, W. C. & HORNE, R. W. (1960). The morphology of herpes virus. *Virology* **12**, 204-222.

WILLIAMS, D. L., GERSHON, A. A., GELB, L. D., SPRAKER, M. K., STEINBERG, S. & RAGAB, A. H. (1985). Herpes zoster following varicella vaccine in a child with acute lymphocytic leukemia. *Journal of Pediatrics* **106**, 259-261.

WILLIAMSON, A. P. (1975). The varicella-zoster virus in the etiology severe congenital defects. *Clinical Pediatrics* **14**, 533-555.

WILSON, M. C., FARMER, J. L. & ROTHMAN, F. (1966). Thymidylate synthesis and aminopterin resistance in *Bacillus subtilis. Journal of Bacteriology* **92**, 186-196.

WU, C. A., NELSON, N. J., McGEOCH, D. J. & CHALLBERG, M. D. (1988). Identification of herpes simplex virus type 1 genes required for origin-dependent DNA synthesis. *Virology* **62**, 435-443.

WU, J. C. & SANTI, D. V. (1987). Kinetic and catalytic mechanism of *Hha*I methyltransferase. *Journal of Biological Chemistry* **262**, 4778-4786.

YOSHIMURA, A., KUWAZURU, Y., FURKAWA, T., YOSHIDA, H., YAMADA, K. & AKIYAMA, S. (1990). Purification and tissue distribution of human thymidine phosphorylase; high expression in lymphocytes, reticulocytes and tumours. *Biochimica et Biophysica Acta* **1034**, 107-113.

YOSHIOKA, A., TANAKA, K., HIRAOKA, O., KOYAMA, Y., HIROTA, Y., AYUSAWA, D., SENO, T., GARRETT, C. & WATAYA, Y. (1987). Deoxyribonucleoside-triphosphate imbalance: 5-fluorodeoxyuridine-induced DNA double-strand breaks in mouse FM3A cells and the mechanism of cell death. *Journal of Biological Chemistry* **262**, 8235-8241.

ZAIA, J. A., LEVIN, M. J., PREBLUD, S. R., *et al.*, (1983). Evaluation of varicella-zoster immune globulin: protection of immunosuppressed children after household exposure to varicella. *Journal of Infectious Diseases* **147**, 737-743.

ZHANG, H., CISNEROS, R. J., DENG, W., JOHNSON, L. F. & DUNLAP, R. B. (1990). Site-directed mutagenesis of mouse thymidylate synthase: alteration of Arg44 to Val44 in a conserved loop guarding the active site has striking effects on catalysis and nucleotide binding. *Biochemical and Biophysical Research Communications* **167**, 869-875.

ZUSMAN, T., ROSENSHINE, I., BOEHM, G., JAENICKE, R., LESKIW, B. & MEVARECH, M. (1989). Dihydrofolate reductase of the extremely halophilic archaebacterium *Halobacterium volcanii*. *Journal of Biological Chemistry* **264**, 18878-18883.

ZWEERINK, H. J., MORTON, D. H., STANTON, L. W. & NEFF, B. J. (1984). Restriction endonuclease analysis of the DNA from varicella-zoster virus. Stability of the DNA after passage *in vitro*. *Journal of General Virology* **55**, 207-211.

*Journal of General Virology* (1991), **72**, 2583–2586. *Printed in Great Britain*



# Evolution of herpesvirus thymidine kinases from cellular deoxycytidine kinase

**Patrick T. Harrison, Russell Thompson and Andrew J. Davison***

*MRC Virology Unit and Department of Virology, Institute of Virology, University of Glasgow, Church Street, Glasgow G11 5JR, U.K.*

The thymidine kinases encoded by herpesviruses of higher vertebrates form a distinct group and are unrelated to the thymidine kinases (TKs) of other organisms. Their evolutionary source has not been identified, but our analysis has revealed a clear relationship with a sequence of human deoxycytidine kinase (dCK) published recently. We report the sequence of the putative TK of channel catfish virus, a herpesvirus of a lower vertebrate, and show that it is also related to dCK. We propose, therefore, that the TKs of herpesviruses of higher and lower vertebrates have evolved, either independently or successively, from a cellular dCK.

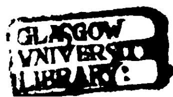
Thymidine kinase (TK; EC 2.7.1.21) catalyses the formation of thymidylate by the phosphorylation of thymidine using ATP as the phosphate donor. Herpesvirus TKs differ from cell cytosolic TKs in both their substrate specificity and the phosphate donors that they can utilize. Herpesvirus and certain mitochondrial TKs have the ability to use CTP as an alternative phosphate donor, but differ in their sensitivity to dCTP inhibition and physical properties (Kit, 1985; Kit *et al.*, 1974).

Unlike cellular TKs, herpesvirus TKs can utilize a broad range of substrates in addition to thymidine, including deoxycytidine, a capability which is essential for the recognition of several antiviral nucleoside analogues, such as acyclovir and (*E*)-5-(2-bromovinyl)-2′-deoxyuridine (BVdU; Fyfe *et al.*, 1978; Cheng *et al.*, 1981; Larder *et al.*, 1983). Moreover, most herpesvirus TKs are also able to phosphorylate thymidylate (Chen *et al.*, 1979), which has a crucial role in the activation of certain nucleoside analogues. For example, herpes simplex virus type 1 (HSV-1) and varicella-zoster virus (VZV), which are sensitive to BVdU, express TKs that phosphorylate this compound to both the mono- and diphosphate forms, whereas HSV-2, which is resistant to BVdU, expresses a TK that can catalyse production of the monophosphate form only (Fyfe, 1981). Most nucleoside analogues are active in the nucleoside triphosphate form, which inhibits virus replication by interacting with DNA polymerase and presumably is formed by the action of a cellular kinase (reviewed in Larder & Darby, 1984).

Poxviruses encode TKs with substrate specificities and physical properties similar to those of cell cytosolic TKs (Kit, 1985; Kit *et al.*, 1974), and these features, as well as clear amino acid sequence similarity, have led to the proposal that poxvirus TKs evolved from cell cytosolic TKs (Kwoh & Engler, 1984; Boyle *et al.*, 1987 and references therein). However, there is no apparent similarity between the sequences of herpesvirus and cell cytosolic TKs, with the exception of a single motif common to many enzymes that bind ATP (Walker *et al.*, 1982; Gentry, 1985), and cellular TK is a tetramer whereas herpesvirus TKs are dimers. To our knowledge, no sequence data are available for mitochondrial TKs. Thus, on the basis of sequence, physical and enzymic properties, herpesvirus TKs form a distinct group, the evolutionary source of which is obscure.

Comparisons between herpesvirus TK sequences have revealed several highly conserved amino acid residues (Robertson & Whalley, 1988; Honess *et al.*, 1989), which have been grouped into six 'sites' (Balasubramaniam *et al.*, 1990). The locations of the sites in the 15 published herpesvirus TK sequences are shown in Fig. 1. The herpesvirus TK sequences reported to date are encoded by viruses that infect mammals or birds, but information has recently become available on the TK of channel catfish virus (CCV). This enzyme has been characterized biochemically and the gene encoding it has been located on the genome to within 3 kbp by marker rescue of a TK deficient CCV mutant (L. A. Hanson, R. L. Thune & K. G. Kousoulas, personal communication). CCV TK can be distinguished biochemically from the channel catfish cellular TK and, in contrast to other herpesvirus TKs, cannot use CTP as a phosphate donor. CCV DNA has been completely sequenced (A. J. Davison, unpub-

0001-0346 © 1991 SGM


GLASGOW UNIVERSITY LIBRARY

```
                          Site 1                 Site 2
                          *******                ******
HSV1  NH2  -50-  rVYIDGphGmGKTtttqlL -13-  EPMtYW -73-
HSV2  NH2  -50-  rVYIDGphGvGKTttsaqL -14-  EPMtYW -73-
VZV   NH2  -13-  rIYLDGayGiGKTtaaeef -15-  EPLsYW -75-
MHV   NH2  -11-  rVYLDGphGvGKSttaeaL -15-  EPMaYW -78-
PRV   NH2  -04-  rIYLDGayGtGKSttarvM -09-  EPMaYW -69-
FHV   NH2  -21-  rIYIDGayGiGKSltakyL -15-  EPMlYW -72-
EHV1  NH2  -26-  rIYLDGvyGiGKSttgrvM -14-  EPMaYW -72-
EHV4  NH2  -26-  rIYLDGvyGiGKSttgrvM -14-  EPMaYW -72-
BHV1  NH2  -11-  rIYLDGahGlGKTttgraL -14-  EPMaYW -82-
BHV2  NH2  -05-  rVYVDGphGlGKTtaasrL -11-  EPMsYW -73-
ILTV  NH2  -30-  lLYVDGpfGvGKTvtaktL -14-  EPMqaW -77-
MDV   NH2  -20-  rVYLDGsmGiGKTsmlneI -14-  EPMkYW -70-
HVT   NH2  -11-  rVYLDGpfGiGKTsilnaM -14-  EPMkYW -70-
EBV   NH2 -285-  sLFLEGapGvGKTtmlnhL -12-  EPMrYW -69-
HVS   NH2 -210-  fIFLEGsiGvGKTtllksM -13-  EPIaYW -66-

CCV   NH2  -17-  vfcVEGniGcGKStlvkaL -13-  EPVdqW -43-

dCK   NH2  -22-  kIsIEGniaaGKStfvniL -11-  EPVarW -74-

      Site 3    Site 4       Site 5            Site 6
      ***       ***          *******           ******
HSV1  DRHPIaalLCYP -42-  RLakRqRpGE -58-  igdTLF -87-  COOH
HSV2  DRHPIaslLCYP -42-  RLarRqRpGE -57-  iedTLF -86-  COOH
VZV   DRHPIastICFP -42-  RVskRaRpGE -60-  iedTLF -83-  COOH
MHV   DRHaVasmVCYP -42-  RLraRaRpGE -63-  ladTLF -99-  COOH
PRV   DRHPVaatVCFP -42-  RLraRaRaGE -61-  lqdTLF -82-  COOH
FHV   DRHPLaslVCFP -42-  RLrgRsRtGE -61-  lcdTLF -79-  COOH
EHV1  DRHPVasaVCFP -42-  RLrtRaRiGE -61-  lsdTLF -84-  COOH
EHV4  DRHPVastVCFP -42-  RLrtRaRiGE -61-  lgdTLF -84-  COOH
BHV1  DRHPVaacLCYP -42-  RLaaRaRpGD -63-  lrdTLF -94-  COOH
BHV2  DRHPtaslLCYP -42-  RLvaRgppGE -55-  ghnTLL -67-  COOH
ILTV  DRHPLaacLCFP -42-  RIiqRgRpGE -63-  idqTLF -85-  COOH
MDV   DRHPIsatVCFP -43-  RLssRnRtGE -67-  lhhTLL -85-  COOH
HVT   DRHPVaaiLCFP -43-  RIcsRdRpGE -68-  ldrTLL -91-  COOH
EBV   DRHlLsasVvFP -41-  RLkkRgRkhE -63-  yknSIF -84-  COOH
HVS   DRHPLsatVvFP -41-  RVkkRnRkeE -63-  lksSIF -81-  COOH

CCV   ERsPMsatrvFc -46-  RMrrRdRtGE -62-  COOH

dCK   DRyifasn.lYe -44-  RIylRgRneE -63-  COOH
```

Fig. 1. Alignment of the amino acid sequences of six conserved sites in 15 higher vertebrate herpesvirus TKs, and comparison with CCV TK and human dCK. MHV, marmoset herpesvirus; PRV, pseudorabies virus; FHV, feline herpes virus; EHV1 and -4, equine herpesvirus types 1 and 4; BHV1 and -2, bovine herpesvirus types 1 and 2; ILTV, infectious laryngotracheitis virus; MDV, Marek's disease virus; HVT, turkey herpesvirus; EBV, Epstein–Barr virus; HVS, herpesvirus saimirii. The TK sequences are from Balasubramaniam *et al.* (1990) with the exception of EHV-4, Nicolson *et al.* (1990); BHV-2, Sheppard & May (1989); ILTV, Griffin & Boursnell (1990). The BHV-1 sequence used by Balasubramaniam *et al.* (1990) is replaced by that of Smith *et al.* (1990). The HVT sequence used in this analysis and in that of Balasubramaniam *et al.* (1990) differs from the original sequence published by Martin *et al.* (1989) by the replacement of the two amino acids at the carboxy terminus with 42 amino acids, resulting from the insertion of a G residue between nucleotides 1532 and 1533 in the original DNA sequence. The considerable similarity between the C-terminal extension and the TK of the closely related MDV suggests that there was an error in the original sequence. The human dCK sequence is from Chottiner *et al.* (1991). The sequences were aligned using programs GAP, LINEUP and PRETTY (Sequence Analysis Software Package, Genetics Computer Group, University of Wisconsin). Upper case letters indicate identical residues in at least 13 of the TK sequences, or a residue in one of the groups (I/L/M/V) (D/E) (S/T) (F/Y) (F/L) (K/R) in at least 14 sequences. Upper case letters are shown for CCV TK and dCK when residues are identical to, or in the same group as, those residues in the other TKs. The six conserved sites proposed by Balasubramaniam *et al.* (1990) are indicated. Dots represent characters introduced to produce the alignment.

```
 M   A   L   R   E   G   P   T   P   A   R   F   L   P   E   G   L   V   F   C     20
ATGGCGCTCAGGGAAGGTCCCACACCCGCCAGGTTTCTCCCCGAAGGGCTAGTGTTCTGT 8844

 V   E   G   N   I   G   C   G   K   S   T   L   V   K   A   L   M   E   R   V     40
GTCGAGGGTAACATAGGATGCGGGAAGAGCACCCTGGTAAAAGCTCTGATGGAGCGCGTG 8904

 A   G   S   G   V   N   V   V   E   E   P   V   D   Q   W   V   N   H   N   G     60
GCGGGGTCCGGGGTCAACGTGGTCGAAGAGCCGGTAGATCAGTGGGTGAATCACAATGGG 8964

 K   N   Y   L   E   L   S   Y   T   D   P   T   G   Y   A   V   P   F   Q   N     80
AAGAACTATCTCGAACTGTCGTACACGGATCCCACGGGGTACGCCGTCCCGTTTCAGAAC 9024

 L   V   F   D   S   Y   V   N   V   Q   R   L   Q   N   P   D   I   M   E   R    100
CTGGTGTTCGACAGTTACGTGAACGTGCAGCGCCTGCAGAATCCCGACATCATGGAGCGC 9084

 S   P   M   S   A   T   R   V   F   C   A   V   N   G   S   R   G   V   I   P    120
TCCCCGATGAGCGCGACCCGGGTCTTTTGCGCGGTCAACGGGAGCCGCGGGGTCATCCCC 9144

 A   T   A   L   P   G   M   A   A   R   G   E   A   V   M   R   T   I   A   T    140
GCGACCGCGCTCCCGGGGATGGCGGCGAGGGGGGAGGCCGTGATGCGAACCATCGCCACC 9204

 R   P   V   F   V   Y   L   E   L   P   P   E   E   C   L   R   R   M   R   R    160
CGTCCCGTCTTCGTGTACCTGGAGCTACCCCCCGAGGAGTGTCTGCGCAGGATGCGTCGT 9264

 R   D   R   T   G   E   A   G   V   G   L   D   Y   L   R   L   L   H   E   R    180
CGGGACAGGACCGGGGAGGCGGGCGTGGGGCTGGACTACCTGCGTCTCCTCCACGAAAGA 9324

 Y   E   A   W   L   S   S   A   E   D   V   E   R   V   D   A   S   R   S   R    200
TACGAGGCGTGGCTATCCTCCGCGGAGGACGTGGAGCGCGTGGACGCGTCCCGCTCCCGA 9384

 E   E   I   V   D   R   V   I   E   I   L   C   R   R   H   P   R   L   R   A    220
GAGGAGATTGTCGATCGCGTGATAGAGATACTGTGCCGCCGCCACCCACGACTCCGCGCC 9444

 P   L   T   R   K   S   Q   L   -                                                228
CCCCTAACACGAAAAAGTCAGTTATAA                                  9471
```

Fig. 2. Nucleotide and predicted amino acid sequences of the putative CCV TK. Nucleotides are numbered with respect to the complete CCV DNA sequence. Residues in the amino acid sequence are also numbered.

lished data) showing that the 3 kbp region including the TK gene contains three complete open reading frames which are predicted to encode proteins, one of which is related to other herpesvirus TKs (Fig. 1). The nucleotide and predicted amino acid sequences of the putative CCV TK are shown in Fig. 2. The herpesvirus TKs are related in four or five of the six conserved sites to the human deoxycytidine kinase (dCK) sequence reported recently (Chottiner *et al.*, 1991) (Fig. 1). Human dCK, like herpesvirus TKs, is a dimeric enzyme with a broad substrate specificity, including purine nucleosides and several cytidine analogues. Both CTP and GTP, as well as ATP, can act as phosphate donors.

Site 1 is the proposed ATP-binding motif of higher vertebrate herpesvirus TKs and has the consensus sequence (D/E)GXXGXGK(T/S) (Gentry, 1985); CCV TK site 1 is identical to this consensus sequence and the dCK sequence differs only by a single substitution (alanine for the second glycine). Replacement of this glycine by valine in HSV-1 TK results in inactivation of the enzyme (Liu & Summers, 1988), although a strict requirement for this amino acid is not observed in other ATP-binding proteins (reviewed in Saraste *et al.*, 1990). Therefore it seems likely that dCK site 1 is part of a functional ATP-binding motif.

Site 2 in the higher vertebrate herpesvirus TKs has the consensus sequence EP(M/L/I)XYW (Balasubramaniam *et al.*, 1990). In ILTV TK, an alanine replaces the tyrosine (Griffin & Boursnell, 1990), and that of CCV possesses a glutamine at this position and a valine at the third position, at which the consensus sequence contains other small hydrophobic amino acids. As small hydrophobic residues are often considered to be similar for the purpose of denoting amino acid conservation (Taylor, 1986; Bordo & Argos, 1991), the consensus sequence

could be modified to EP(M/L/I/V) XXW, which is identical to the site 2 sequence of dCK. The distance between sites 1 and 2 of CCV TK and dCK is similar to that between those of the other herpesvirus TKs.

The region containing sites 3 and 4 has been proposed to have a role in thymidine recognition (Darby *et al.*, 1986). The triplet DRH is conserved at site 3 of all herpesvirus TKs described. Indeed, it was its similarity to a DRY sequence in yeast thymidylate kinase (TmpK) that was used, in part, by Robertson & Whalley (1988) to suggest an evolutionary relationship between this enzyme and herpesvirus TKs. In site 3 of CCV TK, glutamate replaces aspartate and serine replaces histidine; in site 4, the conserved proline is replaced by cysteine. Despite a lower level of similarity between higher vertebrate herpesvirus and CCV TK sites 3 and 4, the CCV enzyme does appear to be related to herpesvirus TKs in this region, most notably to that of HVS. dCK has DRY in place of the DRH motif in site 3 but lacks convincing similarity in site 4. Nevertheless, the similarity between dCK and herpesvirus TKs at sites 3 and 4 is greater than that between yeast TmpK and herpesvirus TKs (Robertson & Whalley, 1988), and that between herpesvirus TKs and the sequences reported recently for vaccinia virus (Smith *et al.*, 1989) and human TmpKs (Su & Sclafani, 1991). We could detect no correlation between the ability of the enzyme to phosphorylate substrates other than thymidine and the amino acid sequence of this region. The distance between sites 2 and 3 of dCK is similar between those of higher vertebrate herpesvirus TKs, but that between those of CCV TK is significantly shorter.

Arginine-rich site 5 is similar to a region of porcine adenylate kinase that is thought to be involved in phosphoryl group binding (Dreusike *et al.*, 1988; Balasubramaniam *et al.*, 1990). Three conserved arginine residues are present in all herpesvirus TKs reported previously with the exception of BHV-2, which has a proline in place of the third arginine, and there is also a well-conserved GE motif; CCV TK contains the arginines and the GE motif. dCK also possesses the three arginines and the glutamate, but the glycine is replaced by a glutamate. The distance between sites 4 and 5 of CCV TK and dCK is marginally greater than between those of the other herpesvirus TKs.

CCV TK and dCK are similar in length and, unlike other herpesvirus TKs, show similarity outside the five conserved sites, particularly in the regions between sites 2 and 3, and surrounding site 5 (Fig. 3). They are significantly shorter than higher vertebrate herpesvirus TKs, and lack site 6 and at least one other residue of functional importance (Darby *et al.*, 1986).

The sequence alignments discussed in this paper demonstrate significant similarity between dCK and herpesvirus TKs. Although it is conceivable that this situation has arisen by convergent evolution, we consider that the data strongly suggest that higher or lower vertebrate herpesvirus TKs have evolved from a captured cellular dCK gene. In developing the ability to phosphorylate thymidine, the enzyme specified by the captured gene in most cases has retained the ability to phosphorylate deoxycytidine. Thus, the proposed origin of herpesvirus TKs is consistent with their distinct biochemical properties.

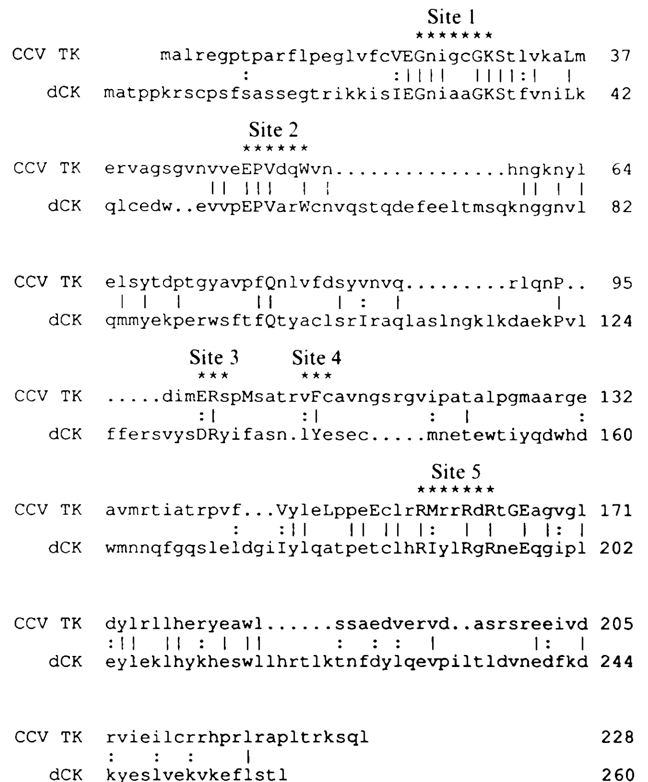



Fig. 3. Alignment of the entire human dCK and CCV TK sequences. Vertical bars indicate identical amino acid residues and colons indicate similar residues. Dots represent characters introduced to produce the alignment. Upper case letters are used as described in the legend to Fig. 1.

P. T. Harrison was the recipient of a C.A.S.E. research studentship from the Science and Engineering Research Council, and the Wellcome Research Laboratories.

## References

Balasubramaniam, N. K., Veerisetty, V. & Gentry, G. A. (1990). Herpesviral deoxythymidine kinases contain a site analogous to the phosphoryl-binding arginine-rich region of porcine adenylate kinase; comparison of secondary structure predictions and conservation. *Journal of General Virology* **71**, 2979–2987.

BORDO, D. & ARGOS, P. (1991). Suggestions for "safe" residue substitutions in site-directed mutagenesis. *Journal of Molecular Biology* **217**, 721–729.

BOYLE, D. B., COUPAR, B. E. H., GIBBS, A. J., SEIGMAN, L. J. & BOTH, G. W. (1987). Fowlpox virus thymidine kinase: nucleotide sequence and relationship to other thymidine kinases. *Virology* **156**, 355–365.

CHEN, M. S., SUMMERS, W. P., WALKER, J., SUMMERS, W. C. & PRUSSOFF, W. H. (1979). Characterization of pyrimidine deoxyribonucleoside kinase (thymidine kinase) and thymidylate kinase as a multifunctional enzyme in cells transformed by herpes simplex type 1 and in cells infected with mutant strains of herpes simplex virus. *Journal of Virology* **30**, 942–945.

CHENG, Y.-C., DUTSCHMAN, G., DECLERCQ, E., JONES, A. S., RAHIM, S. G., VERHELST, G. & WALKER, R. T. (1981). Differential affinities of 5-(2-halogenovinyl)-2′-deoxyuridines for deoxythymidine kinases of various origins. *Molecular Pharmacology* **20**, 230–233.

CHOTTINER, E. G., SHEWACH, D. S., DATTA, N. S., ASHCRAFT, E., GRIBBIN, D., GINSBURG, D., FOX, I. H. & MITCHELL, B. S. (1991). Cloning and expression of human deoxycytidine kinase cDNA. *Proceedings of the National Academy of Sciences, U.S.A.* **88**, 1531–1535.

DARBY, G., LARDER, B. A. & INGLIS, M. M. (1986). Evidence that the 'active centre' of the herpes simplex virus thymidine kinase involves an interaction between three distinct regions of the polypeptide. *Journal of General Virology* **67**, 753–758.

DREUSIKE, D., KARPLUS, P. A. & SCHULZ, G. E. (1988). Refined structure of porcine adenylate kinase at 2·1 Å resolution. *Journal of Molecular Biology* **199**, 359–371.

FYFE, J. A. (1981). Differential phosphorylation of (*E*)-5-(2-bromovinyl)-2′-deoxyuridine monophosphate by thymidine kinases from herpes simplex viruses types 1 and 2 and varicella-zoster virus. *Molecular Pharmacology* **21**, 432–437.

FYFE, J. A., KELLAR, P. M., FURMAN, P. A., MILLER, R. L. & ELION, G. B. (1978). Thymidine kinase from herpes simplex virus phosphorylates the new antiviral compound 9-(2-hydroxyethoxymethyl)guanine. *Journal of Biological Chemistry* **253**, 8721–8727.

GENTRY, G. A. (1985). Locating a nucleotide-binding site in the thymidine kinase of vaccinia virus and of herpes simplex virus by scoring triply aligned protein sequences. *Proceedings of the National Academy of Sciences, U.S.A.* **85**, 6815–6819.

GRIFFIN, A. M. & BOURSNELL, M. E. G. (1990). Analysis of the nucleotide sequence of DNA from the region of the thymidine kinase gene of infectious laryngotracheitis virus; potential evolutionary relationships between the herpesvirus subfamilies. *Journal of General Virology* **71**, 841–850.

HONESS, R. W., CRAXTON, M. A., WILLIAMS, L. & GOMPELS, U. A. (1989). A comparative analysis of the sequence of the thymidine kinase gene of a gammaherpesvirus, herpesvirus saimiri. *Journal of General Virology* **70**, 3003–3013.

KIT, S. (1985). Thymidine kinase. *Microbiological Sciences* **2**, 369–375.

KIT, S., LEUNG, W.-C., JORGENSEN, G. N., TRKULA, D. & DUBBS, D. R. (1974). Thymidine kinase isozymes of normal and virus-infected cells. *Cold Spring Harbor Symposia on Quantitative Biology* **34**, 703–715.

KWOH, T. J. & ENGLER, J. A. (1984). The nucleotide sequence of the chicken thymidine kinase gene and the relationship of its predicted polypeptide to that of the vaccinia virus thymidine kinase. *Nucleic Acids Research* **12**, 3959–3971.

LARDER, B. A. & DARBY, G. (1984). Virus drug-resistance: mechanisms and consequences. *Antiviral Research* **4**, 1–42.

LARDER, B. A., CHENG, Y.-C. & DARBY, G. (1983). Characterization of abnormal thymidine kinases induced by drug-resistant strains of herpes simplex virus type 1. *Journal of General Virology* **64**, 523–532.

LIU, Q. & SUMMERS, W. C. (1988). Site-directed mutagenesis of a nucleotide-binding domain in HSV-1 thymidine kinase: effects on catalytic activity. *Virology* **163**, 638–642.

MARTIN, S. L., APARISIO, D. I. & BANDYOPADHYAY, P. K. (1989). Genetic and biochemical characterization of the thymidine kinase gene from herpesvirus of turkeys. *Journal of Virology* **63**, 2847–2852.

NICOLSON, L., CULLINANE, A. A. & ONIONS, D. E. (1990). The nucleotide sequence of the equine herpesvirus 4 thymidine kinase gene. *Journal of General Virology* **71**, 1801–1805.

ROBERTSON, G. R. & WHALLEY, J. M. (1988). Evolution of the herpes thymidine kinase: identification and comparison of the equine herpesvirus 1 thymidine kinase gene reveals similarity to a cell encoded thymidylate kinase. *Nucleic Acids Research* **16**, 11303–11317.

SARASTE, M., SIBBALD, P. R. & HITTINGHOFER, A. (1990). The P-loop – a common motif in ATP- and GTP-binding proteins. *Trends in Biochemical Sciences* **15**, 430–434.

SHEPPARD, M. & MAY, J. T. (1989). Location and characterization of the bovine herpesvirus type 2 thymidine kinase gene. *Journal of General Virology* **70**, 3067–3071.

SMITH, G. A., YOUNG, P. L. & MATTICK, J. S. (1990). The location and nucleotide sequence of the thymidine kinase gene of bovine herpesvirus type 1.2. *Journal of General Virology* **71**, 2417–2424.

SMITH, G. L., DE CARLOS, A. & CHAN, Y. S. (1989). Vaccinia virus encodes a thymidylate kinase gene: sequence and transcriptional mapping. *Nucleic Acids Research* **17**, 7581–7590.

SU, J.-Y. & SCLAFANI, R. A. (1991). Molecular cloning and expression of the human deoxythymidylate kinase gene in yeast. *Nucleic Acids Research* **19**, 832–837.

TAYLOR, W. R. (1986). The classification of amino acid conservation. *Journal of Theoretical Biology* **119**, 205–218.

WALKER, J. E., SARASTE, M., RUNSWICK, M. J. & GAY, N. J. (1982). Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases, and other ATP-requiring enzymes and a common nucleotide fold. *EMBO Journal* **1**, 945–951.

(*Received 23 April 1991; Accepted 17 June 1991*)

```
                        147      H       159 165 v   173    I    183   iv   192                 195    iii     210
                          |-----------|  |^^^|  |------|  ↘|^^^^^^^^|                          |^^^^^^^^^^^|  |----
VZV             aeykdcqsnylqqgiDQLqtVIdtIKtnPesRRmIissWNpkdip.lMvLPPCHtLcQFYVang.................eLscqVYQRSgDMgLGVPFN
HVS             aeykgvgrdykgegvDQLkqLIdtIKtnPtdRRmLmcaWNvsdip.kMvLPPCHvLsQFYVcdg.................kLscqLYQRSaDMgLGVPFN
HVA             aeyqglkhnyggegvDQLkqIIntIHtnPtdRRmLmcaWNvldvp.kMALPPCHvLsQFYVcdg.................kLscqLYQRSaDMgLGVPFN
Human           aeyrdmesdysgqgvDQLqrVIdtIKtnPddRRiImcaWNprdlp.lMALPPCHaLcQFYVvns.................eLscqLYQRSgDMgLGVPFN
Mouse           aeykdmdsdysgqgvDQLqkVIdtIKtnPddRRiImcaWNpkdlp.lMALPPCHaLcQFYVvng.................eLscqLYQRSgDMgLGVPFN
C.albicans      aeykdcdsdytgqgfDQLqdVIkkLKtnPydRRiImsaWNppdfa.kMALPPCHvFcQFYVnfptslpdpnnpkqaktakpkLsclLYQRScDMgLGVPFN
P.carinii       aeyidcktnyigqgvDQLanIIqkIRtsPydRRlIlsaWNpadle.kMALPPCHmFcQFYVhipsnnh...........rpeLscqLYQRScDMgLGVPFN
S.cerevisiae    akykktcdddytgqgiDQLkqVIhkLKtnPydRRiImsaWNpadfd.kMALPPCHiFsQFYVsfpkegeg.......sgkprLsclLYQRScDMgLGVPFN
C.fasiculata    adykgfdanyd.egvDQIktIVetLKtn..dRRlLvtaWNpcalh.kMAVrPCHlLgQFYVntqtk...............eLscmLYQRccDMgLGVPFN
L.amazonensis   aeyrgleanydgegvDQIkfIVetIKanPndRRlLftaWNpcalh.kMALPPCHlLaQFYVnteks...............eLscmLYQRScDMgLGVPFN
L.major         adykgfeanydgegvDQIklIVetIKtnPndRRlLvtaWNpcalq.kMALPPCHlLaQFYVntdts...............eLscmLYQRScDMgLGVPFN
L.tropica       adykgfeanydgegvDQIklIVetIKtnPndRRlLvtaWNpcalq.kMAVPPCHlLaQFYVntdts...............eLscmLYQRScDMgLGVPFN
P.chabaudi      aeytdmhadykdkgvDQLknIInlIKndPtcRRiIlcaWNvkdld.qMALPPCHiLcQFYVfdg.................kLsciMYQRScDLgLGVPFN
P.falciparum    aeytnmydnyenkgvDQLknIInlIKndPtsRRiLlcaWNvkdld.qMALPPCHiLcQFYVfdg.................kLsciMYQRScDLgLGVPFN
E.coli          tp.......dgrhiDQIttVLnqLKndPdsRRiIvsaWNvgeld.kMALaPCHaFfQFYVadg.................kLscqLYQRScDVfLGLPFN
B.subtilis      ga........dgetiDQIsrLIedIKtnPnsRRlIvsaWNvgeid.kMALPPCHcLfQFYVsdg.................kLscqLYQRSaDVfLGVPFN
L.lactis        i...............IgkLLegLaknPwnRRnIinlWqyedfeetegLlPCafqtmFdVrrekdgqi............yLdatLiQRSnDMlVahhiN
L.casei         ts........kgdtiDQLgdVIeqIKthPysRRlIvsaWNpedvp.tMALPPCHtLyQFYVndg.................kLslqLYQRSaDIfLGVPFN
Tn4003          dk........ngnhyDQLksVIqqIKtnPnsRRhIvsaWNpteid.sMALPPCHtMfQFYVqeg.................kLncqLYQRSaDIfLGVPFN
Phage T4        ............ggvDQIieVIdrIKklPndRRqIvsaWNpaelk.yMALPPCHmFyQFnVrng.................yLdlqwYQRSvDVfLGLPFN
Phage phi3T     rs.......lngekvDQVdyLLhqLKnnPssRRhItmlWNpddld.aMALtPCvyetQWYVkqg.................kLhleVraRSnDMaLGnPFN

                        J          230      237     ii        251   K  258       267   270
                 ---------------|       |^^^^^^^^^|      |------|          |---|
VZV             IAgYaLLTyIVAhvtgLktGDLIhtmGdaHIYlnHIDaLKvQLaRsPkpf.PcLkiirnv....tdIndF........kwdDfqLdgYnp.hppLKmeMAL
HVS             IASYsLLTcMIAhvtnLvpGEFIhtiGdaHIYvdHIDaLKmQLtRtPrpf.PtLrfarnv....scIddF........kadDiiLenYnp.hpiIKmhMAV
HVA             IASYsLLTcMIAhvtdLvpGEFIhtlGdaHVYvnHVDaLteQLtRtPrpf.PtLkfarkv....asIddF........kanDiiLenYnp.ypsIKmpMAV
Human           IASYaLLTyMIAhitgLkpGDFIhtlGdaHIYlnHIEpLKiQLqRePrpf.PkLrilrkv....ekIddF........kaeDfqIegYnp.hptIKmeMAV
Mouse           IASYaLLTyMIAhitgLqpGDFVhtlGdaHIYlnHIEpLKiQLqRePrpf.PkLkilrkv....etIddF........kveDfqIegYnp.hptIKmeMAV
C.albicans      IASYaLLTkMIAhvvdMdcGEFIhtlGdaHVYldHIDaLKeQFeRiPkqf.PkLvikeerkneiksIddF........kfeDfeIvgYep.yppIKmkMsV
P.carinii       IASYaLLTcMIAhvcdLdpGDFIhvmGdcHIYkdHIEaLqqQLtRsPrpf.PtLslnrsi....tdIedF........tldDfnIqnYhp.yetIKmkMsI
S.cerevisiae    IASYaLLTrMIAkvvdMepGEFIhtlGdaHVYkdHIDaLKeQitRnPrpf.PkLkikrdv....kdIddF........kltDfeIedYnp.hprIqmkMsV
C.fasciculata   IASYaLLTiLIAkatgLrpGELVhtlGtaHVYsnHVEaLKeQLqRvPvaf.PvLvfkker....efLedY........estDmeVvdYvp.yppIKmeMAV
L.amazonensis   IASYaLLTiLIAkatgLrpGELVhtlGdaHVYrsHIDaLKaQLeRvPhaf.PtLvfkeer....qfLedY........elmDmeVidYvp.hppIKmeMAV
L.major         IASYaLLTiLIAkatgLrpGELVhtlGdaHVYrnHVDaLKaQLeRvPhaf.PtLifkeer....qyLedY........eltDmeVidYvp.hpaIKmeMAV
L.tropica       IASYaLLTiLIAkatgLrpGELVhtlGdaHVYrnHVDaLKaQLeRvPhaf.PtLifkeer....qyLedY........eltDmeVidYvp.hpaIKmeMAV
P.chabaudi      IASYsIFTyMIAqvcnLqpaEFIhvlGnaHVYnnHVEsLKvQLnRtPypf.PtLklnpei....knIedF........tisDftVqnYvh.hdkIsmdMAa
P.falciparum    IASYsIFThMIAqvcnLqpaqFIhvlGnaHVYnnHIDsLKiQLnRiPypf.PtLklnpdi....knIedF........tisDftIqnYvh.hekIsmdMAa
E.coli          IASYaLLvhMmAqqcdLevGDFVwtgGdtHlYsnHMDqtHlQLsRePrpl.PkLiikrkp....esIfdY........rfeDfeIegYdp.hpgIKapVAI
B.subtilis      IASYaLLTmIIAhvtgLepGEFIhtfGdvHIYqnHIEqVnlQLeRdvrpl.PqLrfarkv....dsIfnF........afeDfiIedYdp.hphIKgaVsV
L.lactis        amqYvaLqmMIAkhfswkvGkFfyfvnnlHIYdnqfEqanelMkRtasekePrLvlnvp.....dgtnfF.....dikpeDfeLvdYepvkpqLKfdLAI
L.casei         IASYaLLThLVAhecgLevGEFIhtfGdaHLYvnHLDqIKeQLsRtPrpa.PtLqlnpdk....hdIfdF........dmkDikLlnYdp.ypaIKapVAV
Tn4003          IASYaLLThLVAkecgLevGEFIhtfGdaHIYsnHMDaIHtQLsRdsylp.PqLkintdk.....sIfdi........nyeDleLinYes.hpaIKapIAV
Phage T4        IASYatLvhIVAkmcnLipGDLIfsgGntHIYmnHVEqcKeiLrRePkel.ceLvisglpykfrylstkeqlkyvlklrpkDfvLnnYvs.hppIKgkMAV
Phage phi3T     VfqYnVLqrMIAqvtgyelGEYIfniGdcHVYtrHIDnLKiQMeReqfea.PeLwinpev....kdfynF........tvdDfkLinYkh.gdkLlfeVAV
```

```
                                                          17       A       31  -i              42       i    54                  68   B
                                                          |-------------|  |^|              |^^^^^^^^^^^|                    |--------
VZV                          mgdlscwtkvpgftltgelqYlkqvddILryGvrkr..........DRTGiGTLSlFG.mqaRYnLrne.FPLLTTKR.VfwraVveEL
HVS                                  msthteeqhgehqYlsqvqhILnyGsfkn..........DRTGtGTLSiFG.tqsRFsLene.FPLLTTKR.VfwrgVveEL
HVA                                       meelhaehqYlsqvkhILncGnfkh..........DRTGvGTLSvFG.mqsRYsLekd.FPLLTTKR.VfwrgVveEL
Human           mpvagselprrplppaaqerdaeprpphgelqYlgqiqhILrcGvrkd..........DRTGtGTLSvFG.mqaRYsLrde.FPLLTTKR.VfwkgVleEL
Mouse           mlvvgsel......qsdaqqlsaeaprhgelqYlrqvehILrcGfkke..........DRTGtGTLSvFG.mqaRYsLrde.FPLLTTKR.vfwkgVleEL
C.albicans                          mtvspntaeqaYldlckrIIdeGehrp..........DRTGtGTkSlFappqlRFdLsndtFPLLTTKK.VfskgIihEL
P.carinii                             mvnaeeqqYlnlvqyIInhGedrp..........DRTGtGTLSvpapsplKFsLrnktFPLLTTKR.VflrgVieEL
S.cervisiae                         mtmdgknkeeeqYldlckrIIdeGefrp..........DRTGtGTLSlFappqlRFsLrddtFPLLTTKK.VftrgIilEL
C.fasiculata          NH2 -223- mkyvphnaeerqYlelidrIMktGlvke..........DRTGvGTISlFG.aqm.FsLrdnqLPLLTTKR.VfwrgVceEL
L.amazonensis         NH2 -224- ckyvprnheerqYlelidrIMktGiake..........DRTGvGTLSlFG.aqmRFsLrdnrLPLLTTKR.VfwrgVceEL
L.major               NH2 -224- ckyvprnheerqYlelidrIMktGivke..........DRTGvGTISlFG.aqmRFsLrdnrLPLLTTKR.VfwrgVceEL
L.tropica             NH2 -224- ckyvprnheerqYlelidrIMktGivke..........DRTGvGTISlFG.aqmRFsLrdnrLPLLTTKR.VfwrgVceEL
P.chabaudi            NH2 -280- nsikykhhpeyqYlniiydIImhGnkqd..........DRTGvGvLSkFG.ymmKFnLsey.FPLLTTKK.LfvrgIieEL
P.falciparum          NH2 -315- nslkykyhpeyqYlniiydIMmnGnkqs..........DRTGvGvLSkFG.yimKFdLsqy.FPLLTTKK.LflrgIieEL
E.coli                                          mkqYlelmqkVLdeGtqkn..........DRTGtGTLSiFG.hqmRFnLqdg.FPLVTTKR.chlrsIihEL
B.subtilis                                      mkqYkdfcrhVLehGekkg..........DRTGtGTIStFG.yqmRFnLreg.FPMLTTKK.LhfksIahEL
L.lactis                                 mtyadqvFkqniqnILdnGvfsenarpkyk...DgqmanSkyvtG.sfvtYdLqkgeFPI.TTlRpIpiksaikEL
L.casei                                      mleqpYldlakkVLdeGhfkp..........DRThtGTySiFG.hqmRFdLskg.FPLLTTKK.VpfglIksEL
Tn4003                                   mynpfdeaYhglceeILeiGnrrd..........DRThtGTISkFG.hqlRFdLtkg.FPLLTTKK.VsfklVatEL
Phage T4                                        mkqYqdlikdIfenGyetd..........DRTGtGTIalFG.sklRWdLtkg.FPaVTTKK.LawkaciaEL
Phage phi3T                              mtqfdkqYnsiikdIInnGisdeefdvrtkwdsDgTpahTLSvms.kqmRFdnse..vPILTTKK.VawktaikEL

                  82   86  C  93         99 D 103                                                           122 G  127
                -----|   |------|         |---|                                                              |----|
VZV             LWFIrGsTd.skeLaakd.......IHIWdiygsskflnrngfhkrht..............................gDLGPIYGfQWRhFg
HVS             LWFIrGsTd.skeLsaag.......VHIWdangsrsfldklgfydrde..............................gDLGPVYGfQWRhFg
HVA             LWFIrGsTd.skeLaasg.......VHIWdangsrsyldklglfdree..............................gDLGPVYGfQWRhFg
Human           LWFIkGsTn.akeLsskg.......VKIWdangsrdfldslgfstree..............................gDLGPVYGfQWRhFg
Mouse           LWFIkGsTn.akeLsskg.......VRIWdangsrdfldslgfsarqe..............................gDLGPVYGfQWRhFg
C.albicans      LWFVaGsTd.akiLsekg.......VKIWegngsrefldklglthrre..............................gDLGPVYGfQWRhFg
P.carinii       LWFIrGeTd.slkLrekn.......IHIWdangsreyldsigltkrqe..............................gDLGPIYGfQWRhFg
S.cerevisiae    LWFLaGdTd.anlLseqg.......VKIWdgngsreyldkmgfkdrkv..............................gDLGPVYGfQWRhFg
C.fasiculata    IWFLrGeTn.ahvLadkd.......IHIWdgngsrefldsrgltenke..............................mDLGPVYGfQWRhFg
L.amazonensis   LWFLrGeTn.aqlLadkd.......IHIWdgngsrefldsrgltente..............................mDLGPVYGfQWRhFg
L.major         LWFLrGeTs.aqlLadkd.......IHIWdgngsrefldsrgltenke..............................mDLGPVYGfQWRhFg
L.tropica       LWFLrGeTs.aqlLtdkd.......IHIWdgngsrefldsrgltenke..............................mDLGPVYGfQWRhFg
P.chabaudi      LWFIrGeTn.gntLlekn.......VRIWeangtrefldnrklfhrev..............................nDLGPIYGfQWRhFg
P.falciparum    LWFIrGeTn.gntLlnkn.......VRIWeangtrefldnrklfhrev..............................nDLGPIYGfQWRhFg
E.coli          LWFLqGdTn.iayLhenn.......VtIWdewaden..........................................gDLGPVYGkQWRaWp
B.subtilis      LWFLkGdTn.vryLqeng.......VRIWnewaden..........................................gELGPVYGsQWRsWr
L.lactis        MWiyqdqTselsvLeeky......gVKyWgewgigd..........................................gtiGqrYGatvKkYn
L.casei         LWFLhGdTn.irfLlqhr.......nHIWdewafekwvksdeyhgpdmtdfghrsqkdpefaavyheemakfddrvlhddafaakygDLGlVYGsQWRaWh
Tn4003          LWFIkGdTn.iqyLlkyn.......nnIWnewafenyvqsddyhgpdmtdfghrsqqdpefneqykeemkkfkerilnddafakkygnLGnVYGkQWRdWe
Phage T4        IWFLsGsTn.vndLrliqhdsliqqktVWdenyenqakdlgyhs..................................gELGPIYGkQWRdF.
Phage phi3T     LWiwqlksndvteLnkmg.......VHIWdqwkqed..........................................gtiGhaYGfQlgkkn
```