MISINCORPORATION BY AMV REVERSE TRANSCRIPTASE AND ITS POTENTIAL FOR MUTAGENESIS

Judith Ann Skinner BSc. (Hons. Leeds)

A thesis submitted in accordance with the requirements of the University of Leicester for the degree of Doctor of Philosophy.

Department of Biochemistry

January 1987

UMI Number: U000894

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.

UMI U000894
Published by ProQuest LLC 2015. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.
All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.

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


LEICESTER
UNIVERSITY
LIBRARY

Statement

This thesis is based on work conducted by the author in the Department of Biochemistry of the University of Leicester mainly during the period between August 1984 and July 1986.

All the work recorded in this thesis is original unless otherwise acknowledged in the text or by references. None of the work has been submitted for any other degree in this or any other university.

Signed: J.R. Skinner

Date: 21.12.86

I dedicate this work
to my Saviour,
to Mike
and to my parents.

Without them it would be nothing.

ACKNOWLEDGMENTS

I acknowledge the support of the Cancer Research Campaign in financing this work with a grant SP 1715.

I thank my supervisor, Dr. Ian Eperon, for his invaluable help throughout the time that I spent in his laboratory, and especially for the time and effort that he spent in reading my manuscript.

I also want to express my thanks to all the members of lab 216: Lucy, Peter, Andrew (Griff), Ian (Whizz), Tony and Caroline, for all their advice, help and their unfailing friendship and support. Thanks go also to Kath for the use of her bean-bag!

I thank Professor W.V. Shaw for accepting my application to register for a Ph.D., and also all the other members of the Biochemistry department for their advice and helpful discussions. Particular thanks go to John Keyte for the excellent and rapid synthesis of so many oligonucleotides.

I thank Mrs Bridget Wasilewski for her skilful typing of the bibliography, and I extend my thanks to the Marie Curie Memorial Foundation research institute for their patience whilst I finished writing-up.

Most of all I want to thank Mike for his constant love and care, his exacting reading and re-reading of the manuscript, and for his patience and support.

Misincorporation by AMV-reverse transcriptase and its potential for mutagenesis.

Judith A. Skinner

This thesis describes a systematic investigation of the efficiency of misincorporation by Avian Myeloblastosis Virus reverse transcriptase with all possible combinations of dNTP substrate, template nucleotide, and the nucleotide at the 3' terminus of the primer. Each of a series of 16 synthetic oligonucleotide primers was annealed to single-stranded M13 DNA templates, and a single dNTP was misincorporated at the primer 3' end using AMV reverse transcriptase. The proportion and pattern of misincorporation and incorporation in all 64 situations were assayed using 5'-labelled primers, and the products were separated on denaturing polyacrylamide gels.

Correct incorporations occurred more readily than misincorporations. The efficiency of misincorporation depended on the individual primer but, comparing primers, a clear dependence on the template nucleotide was observed for the preferential misincorporation of different dNTPs. The exact combination of template nucleotide and dNTP was important; although purine:pyrimidine (dNTP substrate: template nucleotide) and pyrimidine:purine misincorporations occurred comparatively readily, some pyrimidine:pyrimidine and purine:purine reactions were equally efficient and yet others were never seen to occur. Some misincorporations were facilitated by subsequent correct incorporations, but despite this the results suggest that the level of misincorporation is limited by the rate of reaction and enzyme inactivation rather than by exonuclease activity.

The recovery of point mutants arising from reverse transcriptase-directed misincorporation of single dNTPs onto single oligonucleotide primers is described and discussed.

Misincorporation of dNTPs is a form of *in vitro* mutagenesis which facilitates the generation of a library of point mutations spread throughout a gene. Conditions have been established in this study for the production of a bank of primers with 3' termini distributed over a region of a gene to be mutated. The misincorporation of single dNTPs onto the termini of such a bank of primers should allow the generation of a library of point mutants.

TABLE OF CONTENTS

INTRODUCTION

MATERIALS AND METHODS

INDEX OF FIGURES AND TABLES

(Figures and Tables are found inserted into the text after the appropriate page)

ABBREVIATIONS

DNA	deoxyribonucleic acid
cDNA	complementary DNA
SS DNA	single-stranded DNA
SS	single-stranded
RF	replicative form
CC DNA	closed-cicular double-stranded DNA
CC	closed-circular
dNMP	2'-deoxyribonucleoside 5'-monophosphate
dNTP	2'-deoxyribonucleoside 5'-triphosphate
ddNTP	2',3'-dideoxyribonucleoside 5'-triphosphate
α-SdNTP	α-phosphorothio 2'-deoxyribonucleoside 5'-triphosphate
N	one of the four bases:
A	adenine
C	cytosine
G	guanine
T	thymine
pyr	pyrimidine
pu	purine
rATP	adenosine triphosphate
RNA	ribonucleic acid
mRNA	messenger RNA
tRNA	transfer RNA
$tRNA^{Phe}$	transfer RNA: phenylalanine specific
nucleotide	deoxyribonucleotide
dideoxynucleotide	dideoxyribonucleotide
thionucleotide	α-phosphorothiodeoxyribonucleotide
bp	base pairs
kb	kilo base pairs
RMLV	Rauscher Murine Leukaemia Virus
RSV	Rous Sarcoma Virus
AMV	Avian Myeloblastosis Virus
SDS	Sodium dodecyl sulphate
EDTA	Ethylene diamine tetraacetic acid
DTT	Dithiothreitol
Tris	Tris (hydroxymethyl)-aminomethane
MES	2-(N-Morpholino) ethane sulphonic acid
LMP	Low melting point
PAGE	Polyacrylamide gel electrophoresis
EtBr	Ethidium bromide
TEMED	N,N,N',N'-tetramethylene diamine
5-Bu	5-Bromouracil
SB	Standard broth
IPTG	Isopropyl-β-D-thiogalactopyranoside
BCIG	5-bromo 4-chloro-3-indolyl β-galactoside
PEG/NaCl	20% polyethylene glycol 6000, 2.5M NaCl
Klenow polymerase	*E. coli* DNA polymerase I, Klenow fragment
Km	Michaelis constant
HPLC	High pressure liquid chromatography
TLC	Thin layer chromatography
PEI	Polyethyleneimine
M.W.	Molecular weight

INTRODUCTION

Chapter 1. Introduction

Mutants have long been one of the major tools in the structural and functional analysis of genes and their regulatory elements. Mutants arising from mutagenesis *in vivo* have only been isolated if a phenotypic change occurred as a result of that mutagenesis; thus the analyses that can result from *in vivo* mutagenesis have been very limited. Furthermore, the non-specific mutagens used to generate mutants *in vivo* can cause multiple alterations in the DNA that may affect subsequent analyses of any mutants that are isolated.

The development of DNA manipulation techniques over the last decade has permitted mutagenesis *in vitro* as a result of direct modification of specified regions of DNA. Initially *in vitro* mutagenesis experiments involved the deletion or insertion of fragments of DNA, but refinements in the methods of DNA manipulation have allowed point mutations to be directed to discrete regions of the DNA by chemical modification of isolated fragments.

The development of suitable single-stranded (SS) bacteriophage vectors has increased the range of specific mutagenesis techniques available for the introduction of both point mutations and deletions, and the rapid sequence determination of SS DNA (Sanger *et al.*, 1977) has facilitated the identification and characterization of

mutants.

Whilst fine structure analysis of DNA regulatory elements and proteins can be carried out following the introduction of specific point mutations, there exists a need for a simple and effective procedure to produce a "library" of mutant molecules each with a single random point mutation. The ability to produce such libraries of mutants would allow more detailed investigations of entire genes, facilitating the understanding of the interactions of different domains within an encoded protein and between proteins. Such an understanding is a necessary prerequisite for the developing field of "protein engineering".

The development of a method of mutagenesis to generate libraries of mutants is presented in this thesis. The method requires the misincorporation of dNTPs by AMV reverse transcriptase at single sites spread throughout a region of DNA. The misincorporation sites within the template sequence are defined by the positions of the 3' termini of primers which span such a region. These primers are generated randomly in a preliminary reaction by dNTP-limited extensions of a single oligonucleotide primer.

A study of reverse transcriptase-directed misincorporation was undertaken in an attempt to establish the parameters controlling the misincorporation of a dNTP at any particular template nucleotide adjacent to each of the four possible primer 3' terminal nucleotides.

The introduction to this thesis comprises Chapters 1 to 7. In Chapter 2 the development and applications of *in*

vitro mutagenesis are introduced and Chapter 3 is an introduction to reverse transcriptase, the error-prone DNA polymerase that can be used in misincorporation mutagenesis. The isolation of mutants is considered in Chapter 4, and Chapter 5 introduces the mechanisms of DNA mismatch repair in *E. coli* that influence the recovery of mutants. Chapter 6 is an introduction to the analysis of misincorporation reactions, and this is followed in Chapter 7 by a brief summary of the Introduction and an outline of the experiments presented in this thesis. Chapter 8 contains the Materials and Methods.

The experimental Results are presented in Chapter 9. In the first section results are presented of the *in vitro* assay of reverse transcriptase-directed misincorporation, followed in the second section by results of the generation of a "bank" of primers annealed to the template DNA for subsequent misincorporation reactions. The results in the third section of Chapter 9 describe the mismatch mutants that have been recovered. The last section of Chapter 9 contains the results of the preparation of a cloned gene for the described misincorporation methodology: the subcloning of the gene into a SS phage vector, the generation of a bank of primers for that template, and the misincorporation by reverse transcriptase.

In the Discussion (Chapter 10) are considered the various factors that were either observed or have been proposed to influence the *in vitro* assay of misincorporation. A comparison is made of the misincorporations

observed *in vitro* and the actual misincorporation mutants recovered, and the results are compared to and contrasted with data available in the literature. The design of mutagenesis strategies involving AMV reverse transcriptase is discussed.

Chapter 11 discusses the application of *in vitro* mutagenesis to protein engineering.

CHAPTER 2. *In vitro* mutagenesis.

In vitro mutagenesis is used to introduce mutations into cloned DNA. It can be executed far more efficiently and specifically than mutagenesis *in vivo* because the DNA has been isolated and cloned into an appropriate vector. Another important advantage is that the screening of the resultant progeny need not depend upon having introduced a phenotypic change.

Mutants so created can be used for a variety of purposes:

1. to investigate the role of specific sequences in DNA and RNA in controlling gene expression,
2. to determine the roles of specific proteins within cells,
3. to study the relationship between structure and function of proteins, and their interactions with other parts of the cell machinery.

There are many different methods of mutagenesis *in vitro*, encompassing both mutagenesis of large regions of DNA, and mutagenesis of single predetermined sites. A wide range of problems can be tackled by mutagenesis *in vitro* and there is no single method appropriate to all situations. The particular method chosen in any one study must depend upon the questions being asked and the extent of the knowledge so far obtained about that particular nucleic acid sequence, protein or interactive system. Over the last decade many varied methods have been devised and these have been reviewed by Shortle *et al.* (1981), Lathe *et al.* (1983),

Botstein and Shortle (1985) and Smith (1985).

Where a gene, or DNA fragment, of interest has been cloned into a suitable vector, then it can be mutated *in vitro* in a number of ways (listed below and described in more detail later in this chapter):

1) deletion mutagenesis and insertion mutagenesis is used where regions are to be identified that may be worthy of further study. Large regions are deleted using specific restriction enzymes, or by careful use of degradative enzymes; in the latter case a range of deletion sizes can be obtained. Inserted DNA is used to disrupt regions of the native DNA;

2) oligonucleotide mutagenesis can be employed when the sequence of the mutation site is known. Although normally used to introduce single specific mismatches, multiple mismatches and short deletions or insertions can also be introduced;

3) a number of different mutants can be obtained from a single mutagenesis reaction if the DNA is subjected to limited chemical modification. The extent of the modification can be controlled by alteration of reaction conditions and by restricting the size of the target DNA, either by using isolated DNA fragments or by protecting regions that are not to be mutated. Thus a number of different point mutants can be recovered from the same reaction;

4) incorporation of dNTP analogues onto the 3' termini of primers annealed to a template DNA strand can be used to

generate specific point mutations at defined positions;

5) misincorporation of normal dNTPs at defined positions can occur by using biased pools of dNTPs, mutagenic reaction conditions for the polymerase, α-SdNTPs and/or an error-prone polymerase. Reverse transcriptase gives higher rates of misincorporation than DNA polymerase I (Battula and Loeb, 1974). The sites for misincorporation can be determined by using the 3' termini of synthetic oligonucleotides or of restriction fragments. A more random spread of termini can be achieved by random single strand nicking of a duplex and progressive digestion from exposed 3' terminal nucleotides, or by random primer-extensions from an oligonucleotide primer or restriction fragment.

2.1 Efficiency of mutagenesis

Where no phenotypic selection of mutants is possible, then any method of *in vitro* mutagenesis chosen should be as efficient as possible so that fewer candidates have to be screened to recover the mutants. However, when a method of mutagenesis such as direct chemical modification is used, the efficiency of mutagenesis must be reduced to such a level that only one mutational event occurs per molecule within that region of DNA. There is, therefore, an optimal efficiency to give single mutations that can be readily detected by screening. For this reason the localised region of DNA exposed to mutagenesis is often limited because the smaller the site of mutagenesis, the more efficient can be

the method. This can be taken to the ultimate situation where the site of mutagenesis is a single base pair which can then be subjected to mutagenesis with a potential efficiency of 100%. This is effectively the case in oligonucleotide-directed mismatch mutagenesis where the site is defined within the oligonucleotide. That the recovery is rarely so good is dependent upon other factors, discussed later in Chapter 5.

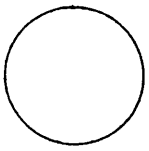
2.2 Deletion mutagenesis.

The deletion of specific sequences from the cloned DNA is often the first approach to identifying regions of particular interest for further study. It is unusual to find two restriction sites conveniently placed either side of the DNA fragment to be deleted and so degradative enzymes such as *Bal* 31 have been employed. *Bal* 31 digests DNA progressively from exposed termini, both 3'-5' and 5'-3', and is used for creating deletions in both directions from a unique restriction site (Gray *et al.*, 1975). *E. coli* exonuclease III is a double stranded 3'-5' exonuclease which can also be used to generate deletions (Guo and Wu, 1983). The DNA end that is not to be digested can be protected either by digestion with *Pst I*, or other restriction enzymes that leave 3' overhangs (Henikoff, 1984), or by the addition of an α-SdNTP. Exonuclease III cannot digest phosphorothioate linkages (Putney *et al.*, 1981). The remaining SS DNA is removed by S1 nuclease digestion before

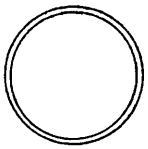
Figure 1.
Generation of point mutations in M13 *in vitro* (excluding mutagenesis of isolated DNA fragments

A Chemical mutagenesis (*e.g.* bisulphite, Skinner *et al.*, 1984)
BI Oligonucleotide mutagenesis
BII Oligonucleotide mutagenesis using the gapped-duplex method
C Misincorporation
CI Misincorporation using Klenow polymerase
CII Misincorporation using reverse transcriptase as described in this study
CIII Misincorporation using reverse transcriptase and a bank of primers as described in this study

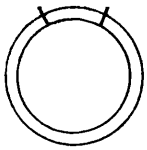
a=anneal; b=copying reaction (Klenow polymerase + 4dNTPs + T4 DNA ligase); c=misincorporation (biased pool of dNTPs + Klenow polymerase); d=misincorporation (single dNTP + reverse transcriptase); e=chase reaction (4dNTPs + reverse transcriptase); f=limited extension reaction (Klenow polymerase + dCTP,dGTP,dTTP + limited dATP)



SS DNA



RF-double stranded DNA



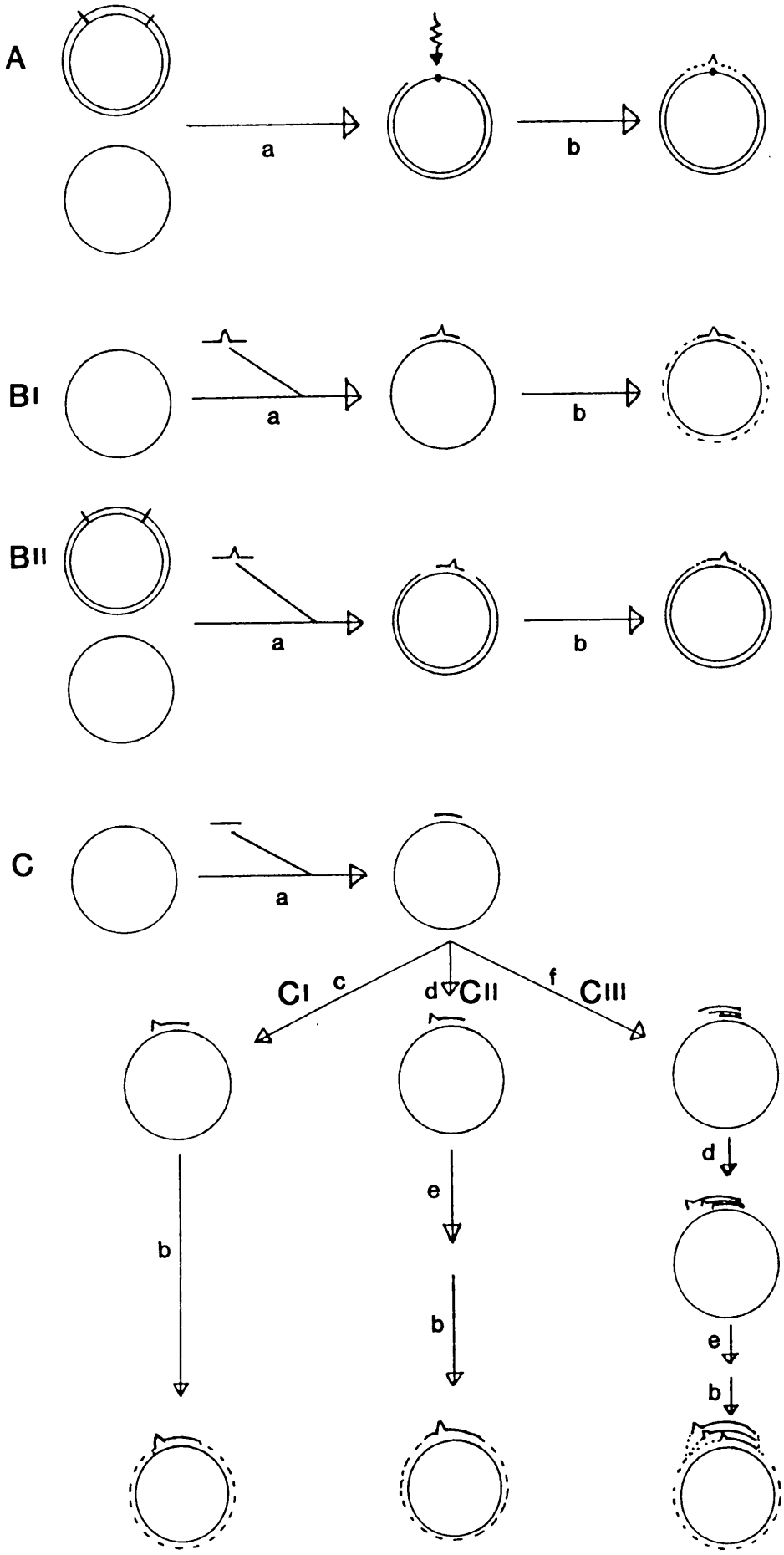
RF linearised at restriction sites

mismatch oligonucleotide primer

primer, can be either an oligonucleotide or a restriction fragment

newly synthesized DNA

chemical attack


A
a
b
BI
a
b
BII
a
b
C
a
CI
c
d
CII
f
CIII
b
e
b
d
e
b

religating the blunt-ended DNA. Where no suitable restriction enzyme sites are available, deletions can be introduced after partial DNase I digestion in the presence of Mn^{2+}, which produces random double strand breaks (Shenk *et al.*, 1976). One problem that does arise when using progressive degradative enzymes to create deletions is that they pause at different positions in the DNA giving preferred end points of digestion.

2.3 Oligonucleotide site-directed point mutagenesis.

(see figure 1)

Single point mutations are most easily introduced by oligonucleotide mutagenesis, reviewed by Smith and Gillam (1981) and by Smith (1985). In the last few years synthetic oligonucleotides have become much more widely available, synthesized to higher purity than before, and often much more cheaply. The very low cost is a particular feature of the paper disc method (Matthes *et al.*, 1984).

The target DNA is cloned in preference into a SS vector for ease of isolation of pure SS DNA and for screening by oligonucleotide hybridisation (Zoller and Smith, 1982) or dideoxynucleotide sequencing (Sanger *et al.*, 1977). Derivatives of the filamentous bacteriophage M13 have been developed as the M13mp series of vectors (Messing, 1983; Yanisch-Perron *et al.*, 1985), which have two distinctive and useful features: the multiple cloning site with a wide range of unique restriction enzyme cleavage sites clustered

together, and also a simple phenotypic screen for the presence of DNA cloned within the phage DNA. This screening works because any cloned insert disrupts the reading frame of the amino-terminal portion of the *E.coli* β-galactosidase (Z) gene (part of the *lac* operon) which forms part of the vector. Vectors carrying no insert DNA allow correct expression of β-galactosidase by complementation with a defective β-galactosidase gene carried on the F episome of the host cell and result in blue plaques when grown in the presence of the inducer isopropyl-β-D-thiogalactopyranoside (IPTG) and the substrate BCIG (5-bromo 4-chloro-3-indolyl β-galactoside). Any insert of foreign DNA cloned into the β-galactosidase gene portion found in the M13 vector will disrupt the gene and so prevent complementation, resulting in colourless (white) plaques.

Oligonucleotide mutagenesis has also been performed with the target DNA cloned in the SS phage ϕX174 (Hutchinson *et al.*, 1978; Razin *et al.*, 1978) or in double stranded vectors (Wallace *et al.*, 1980).

The technique of oligonucleotide mutagenesis requires an oligonucleotide of identical sequence, except for the mutation(s) to be introduced, to a site in the cloned wild-type gene. For a single base substitution the oligonucleotide should have two or three correct nucleotides 3' and six or seven nucleotides 5' of the mismatch for site specificity, efficient priming, and retention of the mutant sequence (Gillam and Smith, 1979). The longer the oligonucleotide is, then the more efficient the priming

becomes, and lengths of 15 and 17 are routinely used. When more than one mismatch is introduced per primer, or when small deletions or insertions are introduced then longer primers are used accordingly. The oligonucleotide is annealed to the SS template which is either SS phage DNA or a gapped double-stranded DNA. The short heteroduplex can act as a primer for extension by DNA polymerase I Klenow fragment (Klenow polymerase) all the way around the template, whereupon the remaining nick is sealed by T4 DNA ligase. The resulting closed circular (CC) DNA has a wild-type sequence in the template strand and a mutant sequence in the newly synthesized strand, and is transfected into competent *E.coli* (JM101 in the case of M13 DNA). The DNA molecules undergo replication and the two different strands segregate allowing subsequent isolation of pure clones of either wild-type or mutant.

Oligonucleotide mutagenesis was developed by Zoller and Smith (1982 and 1983); a second method described by them in 1984 used two oligonucleotides, the first a 5' phosphorylated mismatch primer (as above) and the second an upstream primer (not necessarily 5' phosphorylated). This two primer method enables transfection of copied and ligated material without first requiring isolation of CC DNA because the 5' end of the mismatch primer is protected from *E.coli* 5'-3' exonuclease activity *in vivo*. Recovery of mutants is considered in Chapter 4.

Oligonucleotide-directed mutagenesis is used to create predetermined mutants at a defined site of mutagenesis. All

the newly synthesized strands contain the mutation and the technique is ideal when specific alterations are desired. Each different mutation desired requires a specific oligonucleotide, although mixed populations can be created where the oligonucleotide is synthesized as a mixture of similar oligonucleotides with a low percentage of base substitutions at a number of positions (Matteucci and Heynecker, 1983). However, where point mutations are required throughout a large region of interest with only one base substitution per molecule then misincorporation or chemical methods of mutagenesis, as a first approach, may prove more simple and cost effective than using a mixture of oligonucleotides.

2.4 Chemical mutagenesis

Chemical methods can be used to create point mutations in two different ways:

(i) to damage the bases of a SS template DNA so that *in vitro* synthesis of the complementary strand by DNA polymerase gives misincorporations at the sites of damaged or removed bases;

(ii) to alter the base-pairing specificity of the DNA within the region of interest.

The first strategy involves chemical damage to DNA so preventing subsequent base-pairing. Formic acid removes purines from SS DNA and such depurinated DNA can be used as a template for second strand synthesis by DNA polymerase.

Misincorporation occurs at the apurinic site, with dATP being the misincorporated nucleotide of preference (Schaaper *et al.*, 1983). Reverse transcriptase is used as the DNA polymerase of choice because it lacks a 3'-5' exonuclease activity. Other chemicals are also used, *eg.* nitrous acid, hydrazine, potassium permanganate and dimethylsulphate (Myers *et al.*, 1985c).

The second strategy changes a base so that it pairs differently from the unmodified base. The most established method is deamination of cytosine to uracil with the inorganic salt sodium bisulphite. It is specific to SS DNA and, after replication, results in the change of a CG base pair to a TA base pair (Shortle and Botstein, 1983). The region to be subjected to mutagenesis is made SS usually by creating a gapped-duplex molecule. A small gap can be introduced by a single strand nick of the DNA followed by limited digestion by an exonuclease (Shortle and Nathans, 1978). Randomly introduced small gaps allow for efficient deamination by sodium bisulphite as the number of cytidines accessible on SS DNA will be limited by the size of the gap. Where the region to be mutated is larger it can be defined by annealing a linear vector molecule to a SS recombinant molecule containing the DNA of interest. The SS gap can be more closely defined by judicious use of restriction enzyme sites, on either side of the region to be mutated, to linearise the replicative form of the recombinant molecule.The double-stranded linear DNA is denatured and annealed to the SS form of the recombinant

molecule producing a gapped-duplex with the SS gap available for modification with sodium bisulphite (Skinner *et al.*, 1984; figure 1). The randomness of the distribution of point mutants within such a region is determined by limiting the deamination reaction time and by controlling the concentration of bisulphite. Having modified the cytosine bases in the SS DNA the molecule can be made double-stranded by addition of Klenow polymerase and DNA ligase giving pairing of adenine to uracil (i.e. deaminated cytosine). The mutated DNA is then transfected into *E. coli* and the progeny phage are screened for the appropriate point mutations.

Other chemicals have been used to alter base pairing in DNA so producing mutations: this usually involves subjecting a fragment to base modification and then returning it to the correct position in a wild-type genome by either recombination or by annealing the mutant strand to a SS template. Initial use of a fragment restricts the region which can subsequently carry mutations. If both strands of the fragment are mutagenised independently and then annealed and inserted into a vector, then a higher yield of single point mutations can be achieved (Kadonaga and Knowles, 1985). Hydroxylamine is used to deaminate cytosine to hydroxyaminocytosine, an analogue of thymine, changing CG base pairs to TA base pairs (Busby *et al.*, 1982). Methoxylamine also reacts preferentially with cytosine (Kadonga and Knowles, 1985); nitrous acid deaminates cytosine to uracil, and adenine to hypoxanthine, an analogue of guanine (Myers *et al.*, 1985c). Myers and co-workers

(1985c) also made use of hydrazine which induces non-specific pairing of purines to pyrimidines.

With all the chemical methods of mutagenesis, point mutations can only be produced efficiently if the available site of action is kept small. As the size of target DNA is increased so the efficiency must be decreased to prevent the production of mutants containing more than one mutation. Where selection of mutants or screening for a changed phenotype is possible, then a low efficiency of mutagenesis is not a great problem, but if the mutations are phenotypically silent and all screening is by sequencing then the mutation efficiency should be raised as high as possible.

2.5 Mutagenesis by incorporation of dNTP analogues

DNA polymerases are used to incorporate analogues of the correct dNTP normally incorporated on second strand synthesis *in vitro*. The polymerase requires a region of SS DNA as a template with a free 3' end of a primer to which the dNTP analogue is added. The dNTP analogue has two tautomeric forms, imino and amino, and if incorporated in the amino form, the imino form can pair with a different dNTP during replication *in vivo*. Müller *et al.* (1978) created a SS gap by nicking at a restriction enzyme site which produced a specific priming site, and then used the dNTP analogue N^{6}hydroxy dCTP. The amino form of N^{6}hydroxy dCTP pairs with dGTP but the imino form pairs

with dATP. Other dNTP analogues that have been used are 2-aminopurine deoxyriboside (Grossberger and Clough, 1981), O^6-methyldeoxyguanosine (Dodson *et al.*, 1982; Eadie *et al.* 1984) and 5-bromodeoxyuridine (Mott *et al.*, 1984).

The position of the mutation depends upon the position of the 3' end of the primer (which can be an oligonucleotide, restriction fragment, or the 3'end of a SS gap, see figure 1). The position of the 3' end can be moved either by limited exonuclease III digestion or by extension with DNA polymerase in conditions lacking one or more dNTPs so that the 3' end is at a predetermined position. The incorporation reaction requires addition of the analogue and sometimes also the correct dNTP as a competitor. By varying the ratios of different dNTPs provided, the range of mutants achieved can be extended. AMV reverse transcriptase is sometimes used in preference to Klenow polymerase because it lacks the proofreading 3'-5' exonuclease activity.

2.6 Misincorporation mutagenesis (figure 1)

Misincorporations occur in DNA synthesis when the DNA polymerase incorporates an incorrect dNTP producing a point mutation in the newly synthesized strand. Normally any such misincorporation *in vitro* is removed by the proofreading 3'-5' exonuclease activity of DNA polymerase I. However, the *in vitro* error rate can be greatly increased by altering the reaction conditions. The frequency of misincorporation can be increased by substituting manganese or cobalt ions

for magnesium in the polymerase reaction, by adding beryllium ions, or providing an unequal ratio of the four dNTPs for the synthesis (Kunkel and Loeb, 1979). Proofreading repair can be prevented by the use of α-SdNTPs, alternatively AMV reverse transcriptase can be used instead of Klenow polymerase, as described in the next sections.

As with the incorporation of dNTP analogues, the position of the mutation depends upon the position of the 3' end of the primer. The 3' end of the primer can be positioned at predetermined sites by exonuclease III digestion of the primer, or by extension of the primer with DNA polymerase.

The advantages in using misincorporation for mutagenesis are that, in principle, any misincorporation is possible if the 3' end of the primer is in the correct situation. The particular dNTP to be misincorporated can be specified by only providing one dNTP or, conversely, the template nucleotide to be mutated can be specified by providing only three dNTPs and omitting the complimentary dNTP.

2.6.1 Effect of divalent cation alterations on the fidelity of DNA polymerases

Sirover and Loeb (1976) have shown that the substitution of manganese, cobalt, nickel or zinc ions for magnesium ions reduces the fidelity of DNA synthesis by DNA polymerase I. Similar work (Sirover and Loeb, 1977) showed

that fidelity of AMV reverse transcriptase was also reduced on substitution of manganese or cobalt for magnesium and it was argued that as reverse transcriptase lacks a 3'-5' exonuclease activity then this decrease in fidelity must reflect alterations in the selection of dNTPs.

2.6.2 Biased pools of dNTPs

The dNTP substrates provided for the DNA polymerase reaction significantly affect the fidelity of the enzyme. If one of the four dNTPs is omitted from the gap-filling reaction which is driven by ligation of the resulting nick (preventing proofreading), then different misincorporations can occur at each position where the missing dNTP should be incorporated. Shortle *et al.* (1982) have shown that, using *Micrococcus luteus* DNA polymerase I, such misincorporations at *Cla* I and *Hind* III restriction sites can give rise to both transitions and transversions, when omitting either dATP ,dCTP, or dGTP. If a chosen dNTP were not omitted completely, but provided in very limiting concentrations, then a growing DNA strand could be made with misincorporations at a small number of the potential sites within a defined SS gap.

2.6.3 Misincorporation of α-SdNTPs.

Incorporation of an α-SdNTP at the 3' end of a growing DNA molecule creates a phosphorothioate linkage between that

nucleotide and the adjacent 5' nucleotide. Such phosphorothioate linkages are resistant to proofreading by DNA polymerase I during *in vitro* DNA synthesis (Kunkel *et al.*, 1981). If, therefore, an α-SdNTP is misincorporated during *in vitro* synthesis, it is not corrected and a mutation can result. Shortle and co-workers (1982) have shown that mutations of the type expected were produced when *Micrococcus luteus* DNA polymerase I was provided with a single α-SdNTP followed later by gap repair with all four dNTPs and ligase. Shortle and Lin (1985) misincorporated α-SdNTPs at random sites (generated by short exonuclease III digestion from random SS nicks) using Klenow polymerase. They created 77 unique, single base, missense mutations in the Staphylococcal nuclease gene with almost all possible substitutions represented. The analysis of the mutations recovered is discussed later. Abarzua and Marians (1984) generated 21 independent single base subtitutions by misincorporation of α-SdATP and α-SdTTP with Klenow polymerase. The random sites were generated by limited digestion with exonuclease III so that the new 3' OH termini spanned the region of interest.

2.6.4 Error-prone polymerase: reverse transcriptase

An alternative method of avoiding proofreading repair of newly misincorporated nucleotides is to use an error-prone polymerase lacking a 3'-5' exonuclease activity. The most commonly used error-prone polymerase is AMV reverse

transcriptase. This polymerase was used for mutagenesis *in vitro* by Traboni and co-workers (1984). They misincorporated a single dNTP onto the 3' end of a primer. Extension beyond the misincorporation required the addition of all four dNTPs and Klenow polymerase. Zakour and Loeb (1982) extended a primer fragment to a predetermined position on a ϕX174 SS template with DNA polymerase I, dATP and dCTP. Misincorporation was then induced by incubation with reverse transcriptase and an incomplete set of dNTPs and then the extension was completed by the addition of all four dNTPs. Both transition and transversion mutations were recovered by isolation of revertants of an amber ϕX174 mutation.

Zakour and co-workers (1984) misincorporated single dNTPs at designated positions by provision of only one dNTP and reverse transcriptase in much the same way as Traboni *et al.* (1984). They showed that not only restriction fragments but also synthetic oligonucleotide primers and primers extended with T4 DNA polymerase provided suitable 3' sites for such misincorporation. Furthermore, they suggested that multiple single-step elongations from a single primer should produce a series of primers with differing 3' termini which could be used to produce a series of single substitution mutations at different positions.

2.7 Summary

It is clear that so many varied methods of mutagenesis are now available that the choice of method employed must,

to some extent, be dictated by the nature of the problem to be solved. Where a library of point mutations throughout a region of interest is required, misincorporation can provide a useful tool. Analysis of mutants recovered can give further insight into the mechanisms of action of DNA polymerases which in turn should lead to a clearer understanding of how to utilise such enzymes to produce desired mutations with greater efficiency.

CHAPTER 3. Reverse Transcriptase (DNA polymerase of retrovirus, EC 2.7.7.7)

Reverse transcriptase is a DNA polymerase, found in retroviruses, which can transcribe RNA into complementary DNA (cDNA). Its discovery was reported simultaneously and independently by Baltimore and by Temin and Mizutani in 1970. Baltimore discovered a ribonuclease-sensitive DNA polymerase in purified virions of Rauscher murine leukaemia virus (RMLV), and Temin and Mizutani found the same activity in purified virions of Rous sarcoma virus (RSV). Thereafter this activity was found in many other retroviruses (reviewed in 1972 by Temin and Baltimore).

Most of the research carried out on reverse transcriptase has been restricted to the enzymes of three viruses: avian myeloblastosis virus (AMV), RMLV and RSV. They have been purified by first isolating the virions from the plasma of infected animals (AMV and MLV) or from the supernatant medium of infected cell cultures (AMV, RSV and RMLV). The reverse transcriptase is located in the core of the virion attached to the 60-70S RNA and is released by solubilization with nonionic detergent. The purification procedures are described by Verma (1981, and references therein).

Reverse transcriptase has been shown to have at least three enzymatic activities: (a) RNA-directed DNA polymerase which transcribes DNA from RNA, (b) DNA-directed DNA polymerase which converts SS DNA to double stranded DNA and

(c) RNase H which degrades the RNA strand of a DNA-RNA hybrid (Verma, 1977).

Early investigations on reverse transcriptase failed to detect any endonuclease activity but a DNA endonuclease of molecular weight 32,000 has been purified from AMV and is known as $p32^{endo}$ (Grandgenett *et al.*, 1980).

The DNA polymerase activity of reverse transcriptase can utilize either DNA or RNA templates with either oligodeoxyribonucleotides or oligoribonucleotides as primers, although oligodeoxyribonucleotides are the more efficient primers (Verma, 1977). *In vivo* the primer used for initation of viral DNA synthesis is tRNA. It has been suggested that reverse transcriptase may copy long, single stranded DNA templates inefficiently (Hurwitz and Leis, 1972) because poly(dA).oligo(dT) is almost inactive as a template-primer as compared to poly(A).oligo(dT) template-primer (Wells *et al.*, 1972; Goodman and Spiegelman, 1971 and Robert *et al.*, 1972).

Reverse transcriptase has no 3'-5' exonuclease activity, a major difference between it and the bacterial DNA polymerases, e.g. *E. coli* DNA polymerase I. The 3'-5' exonuclease activity of DNA polymerases is in part responsible for the fidelity of DNA replication, by proofreading the nucleotides newly polymerised onto the 3' end of the extending primer (Kornberg, 1974). Reverse transcriptase lacks this enzymic activity and is often referred to as an "error-prone" polymerase due to the incorporation of incorrect nucleotides (misincorporation)

which cannot be repaired at transcription, giving rise to mismatches in the growing double stranded molecule (Battula and Loeb, 1974 and 1976). Eukaryotic DNA polymerases however, have a much higher fidelity *in vivo* even though they too lack the 3'-5' exonuclease (Loeb and Kunkel, 1982).

The RNase H activity of reverse transcriptase degrades the RNA strand of RNA-DNA hybrids and was discovered by Mölling and co-workers (1971). Unlike cellular RNase H, it requires free 5' or 3' ends and acts as an exoribonuclease, generating oligoribonucleotides of 4 to 20 nucleotides in length (Verma, 1981 and references therein). The RNase H and DNA polymerase activities of reverse transcriptase are found on the same polypeptide, but have different active sites as seen by differing heat stabilities and times of inactivation (Grandgenett *et al.*, 1973 and references in Verma, 1981).

AMV reverse transcriptase has two subunits, α and β of molecular weights 68,000 and 92,000 respectively. The amino acid sequence of α is a subset of the β subunit and the mature form $\alpha\beta$ is believed to be formed by proteolytic cleavage of one subunit of a $\beta\beta$ dimer to give $\alpha\beta$ and p32 (Grandgenett *et al.*, 1980). The β subunit is phosphorylated (Schiff and Grandgenett, 1980). The α subunit contains both the DNA polymerase and RNase H activities that the $\alpha\beta$ holoenzyme contains (Grandgenett *et al.*, 1978). The DNA endonuclease nicks supercoiled DNA randomly in the presence of Mn^{2+} but nicks supercoiled *E. coli* ColE1 plasmid DNA preferentially near the *Eco* R1 site in the presence of Mg^{2+} suggesting some site specificity (Grandgenett *et al.*, 1978).

3.1 Uses of reverse transcriptase

3.1.1 Synthesis of cDNA.

Reverse transcriptase is most commonly used *in vitro* for synthesis of DNA complementary to mRNA. This was first employed in 1972 to produce rabbit globin cDNA (Ross *et al.*, 1972, Kacian *et al.*, 1972 and Verma *et al.*, 1972). The synthesis of cDNA can be used to map transcription initiation sites by reverse transcriptase-directed primer extension on isolated mRNA. Cloning of cDNA molecules is used in the construction of libraries using oligo dT as a primer to make large transcripts from isolated mRNA preparations. Efficient screening of recombinant genomic libraries for specific clones requires cDNA probes.

Cloning cDNA copies of genes usually entails synthesis of double-stranded DNA molecules by reverse transcriptase. Following transcription of a cDNA strand from mRNA, the RNA is removed by alkaline hydrolysis or the RNase H activity. Second strand synthesis requires a primer such as an oligonucleotide (Maniatis *et al.*, 1982), or a primer can be formed by foldback of the 3' end of the first strand cDNA. After opening up the hairpin loop with S1 nuclease, the double-stranded cDNA molecule can be cloned into a suitable vector using linkers or homopolymer tailing. Genomic DNA and cDNA sequences can be compared allowing mapping of introns and exons (Breathnach *et al.*, 1977). Cloned cDNA can be used for a variety of other purposes, e.g. for purification

of mRNA species or for hybrid-arrested translations.

3.1.2 End labelling.

Reverse transcriptase lacks both 3'-5' and 5'-3' exonucleases and so can be used for both selective end-labelling or for gap-filling. Gap-filling is where "sticky-ends" of a restriction fragment with 5' overhangs are filled in by the DNA-directed DNA polymerase activity of reverse transcriptase when all four dNTPs are provided. For selective end-labelling of a fragment which has been cut with two restriction enzymes, a single α-^{32}PdNTP and reverse transcriptase are provided. This may result in just one strand of the molecule becoming labelled at the 3' end if the other end has a 3' overhang, or a 5' overhang lacking a nucleotide complementary to the radioactive dNTP.

3.1.3 *In vitro* mutagenesis.

The lack of proofreading activity of reverse transcriptase has allowed its exploitation for mutagenesis *in vitro* (Kunkel *et al.*, 1979). It is more error-prone than *E. coli* DNA polymerase I and this aspect has been considered in Chapter 2.

Chapter 4. Isolation of mutants produced by mutagenesis *in vitro*.

In vitro mutagenesis is usually carried out for the purpose of isolating mutants which can then be used to further the study of the biochemical or genetical system being investigated. Mutants can be isolated either by selection or by screening. In the latter case, the process of mutagenesis can be studied simultaneously by analysis of the types and frequencies of mutants recovered, provided that no enrichment has been used and that no selective repair of mismatches occurs *in vivo*.

4.1 Selection of mutants.

Isolation of mutants is facilitated where the mutation results in a change of phenotype thus allowing selection of mutants. Zakour and co-workers (1984) studied misincorporation at the *am18* locus of ϕX174. Misincorporations produced viable wild-type revertants which were selected by their growth on host bacteria non-permissive for growth of the ϕX174 amber mutant.

Where a mutation creates an antibiotic resistant phenotype, such mutants can be selected by growth in the presence of that antibiotic. When such a selection procedure exists by virtue of an altered phenotype, then a low efficiency of mutagenesis is not a problem.

4.2 Screening for mutants.

Genotypic changes can be introduced that do not affect viability or growth rate under any conditions, in which case the progeny of the mutagenesis must be screened for mutants. For example, if a piece of foreign DNA is inserted into the multiple cloning site of one of the M13 mp vectors, then screening for the recombinant clones is by virtue of the blue to white colour change of the recombinant plaques. Likewise, if a mutagenesis event converts the phenotype from blue plaques to white plaques (or *vice versa*) then the mutants can be found by screening the progeny for a colour change. Where no phenotypic change occurs on mutagenesis, then the progeny have to be screened by other means to isolate mutants from the background of wild-type.

If mutations have been introduced by oligonucleotide mutagenesis then one can screen for mutants by the differential annealing of the mutant oligonucleotide which has been labelled radioactively. A perfectly matched oligonucleotide duplex has a higher melting temperature than a mismatched duplex. Annealing at low temperatures of the labelled mutagenic oligonucleotide (probe) to filters on which is bound DNA of progeny clones, followed by sequential washes at increasing temperatures to remove unbound probe, permits the detection of mutants because the probe DNA remains annealed to mutant DNA at temperatures above those where probe DNA dissociates from wild-type DNA. (Zoller and Smith, 1983).

The preferred method of mutant identification and characterization is DNA sequencing. This can only be the primary method of screening if candidate progeny produce SS DNA suitable for rapid preparation (Eperon, 1986a) and for DNA sequence determination by the "dideoxysequencing" method of Sanger *et al.*, (1977). Other methods of DNA sequencing such as Maxam and Gilbert's chemical method (Maxam and Gilbert, 1980) or dideoxysequencing of double-stranded DNA are too laborious for detection of rare mutants. Single track determination of DNA sequence allows positive identification of mutants when the substituted base is known. Rapid preparation of SS DNA from 96 clones (Eperon, 1986a) and rapid sequencing using microtitre plates instead of individual tubes allows screening of large numbers of potential mutants thus reducing the problems of detecting mutants when the efficiency of mutagenesis is low. The need to enrich the final yield of mutants by selection, and the need for effective screening procedures increases if less efficient methods of mutagenesis are adopted.

4.3 Enrichment of mutants.

4.3.1 Enrichment of mutants prior to transfection.

Where the efficiency of mutagenesis is low, as with random point mutagenesis over a large region by chemical methods or misincorporation, and where the mutations are phenotypically silent, then mutants will not be detected easily within the large background predominance of wild-type

progeny. If the background of non-mutated molecules can be reduced or eliminated then the mutants should form the majority of the resulting progeny. Myers *et al.* (1985a,b,c) have employed a method of enriching the products of a mutagenesis reaction for mutated molecules by purifying those duplexes which contain a mismatch, prior to transfection into *E. coli*. When such a mismatch is present within a double stranded fragment this alters the strand dissociation of such a fragment. Duplexes containing a mismatch were separated from wild-type duplexes by the different positions of their partial denaturation on gradient gel electrophoresis (Fischer and Lerman, 1985).

Where oligonucleotide mutagenesis results in the removal of a unique restriction cleavage site, digestion of the closed-circular double-stranded (CC) DNA prior to transfection linearises those duplexes which do not contain the mismatch oligonucleotide in the newly synthesized strand. This method can be used to enrich for mutant CC DNA which, on transfection into *E. coli*, will give increased yields of mutants (I.C. Eperon, unpublished).

Strategies for enrichment of mutants have also employed methods whereby the progeny of a mutagenesis reaction preferentially derive from the newly synthesized mutant strand. *In vitro* strand separation of the mutant strand from the wild-type strand has been used by Abarzua and Marians (1984). This method involved the *in vitro* replication of the replicative form (RF) heteroduplex to SS (+) strand circular DNA, using purified *E. coli*

replication proteins, initiated from the ϕX174 origin present on the vector DNA. The purified SS circular DNA was complementary to the newly synthesized strand (the (-) strand) containing the introduced mutation.

Selective cleavage of the template strand after formation of a mutant heteroduplex can be effected by incorporation of α-SdNTPs in the newly synthesized strand. Restriction enzymes preferentially cleave the template strand containing phosphodiester bonds before cleaving strands containing phosphorothioate linkages (Eckstein, 1985). The template strand can be further digested by exonuclease III before transfection when progeny are produced preferentially from the mutant strand.

4.3.2 Enrichment of mutants on transfection (figure 2).

Another method of enrichment utilised the fact that the sequence of methylated DNA is preserved in preference to non-methylated DNA when subject to repair mechanisms *in vivo* (Pukkila *et al.*, 1983), see Chapter 5. The gapped-duplex method was used with the target template strand hypomethylated and the gapped-strand methylated (Kramer *et al.*, 1982; Marmenout *et al.*, 1984). The rationale behind the method was that on annealing a mismatch oligonucleotide within the gap, the template strand would be repaired in preference to the mutant strand. The practicalities of the method are presented by Fritz (1985), also see figure 2.

Kramer and co-workers (1984a and b) used the

Figure 2.
Enrichment of M13 point mutants on transfection.

A Avoidance of the methyl-directed repair of *E. coli*.

AI Point mutagenesis with a methylated template produces a mutant heteroduplex with a non-methylated mutant strand. Transfection into a strain of *E. coli* deficient in the Mut repair system avoids the strand bias of *in vivo* repair.

AII Point mutagenesis with a non-methylated template produces a mutant non-methylated heteroduplex. Transfection into *E. coli*, whether JM101 or a *mut* strain does not give biased repair of the mutant strand.

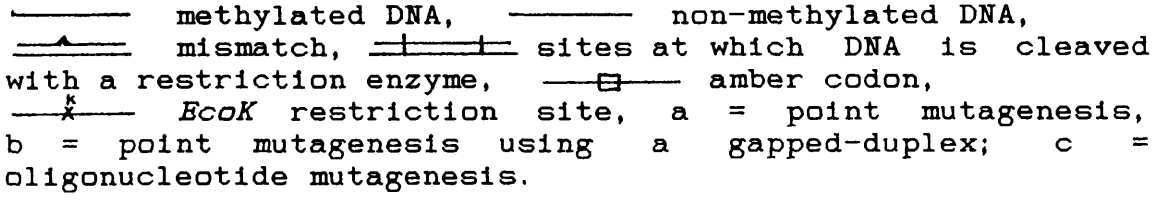
B Exploitation of the methyl-directed repair of *E. coli* (Kramer *et al.*, 1982).
Gapped-duplex point mutagenesis with a non-methylated template strand and a methylated gapped strand. Transfection into a strain of *E. coli* wild-type for the Mut repair system gives biased "repair" of the non-methylated template strand.

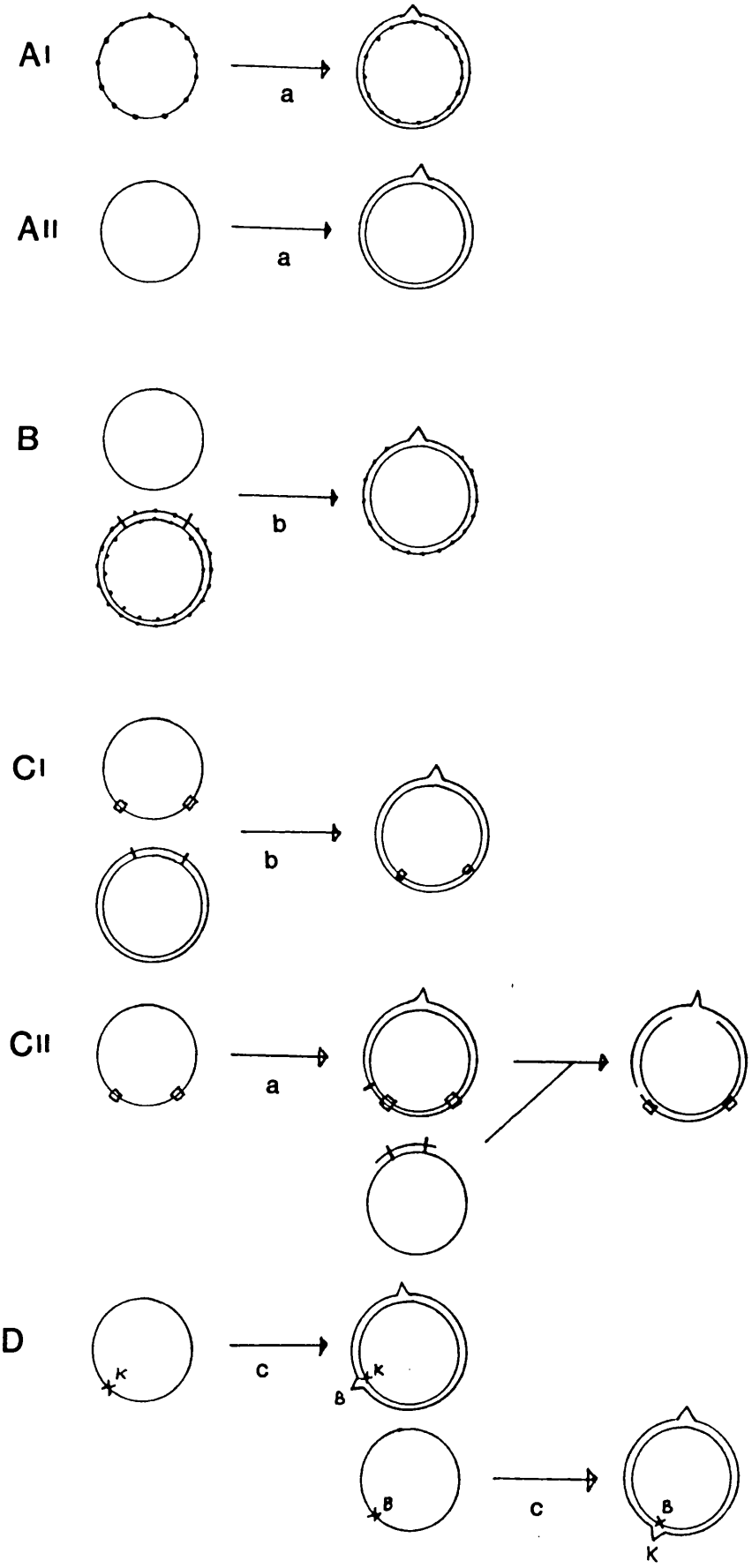
C Exploitation of amber codons.

CI (Kramer *et al.*, 1984a) Point mutagenesis with a gapped duplex molecule using an amber template strand and a wild-type gapped-strand. Transfection into wild-type *E. coli* only allows replication of the mutant (non-amber) strand.

CII (Eperon unpublished). Point mutagenesis with an amber template strand followed by linearisation of the CC DNA and annealing to gapped, SS wild-type DNA. Transfection into wild-type *E. coli* allows replication of only the wild-type strand but the introduced mutation is copied faithfully *in vivo* into the wild-type strand.

D (Carter *et al.*, 1985) Exploitation of restriction sites.
Oligonucleotide mutagenesis with two oligonucleotides introduces a desired mismatch converting an *EcoK* restriction site to an *EcoB* site. Transfection into an *E. coli mut* $r_K^+m_K^+$ strain allows replication of only the mutant strand. A further round of mutagenesis employs a different selection oligonucleotide converting an *EcoB* restriction site to an *EcoK* site and transfection into an *E. coli mut* $r_B^+m_B^+$ strain allows replication of only the new mutant strand.




AI
a
AII
a
B
b
CI
b
CII
a
D
K
c
K
B
B
c
B
K

gapped-duplex method with the template strand carrying two amber mutations in the M13 vector sequences, and the gapped strand being wild-type. The gapped strand acquired the mismatch following annealing of the oligonucleotide to the template, and gap repair. When the heteroduplex was transfected into a host carrying no amber supressor progeny could arise only from this strand. Kramer and co-workers (1984b) also showed that use of a host deficient in DNA repair enhanced yields.

Eperon (unpublished results) has modified the rationale of the gapped duplex/amber selection procedure for use after standard oligonucleotide mutagenesis, figure 2. The CC heteroduplex consists of both strands bearing amber mutations and is linearised at a unique site outside the multiple cloning site. This linear heteroduplex is then annealed to SS non-amber DNA, previously linearised at two sites to exclude the multiple cloning site. On transfection into a host containing no amber suppressor, the mutated strand is copied faithfully into the gap on the wild-type strand so no heteroduplex is available for repair, and only this wild-type strand gives rise to progeny. This method has produced up to 70% mutant recovery.

Enrichment of mutants on transfection into a chosen *E. coli* strain has been effected by Carter *et al.* (1985), figure 2. They have used two oligonucleotide primers for mismatch mutagenesis whereby the first primer contained the desired mismatch, and the second primer was used to eliminate a restriction marker found on the template strand.

The restriction markers in the vectors were either *Eco* K or *Eco* B, and the newly synthesized mutant strand was selected for by growth in *E. coli* K or B strains respectively. Multiple rounds of mutagenesis are described, reciprocating between *Eco* K and *Eco* B selection.

Even with selection of the mutant strand by either the two-primer method or a gapped-duplex method, the recovery of mutants is never 100%. Reasons for this include:

i) the two primers might not always both anneal to the same template strand,

ii) with the gapped-duplex method, some newly synthesized strands may not contain the mutation if the mutant oligonucleotide did not anneal to all the gaps,

iii) repair of DNA mismatches *in vivo*.

4.4 Summary.

It is apparent that the recovery of mutants generated by mutagenesis *in vitro* is complicated by the presence of wild-type clones. Thus the recovery of mutants can be increased by enrichment or selection and their identification and characterization are facilitated by effective screening procedures. An assessment of the results of mutagenesis with regard to the efficiencies of different mutation reactions *in vitro* must include consideration of the factors affecting the recovery of mutants. One of these factors, *in vivo* repair of DNA mismatches, is discussed in the next chapter.

CHAPTER 5. Repair of DNA mismatches in *E. coli*.

Mutagenesis of DNA *in vitro* by chemicals, incorporation of dNTP analogues or misincorporation, introduces changes in the sequence of one strand of DNA, creating a DNA heteroduplex. Mutations introduced chemically are not subject to repair *in vitro*, and repair of misincorporated dNTPs during second strand synthesis *in vitro* can be avoided by judicious use of alternative divalent cations in the reaction, biased pools of dNTPs, α-SdNTPs or AMV reverse transcriptase. However, on introduction of a DNA heteroduplex molecule into *E. coli*, the *in vivo* repair functions of *E. coli* can recognise mismatches and repair one of the strands. The mechanism of strand selection (described below) results in the preferential repair of the newly synthesized, non-methylated DNA strand.

Repair of mismatches before or on replication of the DNA heteroduplex will reduce the yield of mutants obtained from a mutagenesis reaction. Such repair *in vivo* can mask the efficiences of mutagenesis. An understanding of the repair mechanisms involved *in vivo* will allow them to be avoided or exploited in order to recover more mutants.

5.1 The observation of repair of DNA heteroduplexes *in vivo*.

In the process of misincorporation mutagenesis, a mismatched base-pair is introduced into a newly synthesized DNA molecule and the heteroduplex DNA is then introduced

into *E. coli* by transfection. In theory there are eight different, single-base mismatches that can occur in the DNA heteroduplex, these being either transition mismatches (A:C or G:T) or transversion mismatches (A:A, A:G, C:C, C:T, G:G or T:T). The adding and editing functions of DNA polymerase III give an error rate of 10^{-7} to 10^{-6} (Loeb and Kunkel, 1982) and heteroduplex base mismatch correction of newly synthesized DNA gives a further reduction of 10^{-2} to 10^{-3} (Cox, 1976) thus resulting in an overall mutation frequency in *E. coli* of 10^{-8} to 10^{-10} per nucleotide per generation.

Correction of heteroduplex DNA in *E. coli* has been reported with heteroduplexes of bacteriophage: lambda (Wildenberg and Meselson, 1975), ϕX174 (Baas and Jansz, 1972), f1 (Enea *et al.*, 1975), T7 (Bauer *et al.*, 1981) and and M13 (Kramer *et al.*, 1982). These results have implicated the Mut (or methyl-directed) system of *E. coli* mismatch repair, and in particular the *mutH*, *mutL*, *mutS* and *mutU* gene products (Rydberg, 1978). Mutant strains of *E. coli* have been shown to have high mutability with 5-bromouracil (5-Bu); the defect mapped close to the mutator loci *mutH* (*mutR*), *mutS* and *mutU* (Rydberg, 1977 and 1978). *E. coli* *mutL*, *mutH* and *mutS* strains have been reported to be blocked in repair of lambda heteroduplexes (Rydberg, 1978), and *mutL* and *mutS* strains have been shown to be blocked in repair of M13 heteroduplexes (Kramer *et al.*, 1982).

The Mut repair system was deduced to operate by some kind of strand targeting for repair of heteroduplexes because there appeared to be preferential repair of the

incorrect base in the newly synthesized strand. Wagner and Meselson (1976) proposed that this strand recognition was by lack of methylation in the newly synthesized strand. This proposal was based on the observation by Dunn and Smith (1958) that *E. coli* DNA contains methylated bases, and that this methylation occurs shortly after DNA synthesis *in vivo* rather than concurrently (Gold *et al.*, 1963, Srinivasan and Borek, 1964). *E. coli dam* mutants lack dam methylase activity and so do not methylate the DNA after synthesis. Such mutants have been shown to have spontaneous mutation rates of 10 to 100 times higher than those of Dam^+ strains due to loss of specificity of strand recognition for repair (Marinus and Norris, 1974; Glickman *et al.*, 1978). The effect of non-methylation of DNA was shown directly in the preferential repair of the unmethylated strand in hemimethylated heteroduplexes (Radman *et al.*, 1980 and 1981, Pukkila *et al.*, 1983). The *in vivo* repair system of *E. coli* has no strand specificity when the DNA is hypomethylated, *ie.* that from Dam^- strains of *E. coli*; either or both strands can be cleaved and repaired, which can result in either wild-type or mutant progeny. When both strands of chromosomal DNA are cut this can lead to cell death; hence, dam^- strains are very sensitive to mutagens. Mutations in *mutH*, *mutL* or *mutS* in Dam^- cells have been shown to suppress sensitivity of the cells to mutagens (Glickman and Radman, 1980; Jones and Wagner, 1981). Conversely, if the cell has excessive methylase activity, causing hypermethylation, then there is no chance for repair of heteroduplexes before the

nascent strand is methylated and thus the spontaneous mutation rate is increased (Marinus, 1984).

5.2 Heteroduplex repair of M13 phage in *E. coli*

Kramer and co-workers (1984b) transfected methylated, hemimethylated and unmethylated M13 DNA into *E. coli* and demonstrated that hemimethylation at GATC sites can direct the mismatch repair on the unmethylated strand. Unmethylated M13 DNA was also repaired but showed no strand bias. This repair was abolished in *mutL* and *mutS* strains, and *mutH* strains showed no methyl-directed strand bias.

5.2.1 Strand recognition

Lacks and co-workers (1982) proposed that the mismatch repair in *E. coli* be mediated by SS breaks in the DNA acting as a signal for recognition of which strand to repair. This was based upon a comparison with the Hex system of repair of *S. pneumoniae*, in which SS breaks are the signals for strand recognition, as the two systems appear to be quite similar (apart from the methyl-directed strand repair of Mut). It is proposed that an *E. coli* nicking enzyme (which has not as yet been discovered) is responsible for creating SS breaks in hemimethylated or unmethylated DNA; such breaks have been seen in *dam* strains carrying a *lig* mutation reducing ligation (Marinus and Norris, 1974). The nicking enzyme may have arisen from an endonuclease

which, through evolution, has lost its ability to make a double-stranded break at an unmethylated GATC site but can still make a single-stranded cleavage in an unmethylated strand. Claverys and Lacks (1986) suggest that the *mutH* product is responsible for SS breaks at unmethylated or hemimethylated GATC sites, but that other SS gaps can substitute for such GATC sites. This would account for the loss of methyl-directed strand bias seen by Kramer and co-workers in *mutH* cells when using M13 heteroduplexes (Kramer *et al.*, 1984b). The M13 heteroduplexes used in their investigations already contained one SS gap because the material for transfection was not in CC form, and another gap could have been introduced into either strand on entry of the material into the cell. The region between two single-stranded breaks would thus be repaired with no bias to the unmethylated strand. Claverys and Lacks (1986) consider that *mutH* may encode the nicking enzyme which may be a monomer of an ancestral restriction enzyme dimer.

5.2.2 Mismatch repair specificities.

The repair of mismatches in heteroduplexes has been investigated in order to establish whether any particular mismatches are more refractory to repair than others. All transition mismatches appear to be recognised efficiently in M13 (Kramer *et al.*, 1984b), f1 (Lu et al., 1983) and lambda (Wagner *et al.*, 1984). Transversion mutation repair has been investigated in M13 (Kramer *et al.*, 1984b) and in lambda

(Wagner *et al.*, 1984). G:G mutations were repaired as efficiently as the transition mutations, and A:A mutations were repaired with medium efficiency in both cases. T:T mismatches were repaired efficiently in lambda but not at all in M13, and C:C mismatches were repaired with low efficiency in lambda but not at all in M13. T:C and G:A transversion mismatches were not repaired at all.

E. coli does have a repair mechanism for A:G mutations when A is in the template, this occurs at the time of synthesis and requires the *mutT* gene product (Yanofsky *et al.*, 1966). Claverys and Lacks (1986) predict the presence of a similar repair mechanism for repair of the same mismatch when the G is in the template strand.

The Mut repair system also recognises deletions and insertions of one or two bases, but those of three to five nucleotides are not recognised.

G:T mismatches can occur *in vivo* by deamination of 5-methyl cytosine to T and when this occurs the *mutS* gene product is involved in short patch repair. This can act anywhere in the DNA and does not require either SS breaks or undermethylation (Claverys and Lacks, 1986).

5.3 *In vitro* mutagenesis and *in vivo* repair.

Where site directed mutagenesis is used to obtain desired mutations, the processes of *in vivo* repair in *E. coli* can be detrimental, reducing yields. Clearly fidelity of replication *in vivo* is necessary in order to

avoid the introduction of further mutations *in vivo* but preferential repair of introduced mismatches will limit the recovery of the specified mutations. Kramer and co-workers (1984a and b) have demonstrated that the use of Mut^- host strains prevented repair of mismatches introduced by the gapped-duplex method. However, Kramer and co-workers (1982) have also exploited the Mut repair system to preferentially "repair" the wild-type strand of a heteroduplex. This involved the use of a methylated gapped strand into which the mismatch could be introduced, and an unmethylated intact circular single strand which had been propagated in a *dam* host strain. The ligated heteroduplex was introduced into a Mut^+ recipient allowing preferential repair of the wild-type strand.

An alternative method was adopted by Kunkel (1985) who, instead of using an unmethylated wild-type template strand, used one which had been grown in an *ung* host and so contained uracil molecules. After mutagenesis the heteroduplex was introduced into a wild-type (Ung^+) recipient. The *ung* gene of *E. coli* encodes uracil glycosylase which removes uracil bases in a DNA strand so creating abasic sites. When transfected into wild-type *E. coli*, uracil-containing SS DNA gives very low yields of plaques: 5 orders of magnitude lower than when transfected into *ung E. coli*. The heteroduplexes of uracil-containing template DNA and newly synthesized DNA (containing a mutation but no uracil) suffered preferential destruction of the template strand.

A fuller understanding of the mechanisms of *in vivo* repair would allow a greater recovery of mutants following *in vitro* mutagenesis. At present the processes of repair appear to give unequal recovery of different mutations presenting problems in the analysis of the mechanisms of *in vitro* mutagenesis and the action of DNA polymerases.

CHAPTER 6. Analysis of misincorporation reactions.

Most analyses of misincorporation have been carried out by studying the number and type of mutants recovered from a mutagenesis experiment. Where the recovery of mutants is by selection of an altered phenotype it is not possible to analyse the mechanism of misincorporation. Some of the mutations effected by the mutagenesis might be phenotypically silent and so would not be selected.

Screening for mutations by sequencing a proportion or all of the resulting progeny gives a true representation of the different mutations recovered *in vivo*. However, such an analysis can, by definition, only consider those mutations which manifest themselves in the progeny of mutagenesis. Detailed analysis of the misincorporation reaction is therefore not possible by only considering isolated mutants, although some investigators have attempted to do so. In their investigations they have failed to consider a number of relevant factors detailed below.

a) The processes of *in vivo* repair of DNA heteroduplexes are not fully understood (see Chapter 5). Selective repair of certain mismatches by *E. coli* repair systems would mask the specificities of *in vitro* misincorporation.

b) Experiments described in Chapter 2 investigating the misincorporation of dNTPs at ϕX174 amber sites have been limited in the variety of combinations of the template nucleotide, the substrate dNTP and the primer 3'-terminal nucleotide, thus one cannot determine which, if any, of the

combinations are more favourable than others.

c) Experiments where the sites of misincorporation have been spread throughout a gene have used primers generated by exonuclease III digestion. Such primers may be unevenly spread along the template DNA giving an unequal spread of 3' termini and next template nucleotides.

d) Where the mutants have been isolated by screening for a phenotypic change, the proportions of different mutations recovered may not have been representative of the initial misincorporation reaction as some mutations will have given no change in phenotype. Thus no real conclusions can be drawn about the relative efficiencies of particular misincorporations. Multiple mutations may also have occurred but would not have been recovered by the phenotypic screen.

Analyses of the mechanisms of misincorporation mutagenesis would thus be better performed at the initial *in vitro* stage.

Studies of the misincorporation reaction by DNA polymerases have generally been limited in scope. The methods used have involved either synthetic polynucleotides as templates, or natural DNA, and are described in more detail below.

6.1 Synthetic polynucleotide templates.

Synthetic polynucleotides have been useful for *in vitro* assays, and because of their predefined composition have allowed clear distinction between correctly

incorporated and misincorporated dNTPs. Such assays measured incorporation of radioactively labelled complementary or non-complementary dNTPs (Hall and Lehman, 1968; Battula and Loeb, 1974) from which the error rate of the polymerase was calculated. Synthetic polynucleotide templates have allowed a kinetic analysis of stable misincorporations and of turnover of nucleotides (generation of dNMP from dNTP) (Fersht *et al.*, 1983), but they do suffer from a number of limitations. Homopolymer template-primers cannot be used to look at incorporation of all four possible dNTPs at all of the 16 possible sites, defined by different combinations of the 3' terminus of the primer and the next template nucleotide. Homopolymers also differ from natural DNA in their secondary structure and stacking interactions and this may result in altered accuracy of DNA polymerases.

6.2 Natural DNA templates

Assays with natural DNA have been based mainly upon the genetic approach of recovery of mutants *in vivo* after mutation of ϕX174 amber codons by a single base substitution (Loeb and Kunkel, 1982). Such assays are not just a simple measure of misincorporation due to additional factors described earlier. Furthermore, the assays based on production of mutants have perforce been very limited, analysing few combinations of template nucleotide and primer terminus, even though studies with T4 polymerase (Topal *et al.*, 1980) suggested that the efficiency of misincor-

poration may vary depending on the 3' terminal nucleotide of the primer.

Natural DNA has also been used for an *in vitro* study of misincorporation at a number of sites using PAGE analysis. Three of the four dNTPs were provided in each reaction and the patterns of primer extensions were compared to dideoxysequencing tracks (Hillebrand *et al.*, 1984). A further study by Hillebrand and Beattie (1985) considered the effect on misincorporation of the primary and secondary structure of the template. However, in neither study were the misincorporations verified. A change in mobility on PAGE analysis was suggested as an indication that misincorporation had occurred but data showing this were not presented. The substrate dNTP stocks had been purified by HPLC but no control reactions were presented of incubations with DNA polymerase I, AMV reverse transcriptase or T4 DNA polymerase in the absence of added dNTPs. Thus, the possibility of low levels of contaminating dNTPs from other sources could not be excluded. The number of mismatches was not determined. Runs of multiple misincorporations must have occurred in some cases but as the misincorporations were not characterized, it was impossible to assess which particular misincorporations were favoured.

None of the *in vitro* analyses of misincorporation have been able to demonstrate which particular misincorporations of all the possible transitions and transversions were polymerised most efficiently. More information on this would allow a clearer understanding of the action of DNA

polymerases in their misincorporation of dNTPs and hence their fidelity.

CHAPTER 7. Summary

There exists a need for an efficient method of *in vitro* mutagenesis that can create a library of mutant molecules each with a single, different point-mutation such that every possible mutation in the DNA sequence is represented within that library. At present no single method of mutagenesis *in vitro* is ideally suited to this purpose.

Chemical mutagenesis can be used to spread mutations throughout a chosen region, but only certain mutations can be induced by each chemical method. Furthermore, when the region for mutagenesis is large, the efficiency of the reaction has to be kept low to limit mutations to one per molecule. This makes the task of screening for mutants more difficult.

In the development of a method of misincorporation mutagenesis that would generate a library of single point mutants a number of different criteria have to be satisfied:
1) Misincorporation should be efficient so that as high a number of molecules as possible contain a mismatch.
1i) The efficiency of misincorporation must be sufficient to enable easy screening for mutants by DNA sequencing.
1ii) It is important to know whether the DNA polymerase can carry out all possible misincorporations and, if some reactions are more favourable than others, to know what factors control this. This requires an understanding of the mechanism of misincorporation.
2) A single misincorporation should allow further correct

incorporation to proceed unhindered.

3) No more misincorporation should occur after the initial misincorporation.

Recent investigations into the misincorporation of dNTPs by DNA polymerases (see Chapter 2) have succeeded in isolating point mutants spread thoughout a gene using primers generated by exonuclease III digestion and misincorporation of α-SdNTPs by DNA polymerase I (Abarzua and Marians, 1984; Shortle and Lin, 1985). Such investigations have thus failed to consider important factors detailed in Chapter 6 and so have been unable to satisfy criteria 1 and 3 (listed above) by demonstrating the specificities of misincorporation within the reaction.

In vitro analyses of misincorporation reactions have, in general, been limited to the use of synthetic DNA primer-templates (see Chapter 6). These studies have not permitted the investigation of the effects of all different combinations of primer 3' termini, template nucleotide and substrate dNTPs and thus have not satisfied criterion 1ii, the determination of factors that control the different efficiences of misincorporation. Furthermore, the different secondary structure stacking interactions of homopolymer primer-templates do not reflect the situation with natural DNA.

Where *in vitro* analyses of misincorporation have used natural DNA primer-templates and PAGE analysis (Hillebrand *et al.*, 1984), some questions still remain unanswered: the analysis of misincorporation did not allow the

identification of which of three possible dNTPs was misincorporated at any one template position, furthermore, no systematic study of the misincorporation reaction was carried out with different combinations of the site of misincorporation (primer 3' terminal nucleotide and template nucleotide) and the substrate dNTP (criterion 1ii). Moreover, multiple misincorporations cannot be excluded (criterion 3).

The work presented in this thesis is a systematic study of reverse transcriptase-directed misincorporation of specified dNTPs against different template nucleotides to determine whether reverse transcriptase exhibits misincorporation preferences and to satisfy criterion 1ii. An assay is described in which all possible combinations of template nucleotide, substrate dNTP and the 3' terminus of the primer were investigated and the products of the individual misincorporation reactions were analysed directly using PAGE in order to satisfy criteria 2 and 3.

The *in vivo* analysis of some mutations recovered after screening by dideoxysequencing (criterion 1i) is compared with the *in vitro* analysis and discussed in the light of *E. coli* repair of DNA heteroduplexes.

Results are presented that show the generation within one reaction of a range of primers with a common 5' terminus and differing 3' termini (suitable for subsequent misincorporation) of which the majority terminated prior to next-template nucleotides specified within the primer-extension reaction. Such a spread of termini (analysed by

PAGE) allowed for subsequent misincorporation at a number of single sites along the template, a prerequisite of the proposed mutagenesis methodology for the production of a library of mutants.

Many questions still remain unanswered about the structure and function of the mammalian *ras* protein and these require comprehensive mutagenesis of the gene in a way that would alter each single amino acid of the protein independently. This would shed light upon which amino acids are involved at effector sites in protein-protein interactions. Such results combined with a structural map of the protein should enhance our understanding of the role of different domains of the *ras* protein. A library of point mutations of the *ras* gene could be generated by the application of an efficient reverse transcriptase-directed misincorporation methodology to the *ras* gene, using a suitable bank of extended primers to define sites for misincorporation. Thus the method of *in vitro* mutagenesis described here could be applied to protein engineering for the elucidation of protein folding patterns and for the structural and functional analysis of proteins.

CHAPTER 8. MATERIALS AND METHODS

8.1 Buffers

TE(0.1) 10mM Tris-HCl pH7.5, 0.1mM EDTA

8.1.1 Reaction buffers

10xH 100mM Tris-HCl pH7.5, 500mM NaCl, 100mM $MgCl_2$, 10mM DTT

10xC 500mM Tris-HCl pH7.5, 100mM $MgCl_2$, 10mM DTT

10xSeq 100mM Tris-HCl pH7.5, 50mM $MgCl_2$

10xRT 500mM Tris-HCl pH8.3, 60mM $MgCl_2$, 400mM KCl, 10mM DTT

10xExo 600mM Tris-HCl pH8.0, 6.6mM $MgCl_2$, 10mM DTT

5xTMS 200mM Tris-HCl pH7.5, 30mM $MgCl_2$, 10mM spermidine

8.1.2 Electrophoresis buffers

Acrylamide gels contain and are run in 1xTBE:

10xTBE 0.89M Tris-borate, 0.89M boric acid, 0.02M EDTA [per litre: 108g Trizma base, 55g boric acid, 9.3g EDTA (pH8.3)]

Agarose gels contain and are run in 1xE:

50xE 2M Tris-acetate, 0.05M EDTA [per litre: 242g Trizma base, 57.1ml glacial acetic acid, 100ml 0.5M EDTA (pH8.0)]

8.1.3 Electrophoresis dyes

Acrylamide gels: 90% deionised formamide, 50mM EDTA pH8.5, 0.25% xylene cyanol and 0.25% bromophenol blue.

Agarose gels: E dye (6x): 6xE buffer, 30% glycerol, 0.25% xylene cyanol and 0.25% bromophenol blue.

8.2 Bacterial growth media

8.2.1 Broth

TY broth: per litre: 16g Difco bacto tryptone, 10g Difco bacto yeast extract, 5g NaCl

Standard broth(SB): per litre: 10g Difco bacto tryptone, 5g Difco bacto yeast extract, 5g NaCl

8.2.2 Agar

Glucose minimal agar (for maintenance of *E. coli* JM101)

90ml distilled water, 1.5g agar (autoclaved)
10ml 10xM9 salts "
0.1ml 1M $MgCl_2$ "
0.1ml 100mM $CaCl_2$ "
1ml 40% glucose "
0.2ml thiamine (1mg ml^{-1}, filter-sterilized)
0.1ml 10mM $FeCl_3$ "
0.2ml biotin " "

Agar plates were made by supplementing broth with 1.5% Difco bacto agar;

Top (or soft) agar was 0.8% Difco bacto agar in broth.

8.3 Strains of *Escherichia coli*

JM101 *supE, thi, Δ(lac pro)*, F' *traD36, proAB, lacI^{q}ZΔM15*
(Yanisch-Perron *et al.*, 1985)

JM109 *recA1, endA1, gyrA96, thi, hsdR17, supE44, relA1, λ$^-$, Δ(lac pro)*, F' *traD36, proAB, lacI^{q}ZΔM15*
(Yanisch-Perron *et al.*, 1985)

HB101 F$^-$, *hsdS20*($r_B^-m_B^-$), *recA13, ara-14, proA2, lacY1, galK2, rpsL20*(Sm^r) *xyl-5, mtl-1, supE44, λ$^-$*
(Maniatis *et al.*, 1982)

BMH71.18 *mutL* K12, *Δ(lac pro), supE, thi*/F', *proAB, lacI^{q}ZΔM15, mutL::Tn10*
(*Kramer et al.*, 1984b)

8.4 DNA

DNA stocks were stored at -20°C in TE(0.1)

8.4.1 Bacteriophage M13 DNA

mp8 (Messing, 1983)

C6.9 30bp *Bam* HI fragment of U1 gene cloned into the *Bam* HI site of mp9. (Eperon unpublished) Gift of Dr. I.C. Eperon

mICE 10 (Eperon, 1986b) Gift of Dr. I.C. Eperon

mICE 11 (Eperon, 1986b) Gift of Dr. I.C. Eperon

M-*ras* (mEJ-*ras* 2.9:3) 2.9kb *Sac* I fragment of J180 cloned into *Sst* I site of mICE 10 (See Results).

8.4.2 Plasmid DNA

J841 (c-Ha-*ras* 1) (Tabin *et al.*, 1982) Gift of Dr. E. Chang.

J180 (pEJ c-Ha-*ras* 1) (Tabin *et al.*, 1982) Gift of Dr. E. Chang.

8.4.3 Oligonucleotide primers

		Source
Universal primer:	5' GTAAAACGACGGCCAGT3'	Pharmacia
Primers 1-16:	(See Table 1)	J. Keyte
Primer 24:	5'CAGGTCGCGGATCCCC3'	J. Keyte
Primer 42	5'CGTAATCGATGGTCAT3'	J. Keyte
Primer 51 (*ras*)	5'CATCCCTCCTTTCCCAG3'	J. Keyte

8.5 Phenol extraction of DNA solutions.

Phenol was either freshly distilled (Fisons) or HPLC grade (BDH) and was equilibrated twice with 100mM Tris-HCl pH7.5 and then with TE (0.1) and stored at -20°C. DNA solutions were phenol extracted to remove contaminating proteins. In general, volumes were made up to > 50µl and were phenol extracted by addition of half a volume of phenol. The DNA solution/phenol mixture was vortexed and then centrifuged either in a microfuge (12,000g) for five minutes or in corex tubes in an MSE centrifuge at 8,000 rpm for ten minutes. The aqueous (upper) phase was removed and re-extracted with phenol and then with phenol, chloroform, isoamyl alcohol (25:24:1). Sometimes the first organic (lower) phase was back-extracted with an equal volume of TE (0.1). Phenol extraction was repeated until any protein interphase disappeared. Residual phenol was removed by three ether extractions, and the excess ether was driven off by placing the open tubes in a 55°C water bath for 15 minutes.

8.6 Ethanol precipitation of DNA.

DNA in aqueous solution was precipitated by adding a one tenth volume of 3M sodium acetate and 2-2.5 volumes of ethanol. The solution was mixed well and placed at -20°C overnight or in dry ice until frozen (15-30 minutes). The

DNA precipitate was centrifuged in a microfuge (12,000g) for 10 minutes or in an MSE centrifuge in corex tubes at 8,000 rpm for 20 minutes. Ethanol precipitates in microtitre plates were centrifuged in an IEC Centra 4X centrifuge at 2,000 rpm for 30 minutes. The supernatant was aspirated and the pellet washed in ethanol before drying under vacuum. When very small amounts of DNA were to be precipitated then carrier DNA or $tRNA^{Phe}$ were added to aid precipitation.

8.7 Electrophoresis

8.7.1 Polyacrylamide gels

8.7.1.1 Acrylamide preparation (40%)

380g acrylamide, 20g N,N'-methylene bisacrylamide made up to 1l with deionised water. Deionised with Amberlite MB-1 resin (BDH) for 1h and then filtered through Whatman 3MM paper and stored at 4°C.

8.7.1.2 Denaturing gel preparation (containing 7M urea).

For an x% acrylamide gel (up to 22%):

xml 40% acrylamide, 16.8g urea, 4ml 10xTBE and deionised water up to 40ml. Cover and stir until dissolved then add 320µl 10% ammonium persulphate (made monthly and stored at 4°C) and 30µl TEMED (N,N,N',N'-tetramethylene diamine, Serva). Mix gently before pouring between gel plates (20cm x 40cm x 0.38mm), separated by spacers (0.38mm thick Plasticard™) along the sides and taped

together with white Sellotape™. Gels were clamped together until they had polymerised (about 1h). Gels of ≥20% were used the same day otherwise gels could be stored, wrapped in Saranwrap™ up to 2 days. Before use, the bottom of the gel was untaped and the slots were flushed out with 1xTBE. Gels were normally run at 1500-2000V until the bromophenol blue reached the bottom of the gel (gels of ≥20% were also prerun for 30min before loading the samples.)

8.7.1.3 Sample preparation

Samples were mixed with formamide dye (¼vol) and heated, in open tubes, at 80°C for 6 to 10 min to denature the DNA and to reduce the volume. The denatured samples were loaded using a drawn out capillary tube.

8.7.1.4 Gel handling

High percentage acrylamide (≥10%) gels were transferred to the shiny side of Benchkote™, covered in Saranwrap™ and exposed to X-ray film (Fuji RX) at -70°C. Lower percentage gels could be fixed by soaking the gel in 10% glacial acetic acid, 10% ethanol (or industrial methylated spirits) for 20min. Gels were then transferred to Whatman 3MM paper, covered in Saranwrap™ and dried at 80°C under vacuum before being exposed to X-ray film at -70°C or at room temperature. Gels

containing material labelled with [^{35}S] had to be fixed and dried, and the Saranwrap™ was removed before exposing the gel to X-ray film.

8.7.2 Agarose gels

8.7.2.1 Gel preparation

Agarose gels were used at 0.6 to 1.2% in 1xE buffer. Agarose was melted in 1xE buffer then ethidium bromide (EtBr) was added to 50μg ml^{-1} before the gel was poured. Gel plates were taped around the edges with white Sellotape™. BRL Baby Gel tanks were used routinely and required 20ml of gel mix, the gels setting in about 1h.

8.7.2.2 Sample preparation

Samples in 1xE dye were loaded with a drawn-out capillary tube and were subjected to electrophoresis at 12.5mA for 2 to 3 h (Baby Gel). When samples contained SDS, EtBr was not included in the gel but the gel was stained, in 0.5mg ml^{-1} EtBr for 10min, after electrophoresis.

8.7.2.3 Gel handling

DNA bands were observed using an ultraviolet transilluminator or a hand-held ultraviolet lamp and similarly photographed using Polaroid 57 film with the following conditions:

Transilluminator f4.5/0.5sec

Hand-held lamp f4.5/10 to 20sec

Polaroid 55 film, used to obtain negatives as well as positives, requires exposures 60 times longer than does Polaroid 57.

8.7.2.4 Low Melting Point (LMP) Agarose

LMP agarose gels were used when DNA bands were to be excised and extracted from the gels. The gels (BRL ultrapure LMP agarose) were poured at 4°C, covered loosely with Saranwrap™ and allowed to set overnight or for at least two hours. Bands to be excised were observed only using a hand held long wave ultraviolet lamp, to reduce damage to the DNA caused by ultraviolet light.

8.7.2.5 Extraction of DNA from LMP agarose gels

Gel slices were cut as small as possible with a clean scalpel blade and placed into Eppendorf tubes. Initially the DNA was extracted by hot phenol extraction, but latterly the gel slice was diluted in water, melted and used directly for transformations.

8.7.2.5.1 Hot phenol extraction. An equal volume of TE(0.1) was added to the gel slice which was then incubated at 65°C for 15 minutes. An equal volume of phenol was added whilst the tube was still at 65°C,

mixed well with a Gilson P1000 blue tip, and vortexed briefly before leaving at room temperature for 5 minutes. The solution was vortexed once again and centrifuged in a microfuge at 12,000g for five minutes. The aqueous phase was removed and further extracted with phenol and then with phenol, chloroform, isoamyl alcohol until no interphase remained. The aqueous phase was extracted three times with ether to remove any remaining phenol; remaining ether was removed by incubation at 55°C for fifteen minutes with the lids of the Eppendorfs open. A one tenth volume of 3M sodium acetate pH4.5 was added with two volumes of ethanol to precipitate the DNA at -70°C for 15 minutes.

8.7.2.5.2 Direct use of LMP agarose gel slice DNA.

Latterly DNA was used for transformation of competent *E. coli* directly from a melted, diluted gel slice. This avoided losses incurred on phenol extraction of the material, and the presence of LMP agarose did not impair uptake of DNA. Water was added to the gel slice to give a final agarose concentration of 0.1%, and the gel slices were then melted at 65°C for five minutes. The solution was vortexed well and replaced at 65°C for a further ten minutes. This diluted DNA-LMP agarose solution was stored at -20°C and simply thawed prior to use.

8.8 Competent *Escherichia coli*.

8.8.1 Preparation using calcium chloride. (modified from Cohen *et al.*, 1972)

A single colony from a fresh culture plate was placed into a flask of TY broth and incubated at 37°C shaking at 300 rpm. When the culture had an A_{550} of 0.6 to 0.9 the cells were harvested by centrifugation in sterile 50ml tubes in an IEC Centra 4X at 2.5 Krpm for 5 minutes. The supernatant was discarded and the pellet resuspended in ice-cold 50 mM $CaCl_2$ at half the original volume. The cells were left on ice for a minimum of twenty minutes or for up to three hours. The cells were then centrifuged once again at 2.5 Krpm for 5 minutes and resuspended in ice-cold 50 mM $CaCl_2$ at one tenth the original volume.

Competent cells were sometimes stored at -70°C on adding glycerol to 16%, and were subsequently used directly after thawing.

8.8.2.1 Transformation of competent HB101

200µl of competent cells were mixed with plasmid DNA using 1ng or 0.1ng as a control. The cells were incubated on ice for 40 minutes and then heat-shocked at 42°C for 2 minutes. Standard broth (1ml) was added and the cells incubated at 37°C for one hour without shaking. The bacteria were pelleted by brief

centrifugation, most of the supernatant was discarded and then the pellet was resuspended in the remaining broth. The transformed HB101 were then spread onto SB-ampicillin (50μgml^{-1}) agar plates. The plates were incubated at 37°C overnight. Yields of up to 10^6 transformants were recovered per μg plasmid.

8.8.2.2 Transfection of competent JM101, JM109, or BMH 71.18 *mutL* with M13 DNA

200μl of competent cells were mixed with M13 DNA using 1ng or 0.1ng replicative form (RF) as a control. The cells were incubated on ice for 40 minutes and then heat-shocked at 42°C for 2 minutes. The cells were then added to 3ml soft agar at 42°C with 30μl IPTG (24$mgml^{-1}$ in distilled water) and 30μl BCIG (20$mgml^{-1}$ in dimethyl formamide) and poured onto SB agar plates. When set, the plates were inverted and incubated at 37°C overnight. Yields of 10^5 to 10^7 plaques were recovered per μg M13 RF.

Variations:

1) If the cells had been harvested early to make competent cells, then 100μl of dilute exponential JM101 could be added after the heat shock to improve the eventual bacterial lawn,

2) If *E. coli* BMH 71.18 *mut*L was being used then 100μl of dilute exponential JM101 was added after the heat shock to

reduce the chances of mutations occurring when the M13 replicated in the more error prone *mut*L strain.

3) Sometimes, when transfecting with closed circular DNA from mutagenesis copying reactions, 200µl standard broth was added after the heat shock and the mixture incubated at 37°C for one hour before plating in the soft agar. This was to allow a round of replication of M13 before plating in soft agar so that mutant and wild type alleles could segregate.

8.9 Preparation of plasmid DNA

A colony was toothpicked into 100ml standard broth containing ampicillin at 50µgml^{-1} in a 500ml flask and incubated at 37°C shaking at 300 rpm for about eight hours until turbid. Chloramphenicol was added to 100µgml^{-1} and the culture shaken slowly overnight, to amplify the plasmid. The bacteria were harvested by centrifugation in corex tubes in an MSE centrifuge at 8 Krpm for five minutes, and the supernatant was discarded. The pellet was resuspended in 1ml ice-cold 50mM glucose, 10mM EDTA, 25mM Tris-HCl pH 8.0, 4mgml^{-1} lysozyme, and left at room temperature for five minutes. To this was added 2ml 200mM NaOH, 1%SDS and the

mix was placed on ice for five minutes. Then 1.5ml potassium acetate was added (60ml 5M potassium acetate + 11.5ml glacial acetic acid + 28.5ml water), mixed well and once again placed on ice for five minutes. The mixture was centrifuged in a corex tube at 8 Krpm for ten minutes (MSE centrifuge) and the plasmid-containing supernatant was passed through a 0.45 μm pore acrodisc (Gelman). The volume of the supernatant was measured and CsCl was added at 1.22g per ml of supernatant, likewise 200μl of 750μgml^{-1} ethidium bromide per ml of supernatant.

The total plasmid preparation was sealed into a 5ml ultracentrifuge tube and centrifuged in a Beckmann vTi65.2 rotor at 55 Krpm overnight. The centrifuge was then switched off and after braking to 1 Krpm the rotor was allowed to coast down to stationary.

Plasmid bands were observed using longwave ultraviolet light and extracted from the tubes with a syringe and 22 guage needle. The ethidium bromide was extracted from the sample with water-saturated isopropanol before adding two volumes of water. The DNA was precipitated at -70°C by addition of a one tenth volume of 3M sodium acetate and 2-2.5 volumes of ethanol. In order to prevent precipitation of salt with the DNA, the frozen solution was first warmed to room temperature before centrifugation. The DNA pellet was washed in ethanol and resuspended in TE (0.1).

8.10 Preparation of bacteriophage M13 replicative form (RF) DNA.

A single plaque (less than four days old) was toothpicked into 10ml TY broth in a 50ml flask and grown for four hours shaking at 300 rpm at 37°C. This culture was then added to 100ml of very dilute JM101 in TY broth and grown for a further two to four hours. The bacteria were centrifuged in corex tubes at 6 Krpm for five minutes and the supernatant broth discarded. The pellet was washed in TE (0.1) to remove contaminating phage and was then treated as for a plasmid peparation. The DNA was purified on a caesium chloride density gradient as for plasmid preparations.

8.11 Preparation of bacteriophage M13 single-stranded (SS) DNA.

8.11.1 Rapid preparation in sets of 96 clones (I.C. Eperon, 1986a)

Small preparations of M13 single-stranded (SS) DNA were grown in Nunc 96 well flat bottomed microtitre plates in 250µl TY broth. Single plaques (less than four days old) were toothpicked into individual wells and the plate (without its lid) was then placed securely in a suitably sized air-tight plastic lunch box. The

cultures were incubated for seven hours at 37°C, shaking at 280 to 300 rpm.

In order to separate the bacteria from the M13-containing broth the cultures were filtered through a millititre-GV 96 well microtitre filter plate (Millipore) into a recipient flat bottomed microtitre plate (Nunc). To prevent cross contamination of the individual phage preparations an intermediate transfer plate was made from a Nunc round bottomed 96 well microtitre plate (with the edges cut off and a hole pierced at the bottom of each well) fixed to the uppermost filter plate with a lattice of silicone adhesive. This assembly was fitted into a millipore vacuum manifold and the filtration carried out by vacuum.

Following the filtration, the uppermost filter plate was prised off the intermediate transfer plate and discarded. Any residual filtrate still in the intermediate transfer plate could be centrifuged into the recipient plate using the appropriate rotor in an IEC Centra 4X bench centrifuge at 500 rpm for two minutes. The intermediate transfer plate was then removed and, after cleaning, could be re-used.

M13 phage from the filtrate were then precipitated by addition of 50µl 20% polyethylene glycol 6000, 2.5M NaCl (PEG/NaCl) to each well. The wells of the plates were sealed with white or yellow Sellotape, inverted to

mix the PEG/NaCl solution, and then left at room temperature for 20 minutes. Phage precipitates were pelleted by centrifugation at 2,500 rpm (600g) in the IEC Centra 4X for 20 minutes. The supernatant was discarded by aspiration into a side arm trap flask and the soft pellets resuspended in 100µl TE (0.1). The phage were precipitated once again as above by the addition of 25µl PEG/NaCl. Following aspiration of the supernatant the pellets were respun briefly to bring down residual PEG/NaCl solution from the sides of the wells which in turn was aspirated.

The precipitated phage were lysed by resuspending them in 40µl TE (0.1), 0.5% SDS, covering the plate with Nescofilm and floating it on a 80°C water bath for 10 minutes (with the stirrer switched off).

The liberated SS M13 DNA was precipitated by the addition of 4µl 3M sodium acetate pH 4.5 and 100µl ethanol then placing the covered plate at -20°C overnight or at -70°C for more than one hour. The precipitated DNA was centrifuged at 2,500 rpm in the IEC Centra 4X for 30 minutes, the supernatant aspirated and discarded and the pellets washed in 100µl ice-cold ethanol. The pellets were dried under vacuum and then resuspended in 50µl TE (0.1) resulting in sufficient single-stranded template in each well for 25 sequencing tracks.

8.11.2 Larger scale preparation

A single plaque (less than four days old) was picked into 10ml of TY broth in a 50ml flask and shaken at 300 rpm at 37°C flask for seven hours. When a yet larger culture was required, the 10ml culture was added to a larger flask containing TY and very few exponential JM101 after two hours, and grown for a further five hours. The bacteria were centrifuged in corex tubes at 8,000 rpm in an MSE centrifuge for five minutes and the phage-containing supernatant filtered through a 0.45μm pore acrodisc (Gelman) to remove any remaining bacteria.

A quarter volume of PEG/NaCl was added to the phage-containing filtrate, the solution mixed well and allowed to stand at room temperature for 20 minutes. The precipitated phage were centrifuged in corex tubes at 8,000 rpm in an MSE centrifuge for 15 minutes and the supernatant discarded. Residual PEG/NaCl solution was centrifuged from the sides to the bottom and aspirated. The phage pellet was resuspended in 100-200μl TE (0.1) and transferred to a 1.5ml Eppendorf tube. A quarter volume of PEG/NaCl was added. The precipitation procedure was repeated to remove any contaminants brought down with the first pellet. The precipitates were centrifuged in a microfuge at 12,000g for 15 minutes, then the supernatant was aspirated. The phage pellet was once again resuspended in 100-200μl TE (0.1)

and phenol extracted to remove the protein, so leaving the M13 SS DNA. The DNA was precipitated with ethanol and sodium acetate, washed in ethanol and resuspended in TE (0.1).

8.12 Dideoxysequencing of DNA cloned into bacteriophage M13.

Bacteriophage M13 DNA was sequenced by the dideoxy (chain termination) method of Sanger *et al.* (1977). This involved annealing an oligonucleotide "primer" to the single-stranded form and extending from the 3' end of the primer using DNA polymerase I large fragment (Klenow polymerase). Four reactions were performed, namely A,C,G and T, each containing the four dNTPs, one of which was at a low concentration. The corresponding dideoxyribonucleoside triphosphate (ddNTP) was included and competed for incorporation on the extending 3' end. Once incorporated, the dideoxynucleotide prevented further extension and was resistant to the 3'-5' exonuclease of the Klenow polymerase. The products were made radioactive either by use of a phosphorylated primer labelled with a ^{32}P at the 5' end (see later), or by incorporating either $\alpha^{32}P$-dATP or $\alpha^{35}S$-dATP into the reaction mixes and excluding unlabelled dATP. The four reactions were analysed by electrophoresis on denaturing poyacrylamide gels.

Mixes required for dideoxysequencing (stored at -20°C):

ddATP	0.04mM in TE (0.1)
ddCTP	0.25mM in TE (0.1)
ddGTP	0.16mM in TE (0.1)
ddTTP	0.50mM in TE (0.1)

A°	200µl	0.5mM dCTP	+	200µl	0.5mM dGTP	+	200µl	0.5mMdTTP
C°	10µl	"	+	200µl	"	+	200µl	"
G°	200µl	"	+	10µl	"	+	200µl	"
T°	200µl	"	+	200µl	"	+	10µl	"

α^{32}P-dATP	4000Cimmol^{-1}
α^{35}S-dATP	1000Cimmol^{-1}

Each reaction required 1µl of N° +1µl ddNTP + 1µCi α^{32}P-dATP or 2µCi α^{35}S-dATP where appropriate. When using 5' labelled primers, unlabelled dATP was added to the reaction mixes at the same concentration as the usual labelled dATP.

Sequencing reactions were sometimes carried out in Eppendorf tubes but were more usually performed in Nunc flat bottomed 96 well microtitre plates. When sequencing templates that had had been prepared in microtitre plates, 2µl of SS DNA were used for each of the four sequencing reactions. For each reaction 10-40nmoles of template and 30nmoles of primer were annealed at 65°C in 2µl of 2xSeq

(20mM Tris-HCl pH 7.5, 10mM $MgCl_2$) for five minutes. When using microtitre plates this step was performed over a 70°C water bath with the wells covered in Nescofilm. The primer/template mixes were then allowed to cool to room temperature for at least 20 minutes.

To each reaction was added:

2µl (N° + ddNTP)

plus either 2µl (α^{32}P-dATP 0.5µCiµl^{-1} + 0.25 units Klenow polymerase)

or 2µl (α^{35}S-dATP 1.0µCiµl^{-1} + 0.25 units Klenow polymerase)

The Klenow polymerase and labelled dATP (where appropriate) were diluted in ice-cold water or 25% glycerol, 25mM KPO_4 pH 7.5.

The sequencing reactions were incubated at room temperature for 20 minutes or longer with α^{35}S-dATP, 2µl of a "chase" mix (0.5mM each dNTP) was added and then the reactions were incubated at room temperature for a further 15 minutes.

The reactions were stopped by adding 2 to 4µl of formamide dyes and incubating for eight minutes at 80°C with the lids open. This denatured the DNA and reduced the volume by evaporation. Samples were loaded onto denaturing polyacrylamide gels and analysed by electrophoresis.

8.13 Purification of M13 single-stranded DNA

8.13.1 Sephadex G100 columns

SS M13 DNA was passed through a Sephadex G100 column in TE (0.1) to remove free dNTP contaminants. Sephadex G100 was preswollen by autoclaving in TE (0.1) and stored at 4°C.

Columns were prepared in 10ml or 1ml disposable sterile plastic pipettes with a sterile, siliconised, glass wool plug pushed to the bottom with a metal wire. The 1ml pipettes were used with the tapered end at the bottom, but 10ml pipettes were inverted to avoid problems with the dye front running as a tapered V, and the tapered end was cut off to facilitate pouring and loading. The columns were poured at room temperature with the pipettes partially submerged in TE (0.1) to avoid bubbles. The columns were washed with five column volumes of TE (0.1).

The length of the column was measured and two marks were made on the side; the first mark was 1/3 of the way down from the top to bottom and the second mark was 1/2 way down.

The DNA was mixed with a small amount of bromophenol blue and was loaded in as small a volume as possible (about 20µl for a 1ml column, less than 100µl for a 10ml column). The column was kept loaded with running buffer and the effluent was collected from when the dye front

reached the 1/3 mark until it spanned the 1/2 mark. DNA was precipitated from this fraction using sodium acetate and ethanol, washed, dried and resuspended in TE (0.1).

8.13.2 Spun Sephadex columns.

Spun columns were also used to remove free dNTPs from DNA. Sephadex G100 could not be used because the beads would be crushed; Sephadex G50 and G25 were used.

Disposable sterile plastic syringes (1ml) were used to make the columns and were plugged at the bottom with sterile, siliconised, glass wool. Sterile G50 (or G25) beads (pre-swollen in TE (0.1)) were loaded into the syringes until full. These were then suspended in disposable culture tubes (10ml Sarstedt) and centrifuged at 4,500 rpm in an IEC Centra 4X for five minutes. More Sephadex beads were added and the procedure repeated until the column volume was 0.9ml of compacted Sephadex. The column was washed repeatedly with 100µl running buffer (TE (0.1) for SS DNA or 10mM Tris-HCl pH 7.5, 5mM $MgCl_2$ for annealed primer-template DNA), until the effluent recovered after a five minute centrifugation at 4.5 Krpm was exactly 100µl. The DNA was loaded in a total volume of 100µl; the column centrifuged for five minutes at 4,500 rpm and the effluent was ethanol precipitated, washed in ethanol and resuspended in TE (0.1).

8.13.3 NENsorb 20 cartridges (NEN DuPont)

NENsorb 20 cartridges were used to purify DNA according to manufacturers instructions. The sorbant was prewet with methanol and then washed with reagent A (100mM Tris-HCl, 10mM Triethylamine, 1mM EDTA pH 7.7). The DNA was loaded onto the cartridge in 200-400µl reagent A and washed through with 3ml reagent A followed by 3ml sterile distilled water. The DNA was eluted in 20% ethanol and came off in the first 200-400µl. It was then recovered by ethanol precipitation or by lyophilization and resuspended in TE (0.1).

8.14 Synthetic oligonucleotide primers.

8.14.1 Preparation

Synthetic oligonucleotide primers were prepared by J. Keyte using the paper disc method (Matthes *et al.* 1984).

8.14.2 Purification

The primer preparations were dissolved in 100µl TE (0.1) and ethanol precipitated. The resulting pellets were dissolved in 10µl TE (0.1) and 10µl formamide dyes, heated for 10 minutes at 80°C and centrifuged briefly to pellet any residual paper fibres. The supernatants were loaded onto a prewarmed 20% polyacrylamide-7M urea gel and electrophoresed until the bromophenol blue dye front

had migrated about 25cm. Primers 17 nucleotides long migrated about 5cm in advance of the xylene-cyanol dye front. The gels were wrapped in Saranwrap™, placed over a thin-layer chromatography (TLC) plate containing a fluorescent indicator, and the primer bands were detected by shadowing with short wave ultraviolet light. These showed as intense bands of 17-nucleotide primers and lesser amounts of smaller oligonucleotides which had migrated faster on the gel. Primer bands were excised and the DNA eluted from the gel slices in 400µl distilled water at 4°C overnight. The purified primers were precipitated, washed in ethanol, resuspended in 100µl TE (0.1) and their concentration was determined by measuring the optical density at 260nm. The yield ranged from 0.25 to 2.9 nmoles.

8.15 Generation of short extensions from an oligonucleotide to give a range of 3' termini.

Oligonucleotide primers were 5' phosphorylated with T4 polynucleotide kinase and rATP. Short extensions were limited by using very low concentrations of one dNTP and could be analysed by polyacrylamide gel electrophoresis (PAGE) on using α^{32}P-dNTP (usually dATP or dCTP) as the limiting nucleotide. Primer DNA (0.5 pmoles) was annealed to 0.1pmoles SS DNA in 2 x Seq. buffer at 56°C for five minutes and then for twenty minutes at room

temperature. This was sufficient for ten analytical tracks on PAGE.

Dideoxysequencing reactions were performed (using α^{32}P-dATP) for use as molecular weight size markers on an analytical gel.

The conditions for extensions were as follows per analytical track:

0.05pmoles primer DNA annealed to
0.01pmoles SS template DNA
0.25 units Klenow polymerase
0.25μCi (0.25pmol) α^{32}P-dATP (1000Cimmol^{-1})
8μM dCTP, 8μM dGTP, 8μM dTTP
1xSeq. buffer

The total reaction was incubated for 2-5 minutes at room temperature and stopped by adding either: (a) formamide dyes and denaturing at 80°C for 6 minutes, or (b) bromophenol blue in 50mM EDTA.

In some experiments the amount of α^{32}P-dATP was varied as was the time of reaction and concentration of dNTPs.

Unused free nucleotides were separated from the extensions by Sephadex G100 columns or by running the denatured extended primers into denaturing polyacrylamide gels and eluting the products larger than the original primer in elution buffer (500mM sodium acetate pH 5.0, 1mM EDTA, 0.2% SDS) at 55°C for 45 minutes. In preparative reactions the labelled dNTP concentration was decreased and

supplemented with unlabelled dNTP to the same end concentration of 0.1μM.

8.16 Assay of misincorporation directed by reverse transcriptase.

Misincorporation at the 3' end of primers was assayed in three reactions and compared to correct incorporation of the correct dNTP and of the correct ddNTP. In most experiments these were analysed on 20% polyacrylamide 7M urea denaturing gels with dideoxysequencing tracks as markers.

Reverse transcriptase reactions were performed in 5μl 50mM Tris-HCl pH 8.0, 40mM KCl, 6mM $MgCl_2$, 1mM DTT. Reactions, unless otherwise stated, were incubated for 10 minutes at 37°C with 12 units (0.3pmole) of reverse transcriptase (Anglian Biotechnology Ltd, "Super RT") per 0.15pmole template and 0.1pmole 5'-^{32}P-primer (previously annealed in 2xSeq. buffer). In some experiments conditions were varied to perform studies of the kinetics of the reactions and to investigate the effects of different ratios of enzyme/template/primer concentrations. One reaction included 0.1mM ddNTP where N was the base complementary to the next position on the template after the primer 3' terminus. In the four other reactions dATP, dCTP, dGTP or dTTP were included to 0.1mM. Most reactions were performed in Nunc flat bottomed microtitre dishes and were stopped after 10 minutes by the addition of 2μl formamide dyes.

Experiments requiring reactions to be stopped at different times were carried out likewise but using 0.5ml Eppendorf tubes (with a polystyrene block over the lids to reduce condensation). Samples were heated at 80°C for 10 minutes loaded onto prerun 20% denaturing polyacrylamide gels and run at 2000V until the bromophenol blue had migrated 25cm. Gels were wrapped in Saranwrap™ and exposed to X-ray film at -70°C.

Variations:

1) Where αSdATP, αSdCTP, αSdGTP, or αSdTTP were to be misincorporated, these were added (instead of dNTPs) to 0.1mM.

2) Exonuclease III digestions. After a misincorporation reaction the reverse transcriptase was inactivated either at -20°C for six hours or by incubating at 70°C for ten minutes. The reaction mix was adjusted to 1xExo buffer, 5 units exonuclease III were added, and the mixture was incubated for 30 minutes at 37°C.

Scanning of autoradiographs.

An LKB ultroscan laser densitometer was sometimes used to scan autoradiograms in order to quantify the misincorporation products.

8.17 Phosphorylation of oligonucleotide primer 5' termini

Oligonucleotide primers were supplied with 5' hydroxyl ends and required phosphorylation before use in mutagenesis. The phosphate donor was the γ-phosphate of rATP so allowing 5' labelling of primers by use of γ^{32}P-rATP. The enzyme used was T4 polynucleotide kinase and the reaction performed in 1xC (50mM Tris-HCl pH 7.5, 10mM $MgCl_2$, 1mM DTT), 1mM rATP with 1 unit of kinase per 2.5 pmoles primer at 37°C for one hour. When 5' labelling primers no unlabelled rATP was used, but 100μCi γ^{32}P-rATP (600Cimmol^{-1}) was added per 2-5 pmoles primer. When inactivation of the enzyme was required this was carried out at 70°C for 10 minutes.

8.18 Second strand synthesis of M13 DNA ("copying reaction")

Copying reactions were performed after oligonucleotide or misincorporation mutagenesis to create closed circular double stranded DNA (CC DNA). This CC DNA was separated from SS DNA by electrophoresis through ethidium bromide-containing agarose gels. Low melting point agarose was used when the DNA was to be eluted from the gel and used to transfect *E. coli*. By varying the amount of phosphorylated oligonucleotide primer used in different annealing and copying reactions, conditions could be established where all

the available SS DNA template was copied.

Annealing conditions:

0.2pmoles SS DNA template

0.2pmoles phosphorylated primer (or varying amounts)

Total volume of 3µl 2xSeq. buffer

Incubation was at 65°C for five minutes, then at room temperature for at least twenty minutes.

A copy-mix was prepared for n copying reactions:-

n µl of a mix of all four dNTPs, each at 2.5mM

0.5n µl 10mM rATP

0.5n µl 10xSeq. buffer

0.5n units T4 ligase

0.5n units Klenow polymerase

distilled water to 6n µl.

6µl of copy-mix was added to each 3µl annealed template-primer, and incubated at 25°C overnight. A variety of incubation temperatures and times were tested including 37°C for 15 minutes and 30°C for 2 hours. Sometimes the copyings were carried out in 1xTMS buffer instead of 1xSeq. buffer.

The reaction products were analysed on agarose gels.

8.19 Analysis of 3' terminal nucleotides of oligonucleotide primers.

In order to analyse the 3' terminal nucleotide of an oligonucleotide or an extended primer (*eg.* after

misincorporation) a labelled nucleotide was transferred to the 3' hydroxyl end. The nucleotide, α^{32}P-dideoxy ATP, was added to the 3' end of the primer using calf thymus terminal deoxynucleotidyltransferase (terminal transferase).

8.19.1 3' end labelling.

0.25pmole primer

0.5µl 10xcacodylate buffer (4M potassium cacodylate, 500mM Tris-HCl 40mM DTT; pH7.5)

distilled water to 4.5µl

0.5µl 10mM $CoCl_2$

1µCi α^{32}P-ddATP (5500Cimmol^{-1})

4units (0.2µl) terminal transferase

The reaction was performed at 37°C for one hour and then stopped by addition of 100µl 50mM EDTA. The 3' labelled primer was purified from remaining free α^{32}P-ddATP by filtration through a 1ml Sephadex G25 spun column, and ethanol precipitated with 2µg unlabelled SS DNA as carrier at -20°C overnight.

8.19.2 Digestion with calf spleen phosphodiesterase

Calf spleen phosphodiesterase is an exonuclease which requires a free 5' hydroxyl group and releases 3' mononucleotides. Digestion of the 3' end-labelled primers released 3' mononucleotides so that the labelled phosphate was now associated with the 3' terminal

nucleotide of the original primer.

3' labelled oligonucleotide primer was resuspended in 10µl 50mM K-MES pH 8.0. 100µl calf spleen phosphodiesterase, supplied as a suspension in ammonium sulphate solution, was centrifuged and the pellet resuspended in 10µl 50mM K-MES pH 8.0. 1µl of enzyme ($2mgml^{-1}$) was added to the 3' labelled DNA and the mixture was incubated at 37°C for two hours. The reaction was stopped by freezing at -20°C.

8.19.3 Thin layer chromatography (TLC) on polyethyleneimine (PEI) plates. (Volckaert and Fiers, 1977)

PEI plates were washed in 2M formic acid-pyridine pH 2.2, rinsed in water and then dried and stored at 4°C wrapped in aluminium foil.

0.1µl samples were spotted onto the plates which were then developed in 0.5% formic acid and dried. The plates were once more developed, this time in 0.15M lithium acetate/formic acid pH 3.0, dried and exposed to X-ray film.

8.20 Maxam and Gilbert sequencing of oligonucleotides.

Oligonucleotide primers and their extended products of lengths 14 to 24 nucleotides were sequenced by a modified Maxam and Gilbert reaction (Maxam and Gilbert, 1980).

Primers were radiolabelled at the 5' end with T4 polynucleotide kinase and γ^{32}P-rATP before any extensions by misincorporation. These primers and their extended products were purified by electrophoresis through 20% polyacrylamide/7M urea gels and the radioactive primers were excised and eluted in water. The eluate was precipitated and washed in ethanol and then resuspended in water.

Four tracks of sequencing were performed per primer, namely T (thymine), A+C (adenine plus cytosine), G (guanine) and C+T (cytosine plus thymine). In each track up to 0.25pmole of primer was used, and where the amount of 5'-labelled primer to be sequenced was significantly less than 0.25 pmole then non-labelled universal primer was added to ensure standard conditions.

Modification reactions were performed as follows:-

Thymine

the 5'-labelled primer was dissolved in 5µl H_2O to which was added 10µl of a 5% (w/v) solution of OsO_4 and 1µl 98% piperidine. The brown yellow solution was incubated at 25°C for 75 minutes before the reaction was stopped with 100µl 0.3M sodium acetate pH 4.5 and placed in ice.

Adenine plus Cytosine

the 5'-labelled primer was dissolved in 5µl H_2O to which was added 80µl of 1.2N NaOH. The Eppendorf tubes

were sealed with Nescofilm and incubated at 90°C for 60 minutes with lead weights on the lids to prevent the tubes opening. The reaction was stopped with 120µl 1N acetic acid and placed in ice.

Guanine

the 5'-labelled primer was dissolved in 5µl H_2O to which was added 160µl DMS buffer (50mM sodium cacodylate pH 8.0, 10mM $MgCl_2$, 1mM EDTA), and 1µl DMS (dimethyl sulphate). The reaction was incubated for 2 minutes at 37°C and then stopped with 30µl 1M sodium acetate pH 4.5 and placed in ice.

Cytosine plus Thymine

the 5'-labelled primer was dissolved in 20µl H_2O to which was added 30µl hydrazine. The reaction was incubated for 35 minutes at 37°C and then stopped with 50µl 0.3M sodium acetate pH 4.5 and placed in ice.

After the modifications each reaction was desalted by repeated ethanol precipitations: 5µl of 1mgml^{-1} $tRNA^{Phe}$ was added to each tube to aid precipitation along with 4 volumes of ice-cold ethanol. After leaving for 30 minutes in dry ice the tubes were centrifuged at 12,000g for 10 minutes in a cold microfuge. The modified primer was dissolved in 30ul H_2O by vortexing, a further 120µl ethanol was added and the tubes placed in dry ice before centrifuging as before. The

precipitates were washed with 100μl ethanol and the pellets dried under vacuum.

Following desalting, the modified primers were cleaved with piperidine at the sites of base modification. Each desalted and dried pellet of modified primer was dissolved in 50μl 1M (10% v/v) piperidine and incubated for 30 minutes at 90°C with lead weights on the lids of the tubes to prevent them opening and releasing the volatile piperidine. The reactions were stopped by cooling in ice and then the cleaved DNA was dried under vacuum in a Speed Vac. The products were twice washed in 50μl H_2O and dried under vacuum before dissolving in 5μl formamide dyes, heating for 10 minutes at 80°C and applying to 20% polyacrylamide 7M urea gels. The bromophenol blue dye was run about 18cm and then the gel was transferred to the shiny side of Benchkote, wrapped in Saranwrap and exposed to Kodak XAR film for autoradiography at -70°C.

CHAPTER 9. RESULTS

9.1 Introduction

The results presented here are a study of the misincorporation of dNTPs into a DNA strand by the error-prone polymerase AMV reverse transcriptase as it copies a SS DNA template from a primer annealed to that template. The reaction required the provision of AMV reverse transcriptase and a single dNTP as substrate.

This chapter is divided into four sections. The first section is a description of the *in vitro* assay of misincorporation and of the results of that assay. The proportions and patterns of misincorporation and correct incorporation were assayed with all possible combinations of dNTP substrate, template nucleotide and the nucleotide at the 3' terminus of the primer, by analysis of the 5'-^{32}P labelled products on denaturing PAGE.

In section 2 are presented the results of the generation of a bank of primers, by extension of synthetic oligonucleotide primers, for subsequent misincorporations throughout a length of DNA. The primers have a common 5' terminus and random 3' termini which terminated mainly prior to a specified next-template nucleotide.

The third section demonstrates the generation of misincorporation mutants which involved the screening of mutagenesis progeny by observing a colour change of plaques, and/or by sequence determination of the SS DNA.

In section 4 the subcloning of the c-Ha-*ras* gene into the bacteriophage M13 is described. A bank of primers has been synthesized by short extensions of a synthetic oligonucleotide primer in preparation for misincorporation throughout the gene. Results are presented of a misincorporation experiment using one such bank of primers.

9.1.1 Avoidance of contaminating dNTPs.

Specific misincorporation of a given dNTP requires pure template DNA and dNTP stocks. Contaminating dNTPs, even at very low concentrations, could result in low levels of correct incorporation thus reducing the yield of mutants recovered.

9.1.1.1 Purity of template DNA.

SS M13 C6.9 template DNA was tested for the presence of contaminating dNTPs by incubating template-primers with reverse transcriptase in the absence of added dNTPs. The presence of trace amounts of dNTPs was detected by denaturing polyacrylamide gel electrophoresis, where bands of molecular weight greater than that of the primer could be seen. In most cases only one or two additional bands were seen per template-primer but, for the purposes of the misincorporation assay, these contaminants had to be removed. The problem was solved by passing the template DNA through a Sephadex G100 column, following which no

contamination was detectable (see later). This assay was repeated with all template-primer combinations.

9.1.1.2 Purity of dNTP stocks.

It was necessary to distinguish true misincorporation from slow correct incorporation due to very low levels of a contaminating dNTP. This was achieved by Maxam and Gilbert sequencing of the extended primer (see later), or by observing that addition of a single nucleotide to an oligonucleotide primer resulted in different mobilities on polyacrylamide gel electrophoresis for each of the four possible nucleotides. The order of mobility is C>A≥T>G and this can be seen in figure 3.

9.1.1.2.1 Impurity of some stocks of dNTPs: Maxam and Gilbert sequencing.

Initial misincorporation experiments showed more than a single base increase on the 3'OH terminus of a primer when using each different dNTP. Impurity was suspected because of strange banding patterns produced when using either dCTP or dGTP which each seemed to extend readily with both C and G template nucleotides. Figure 4 shows attempted misincorporations against a template 3'...GGGCCCCTA5' with different dNTPs and reaction conditions. The primer-template was free of contaminating dNTPs as seen by incubation with reverse transcriptase which gave no extension. Substrate

Figure 3.
Samples of misincorporation reactions analysed by PAGE.

Misincorporation reactions were performed with reverse transcriptase as described in Materials and Methods (8.16). Additions were as follows: n: no dNTP added, N: correct ddNTP added to give a reference correct incorporation, A: dATP, C: dCTP, G: dGTP, T: dTTP. The different mobilities can be seen with primer 5: C>T>G, and with primer 3: C>A. As an example of the assessment of misincorporation levels, primer 3 shows a good misincorporation with dCTP and a poor misincorporation with dATP. seq: dideoxy sequencing tracks ordered A,C,G,T. The major band at the bottom is unextended primer.

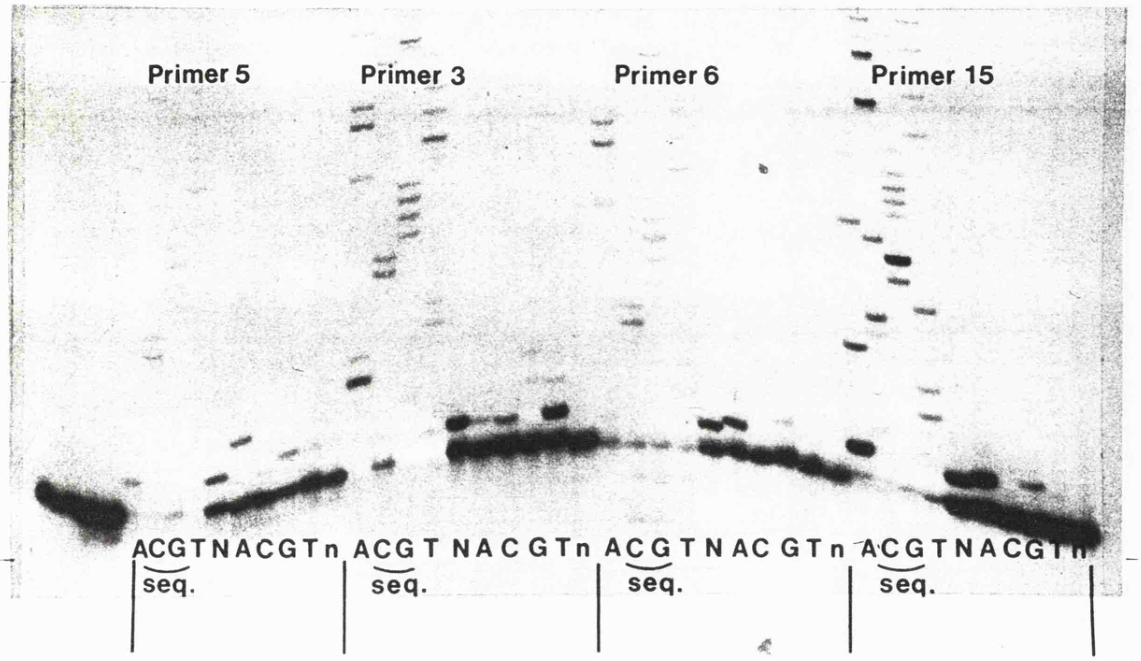

Primer 5
Primer 3
Primer 6
Primer 15
ACGTNACGTn ACGT NACGTn ACGT NACGTn ACGTNACGTn
seq.
seq.
seq.
seq.

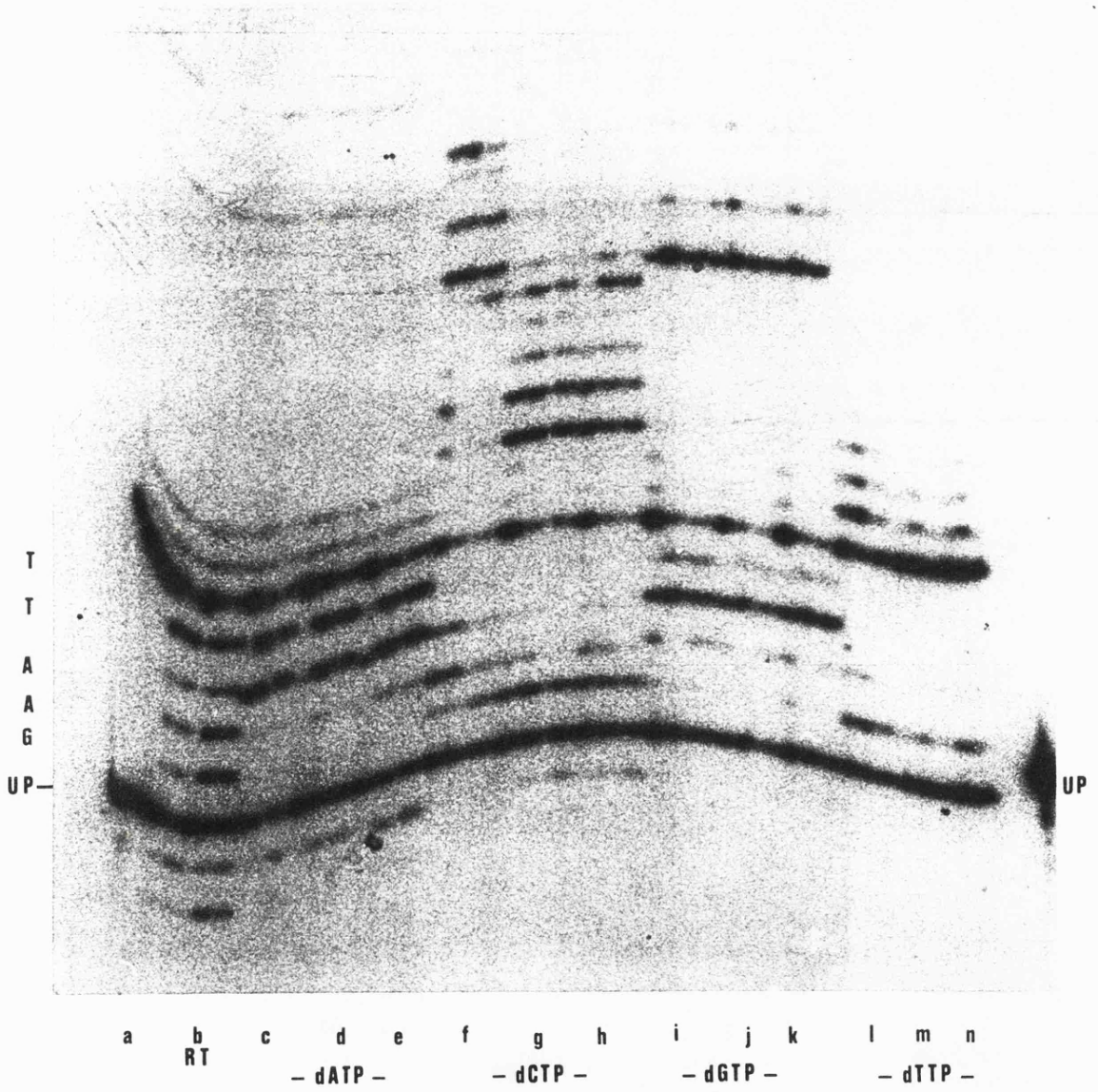
Figure 4.
Analysis of misincorporation with template 3'GGGCCCCTA5'.

Universal primer was annealed to template C6.9 and extended by five nucleotides using Klenow polymerase, dATP, dGTP, and dTTP.

Template: 3'...ACTTAAGGGCCCCTA
Universal primer: 5'...T
Extended primer: 5'...TGAATT

The products were purified through sephadex G100. Subsequent incubation of the extended primer-template alone (a), and with reverse transcriptase (b), gave no further extension, indicating the absence of contaminating dNTPs. The following tracks are samples incubated with reverse transcriptase and a single dNTP at different concentrations for different times:

dATP: c=1.0mM, 10 mins; d=0.1mM, 10 mins; e=0.1mM, 30 mins;
dCTP: f=1.0mM, 10 mins; g=0.1mM, 10 mins; h=0.1mM, 30 mins;
dGTP: i=1.0mM, 10 mins; j=0.1mM, 10 mins; k=0.1mM, 30 mins;
dTTP: l=1.0mM, 10 mins; m=0.1mM, 10 mins; n=0.1mM, 30 mins.


T
T
A
A
G
UP
UP
a
b
RT
c
d
e
f
g
h
i
j
k
l
m
n
– dATP –
– dCTP –
– dGTP –
– dTTP –

dATP was not misincorporated; dTTP was misincorporated at the first three positions, each successive misincorporation being weaker than the preceding one. Substrate dCTP was correctly incorporated at the first three positions and then apparently misincorporated at the following five positions and beyond; dGTP was apparently misincorporated at the first three positions and then further misincorporated beyond. Maxam and Gilbert sequencing later indicated that both dGTP and dCTP stocks were contaminated, shown in figure 5.

Primer 24 (5'CAGGTCGCGGATCCCC3') was annealed to SS mp8 giving next template nucleotides 3'CCCTT5'. Misincorporation reactions with dATP, dGTP and dTTP were performed and run on a 12% polyacrylamide gel. dTTP gave no extension, dATP gave an extension of five bases, the first being strongest, dCTP gave an extension of eight bases, the first being strongest, dGTP gave an extension of six bases, the fifth being strongest. The strongest bands were excised and the DNA eluted, precipitated and sequenced by the Maxam and Gilbert method. The tracks are shown in figure 5 and are in the order T, C+T, A+C, G.

The extensions with substrates dATP and dCTP show some contamination with co-eluted unextended primer. Nevertheless, substrate dATP is seen to have been misincorporated after the 3' terminal C. Substrate dCTP did not give such a clear result but, although some misincorporation cannot altogether be ruled out, it is apparent that some dGTP was correctly incorporated indicating contamination of the dCTP stock. Substrate dGTP

Figure 5.
Maxam and Gilbert sequencing of misincorporation reaction products.

Primer 24 was annealed to template mp8 and misincorporations were performed as described in the text.

Template: 3'...GCCCTT
Primer 24: 5'...C

The extensions from the dATP, dCTP, and dGTP reactions are presented, in each case the tracks are in the order (left to right) T, C+T, A+C, G. The arrow indicates the 3' terminal C of the unextended primer.


C
C
C
C
T
A
G
G
C
G
C
T
G
G
dATP
dCTP
dGTP

was correctly incorporated at the first three positions and although some dGTP misincorporation also occurred at the next two positions (template T), there are also bands in the A+C sequencing track. The most credible conclusion that can be drawn from this is that the dGTP stock contained trace contamination of dATP, giving correct incorporation at these two positions. A further example of this is shown in figure 6 with universal primer extended by three bases on incubation with reverse transcriptase and dGTP (labelled G in figure 6). The extension consisted of a correctly incorporated dGTP and correctly incorporated dATPs against a template 3'CTTAA5' proving contamination of the dGTP stock with dATP. The dGTP stock also contained dCTP contamination, as shown by:

1) an extension that read 5'CCCGGGG (template 3'GGGCCCC) when only dGTP had been added (figure 6: ExG), and

2) an extension that read 5' GGC (template 3'TCG) when only dGTP had been added (figure 6: Primer 42:G3).

One extension with dCTP read 5'GGGAA (template sequence 3'CCCTT) where only dCTP had been added. Therefore the dCTP stock was contaminated with both dGTP and dATP. Other such analyses showed that the dATP and dTTP stocks were pure.

9.1.1.2.2 Analysis of the 3' termini of primers: an attempt to verify misincorporation.

Where misincorporation had been poor (see later for description) then excision of the DNA band from the gel and

Figure 6.
Maxam and Gilbert sequencing of misincorporation reaction products.

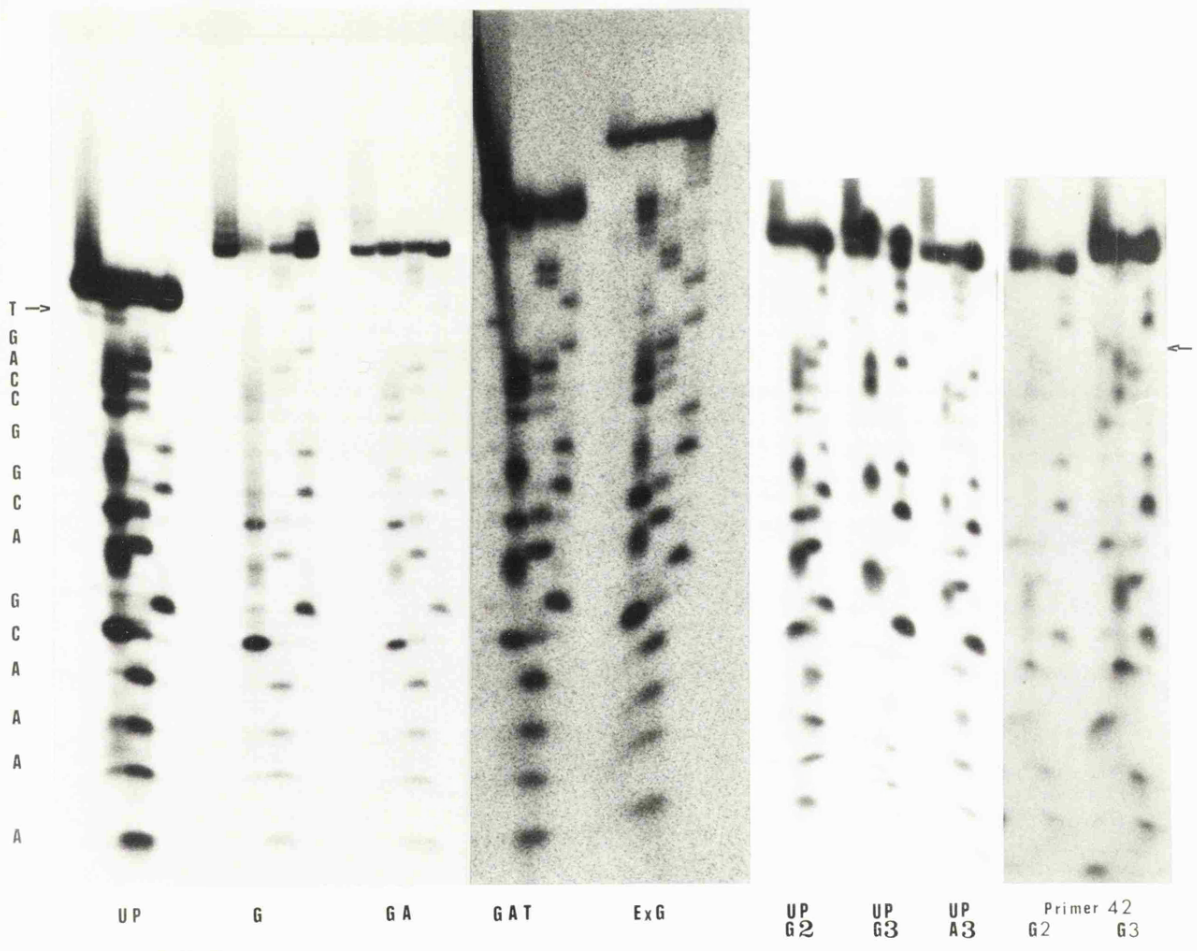
Universal primer (UP) was annealed to template C6.9.
Template: 3'...ACTTAAGGGCCCC
Universal primer: 5'...T
Exensions were with reverse transcriptase as follows:
G: dGTP
GA: dGTP + dATP
GAT: dGTP + dATP + dTTP
ExG: dGTP + dATP + dTTP, gel purify, reanneal, dGTP + reverse transcriptase.

Tracks are in the order (left to right) T, C+T, A+C, G, the sequence of universal primer is shown at the left of the picture, the arrow indicates the 3' terminal T of universal primer. G shows contamination of dGTP with dATP, GAT shows contamination of one of the dNTPs with dCTP, ExG shows contamination of dGTP with dCTP.

Universal primer and template C6.9 (template: 3'CTT):
UP G2: misincorporation with ultra pure dGTP to two bases
UP G3: misincorporation with ultra pure dGTP to three bases
UP A3: misincorporation with ultra pure dATP to three bases

Primer 42 was annealed to template mp8.
Template: 3'...ATCG
Primer 42: 5'...T
Extensions were with reverse transcriptase as follows:
G2: dGTP to two bases
G3: dGTP to three bases

The tracks are in the order (left to right) T, C+T, A+C, G; an arrow indicates the 3' terminal T of primer 42. A misincorporation of dGTP to template T is seen at the first template nucleotide. dCTP was incorporated at the third template nucleotide showing contamination of dGTP with dCTP.


T
G
A
C
C
G
G
C
A
G
C
A
A
A
A
UP
G
GA
GAT
ExG
UP G2
UP G3
UP A3
Primer 42
G2
G3

elution of the DNA almost invariably resulted in some cross contamination of the extended primer with the unextended primer. This caused two-fold problems on Maxam and Gilbert sequencing of the 5'-labelled primers:

1) the recovery was very low, requiring coprecipitation with unlabelled primer, with the result that the autoradiographs required very long exposures with intensifying screens, and

2) the contaminating unextended primer would result in a common band in all four sequencing tracks at almost the same position (sometimes a little higher on the gel) as the expected band showing the 3' terminal nucleotide of the extended primer.

An alternative technique was attempted to confirm that the 3' terminal nucleotide of the extended primer was indeed a misincorporated nucleotide. This technique involved labelling a primer at the 3' end with α-^{32}P dideoxy ATP using terminal transferase. Dideoxy ATP was used to ensure that each primer was only extended by one nucleotide. A primer thus labelled could be digested to completion with calf spleen phosphodiesterase releasing 3'-mononucleotides. In this way the 3' terminal nucleotide of the primer was released bearing the ^{32}P from the α-^{32}P ddATP. The digested primer was then subjected to thin layer chromatography to separate the four different bases thus allowing analysis of the 3' terminal nucleotide of the primer. The conditions for the reactions were established using four primers each with a different 3' terminal

nucleotide, and also by digesting sonicated salmon sperm DNA with calf spleen phosphodiesterase to give unlabelled 3' mononucleotides. Figure 7 shows the four positions of the four different 3'-mononucleotides from primers with 3' terminal nucleotides of G,T,A or C. The order of migration from the bottom of the thin layer chromatography plate to the top was G,T,A,C.

The application of this technique to misincorporated primers was not simple. As mentioned previously, the yields of such primers with poor misincorporations was, by definition, low, and could contain contaminating unextended primers. Moreover, co-precipitation with unlabelled primer was no longer a possibility because the 3' nucleotide of this second primer would also become labelled, and cause a contaminating spot on the eventual autoradiograph. A further problem was that calf spleen phosphodiesterase does not cleave DNA with a 5' phosphate, so 5'-labelled primers were unsuitable (unless phosphatase treatment and subsequent phenol extractions were to be employed). It was therefore decided to perform the preparative misincorporation reactions unlabelled, (with dATP and unlabelled primers) alongside parallel analytical reactions using α-^{32}P dATP, figure 8. The primers 2,6 and 7 were used, as shown below:

Template 3'...TCA

Primer 2 5'...A

Figure 7.
Determination of the 3' terminal nucleotides of primers using TLC.

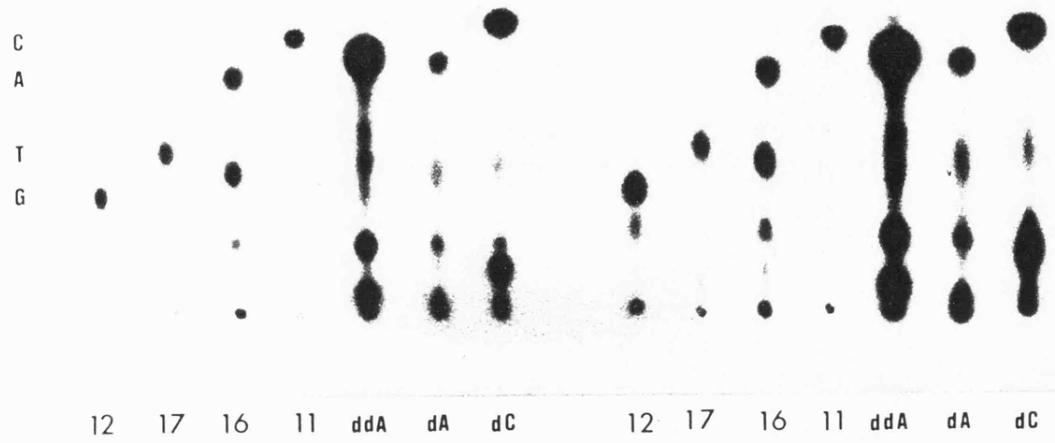
The four primers (numbers 11,12,16 and 17) were incubated with terminal transferase and α^{32}P-ddATP before purification through spun Sephadex G25 columns. All samples were then incubated with calf spleen phosphodiesterase, liberating 3' mononucleotides. These were separated by TLC as described in Materials and Methods (8.19). G,T,A,C indicate the positions to which the respective 3' mononucleotides migrated. The tracks are as follows:
12: primer 12, 3' terminal nucleotide G; 17: primer 17, 3' terminal nucleotide T; 16: primer 16, 3' terminal nucleotide A; 11: primer 11, 3' terminal nucleotide C; ddA: α^{32}P-ddATP; dA: α^{32}P-dATP; dC: α^{32}P-dCTP.

Figure 8.
Preparative gel of unlabelled misincorporation products, with analytical tracks.

Primers 2,6 and 7 were annealed to template C6.9.

Template: 3'...TCACTTAAG
Primer 2: 5'...A
Primer 6: 5'...AGTGA
Primer 7: 5'...AGTGAA

Misincorporation reactions were performed for 30 minutes at 37°C with dATP and reverse transcriptase. Flanking each wide preparative track (unlabelled products) are the products of a similar misincorporation reaction but using α^{32}P-dATP instead of dATP. Primer 2 shows two strong bands, the lower being unextended primer and the second being primer extended by a misincorporated A. A third weak band represents primer extended by two A misincorporations. Primer 6 shows two strong bands, the lower being primer extended by a correctly incorporated A, and the second being primer extended by a correctly incorporated A and a misincorporated A. A third weak band represents primer extended by a correctly incorporated A, and two A misincorporations. Primer 7 shows two strong bands, the lower being unextended primer and the second being primer extended by a misincorporated A. A third weak band represents primer extended by two A misincorporations.


C
A
T
G
12
17
16
11
ddA
dA
dC
12
17
16
11
ddA
dA
dC

Template 3'...TTA

Primer 6 5'...A

Template 3'...TAA

Primer 7 5'...A

Equal amounts of each primer were used, and the uppermost band in each case represented the misincorporated product. The very strong band of primer 6 represented a correctly incorporated dATP at the first template position. The weak bottom band of primers 2 and 7 was the 3' terminal nucleotide A of the primers that had become weakly labelled by virtue of a trace exonuclease activity present in the 30 minute incubation with reverse transcriptase. The products were excised, eluted and precipitated, but gave no results after 3'-labelling and digesting with phosphodiesterase. The misincorporation was repeated but after elution of the DNA from the gel it was desalted by passing it though a spun Sephadex G50 column and drying down the effluent in a Speed-Vac to prevent losses incurred on ethanol precipitation. However, once again no labelled primer was recovered after the terminal transferase reaction and so no results could be seen after TLC.

One final method was used for isolating primers with misincorporations at the 3' termini. This involved removing the dATP after the misincorporation reaction by passing the primer through a spun Sephadex G50 column, and then labelling all the reaction products with α-^{32}P ddATP using

terminal transferase before electrophoresis. All bands of possible misincorporation products were excised, eluted and precipitated with carrier, but so few counts were recovered that subsequent 3' terminal analysis with phosphodiesterase produced no clear results.

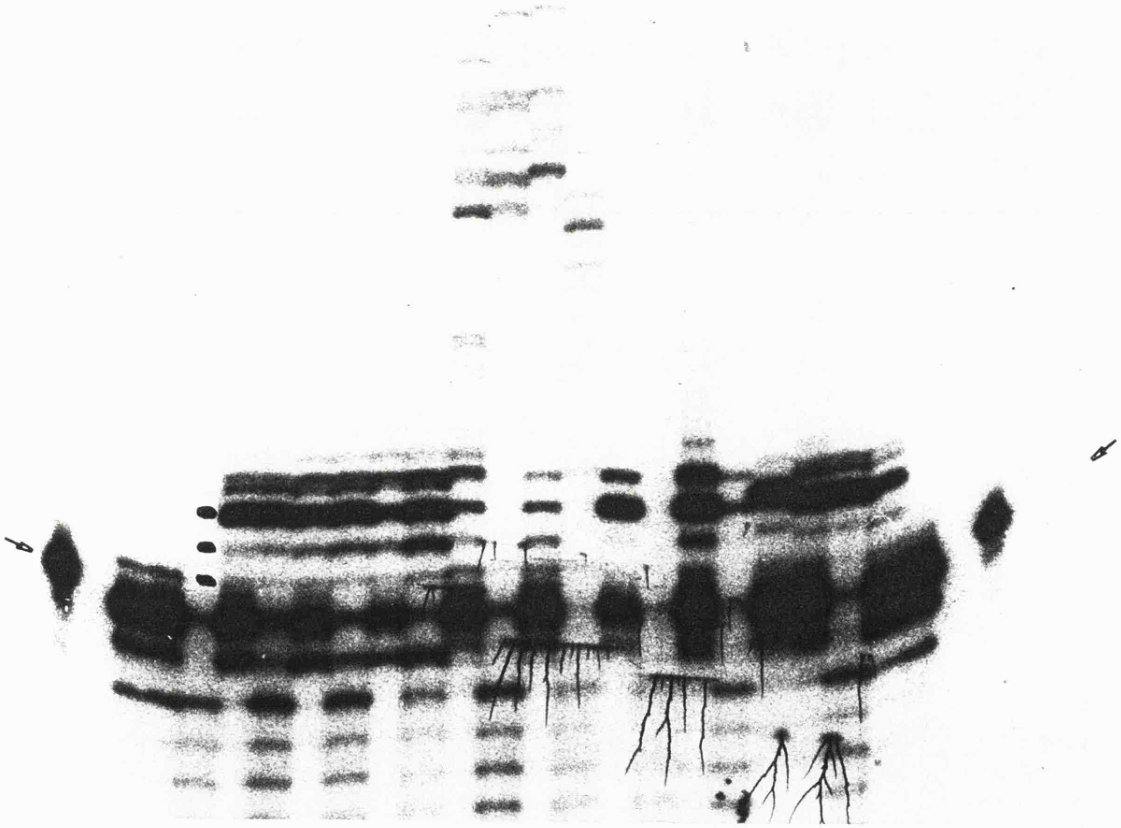
All the experiments reported below (unless otherwise stated) used new "ultrapure" stocks purchased from Pharmacia which indeed proved to be pure by the criteria described here.

9.1.2 The use of α-SdNTPs for misincorporation.

Misincorporation can also use α-SdNTPs as substrates, shown in figure 9. Universal primer was annealed to template C6.9. Misincorporation with either dATP or α-SdATP gave a major product band at position 3 of the template 3'CTTAA5' arising from a misincorporation followed by two correct incorporations. The products of the misincorporation reactions were subjected to exonuclease III digestion which digested all primers and extended primers except those with a thionucleotide at the 3' terminal end. In this way any unextended products present after a misincorporation reaction using α-SdNTPs could be eliminated. Misincorporation of ultra pure dCTP with this primer-template did not occur (not shown in this figure) but use of α-SdCTP gave many extension products resistant to exonuclease III digestion, indicating that the α-SdCTP stock was insufficiently pure for specific misincorporation.

Figure 9.
Misincorporation of α-SdNTPs

Universal primer was annealed to template C6.9.
Template: 3'...ACTTAA
Universal primer: 5'...T
The arrow indicates the position of unextended primer.
a: unextended primer; b: primer + reverse transcriptase, nc dNTPs were added but the weak band of extended primer indicates low level dNTP contamination of the reaction (the template DNA had not been purified through Sephadex G100).
c, c_1, c_{11}: misincorporation of α-SdATP for 10, 30, 160 min.,
d, d_1, d_{11}: exonuclease III digestion of c, c_1, c_{11}.
e, e_1: misincorporation of α-SdCTP for 10 and 160 min.,
f, f_1: exonuclease III digestion of e, e_1.
g, g_1: misincorporation of dATP for 10 and 160 min.,
h, h_1: exonuclease III digestion of g, g_1.
i: incorporation and misincorporation of dGTP for 10 min.



a b b c d c_i d_i c_{ii} d_{ii} e f e_i f_i g h g_i h_i i

9.1.3 Predefined extensions of primers.

As described previously, misincorporations have been achieved using a number of different synthetic oligonucleotide primers. However, initial misincorporation experiments using ladders of primers sometimes failed to show the expected shift in the banding pattern on PAGE analysis when trying to misincorporate single dNTPs. This cast into doubt the ability of AMV reverse transcriptase to direct misincorporation onto a bank of primers previously extended by Klenow polymerase. Such an unexpected problem would necessitate the redesign of the proposed strategy of mutagenesis onto a bank of primers spread throughout a gene.

Accordingly, controlled extensions of the universal primer were carried out for subsequent misincorporations to demonstrate that pre-extended primers could be used as effectively as unextended primers. Universal primer was annealed to template C6.9 and extended with a mix of dATP, dGTP, and dTTP for five minutes at room temperature before heat inactivation, reannealing of primer and template, and purification through a Sephadex G100 column.

Template: 3'...ACTTAAGGGCCCC

Universal Primer: 5'...T

+dGTP, +dATP, +dTTP

Initial Extension: 5'...TGAATT

Sephadex G100 purification

The purified products were incubated without reverse

transcriptase, with reverse transcriptase and with reverse transcriptase plus dGTP. (The use of ultra pure dGTP stocks had not been introduced and so it was expected that reverse transcriptase plus dGTP would give the long extension seen in figure 10 against the template 3'GGGCCCC.)

Template: 3'...ACTTAAGGGCCCC

Extended Primer: 5'...TGAATT

Misincorporation: 5'...TGAATTNNNNNNN

The initial extensions were carried out with reverse transcriptase, Pharmacia Klenow polymerase, Klenow polymerase fraction 1 (gift of Dr. L.P. Eperon), and Klenow polymerase MC1 (gift of M. Sweeney). Only the extension by reverse transcriptase was clean, stopping after extending 5'GAATT. Pharmacia Klenow and Klenow fraction 1 when incubated with dATP, dGTP and dTTP gave many other extensions demonstrating contamination of the enzyme preparations with at least dCTP. Klenow MC1 was incubated with only dGTP and dATP and showed a weak band against a template A.

Purification of the extensions gave dNTP-free products as seen by incubating the extension with reverse transcriptase and no added dNTPs. Addition of dGTP gave new extension products either by misincorporation or by incorporation of trace amounts of contaminating dNTPs. This showed that the extended primers, when purified through Sephadex G100 could be used for further extension by reverse transcriptase, regardless of which polymerase had been used for the initial extension. Figure 10 also shows

Figure 10.
Purified extensions of primers with different DNA polymerases: their purification and use as primers for subsequent extension.

Universal primer (UP) was annealed to template C6.9.

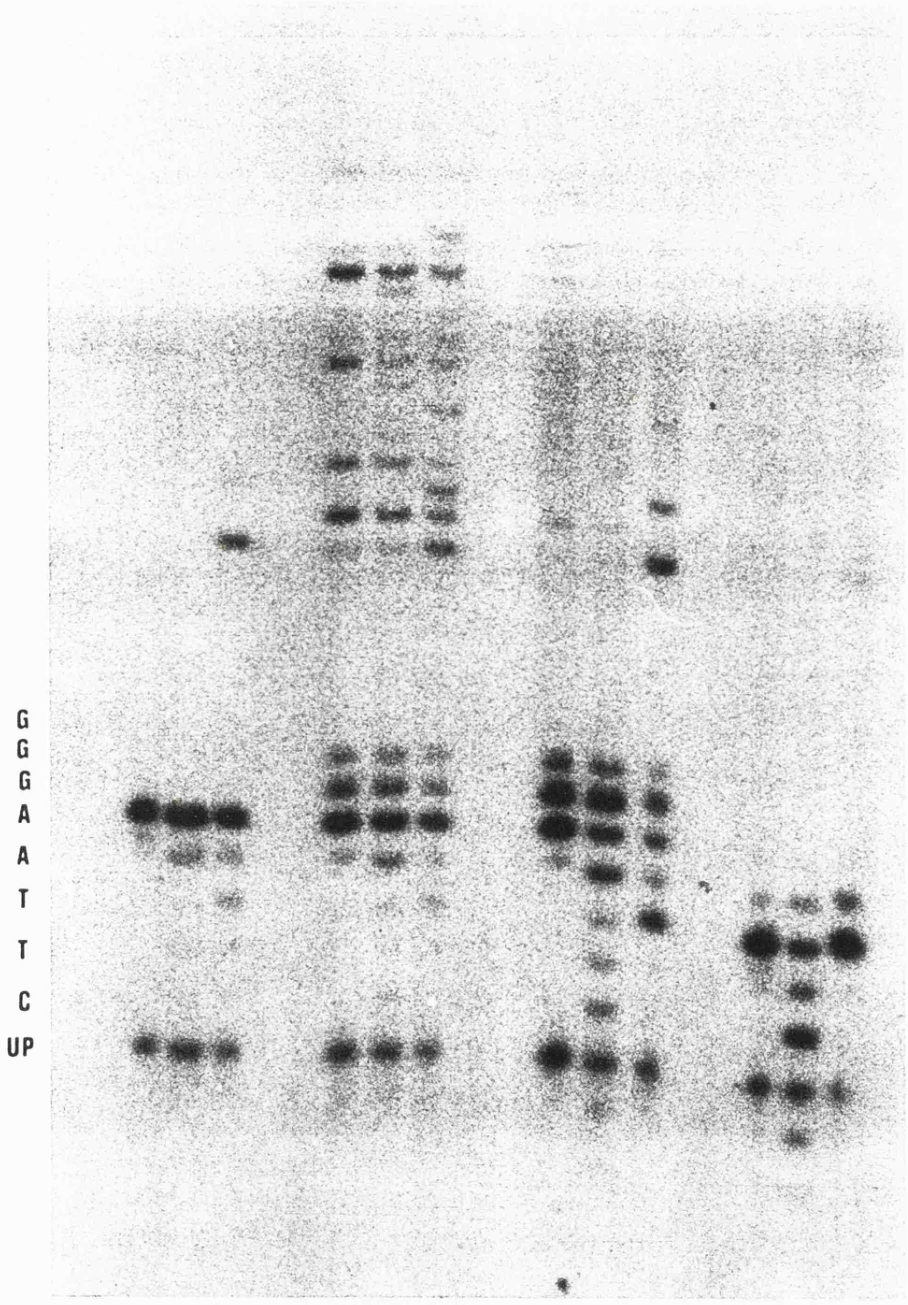
Template: 3'...ACTTAAG

Universal primer: 5'...T

The template sequence is shown at the left side of the picture. Initial extensions were for five minutes at room temperature as follows:

1. reverse transcriptase + dATP, dGTP, dTTP,
2. Pharmacia Klenow polymerase + dATP, dGTP, dTTP,
3. Klenow polymerase fraction 1 + dATP, dGTP, dTTP,
4. Klenow polymerase MC1 + dATP, dGTP,

After purification of the reaction products through Sephadex G100 they were used for misincorporation reactions with dGTP (not ultra pure). The first track of each set of three shows the extended primer-template incubated at 37°C for 30 minutes. The second track of each set of three shows the extended primer-template incubated at 37°C for 30 minutes with reverse transcriptase. The third track of each set of three shows the extended primer-template incubated at 37°C for 30 minutes with reverse transcriptase and with dGTP.


G
G
G
A
A
T
T
C
UP
1
2
3
4

misincorporation by reverse transcriptase from a primer previously extended with reverse transcriptase.

9.1.4 The effect of template primary structure.

Misincorporation of each dNTP was seen to vary according to which primer-template was used. Figure 11 shows the different patterns of misincorporation from the universal primer when annealed to two different templates. However, the next template nucleotide in each case was C, and even when combined with the results from other primers and templates described previously the combinations of primer and template that were available were too few to deduce the factors determining misincorporation patterns.

The effect of the 3'OH terminal nucleotide of the primer was investigated by designing 16 primers. For each set of four primers ending in the same 3'OH nucleotide, each one had a different nucleotide at the next position on the template (Table 1). The results of the misincorporation were analysed by comparing the amount of extended product to an internal reference incorporation for each primer. This internal reference used the correct incorporation of a dideoxy NTP (ddNTP) complementary to the next template nucleotide to assess the maximum possible level of incorporation, i.e. the available annealed primer assuming that the reaction went to completion. In practice, in some cases (Primer 3: dTTP *vs.* ddTTP, Primer 9: dCTP *vs.* ddCTP) dNTP gave up to twice as much incorporation as ddNTP: the

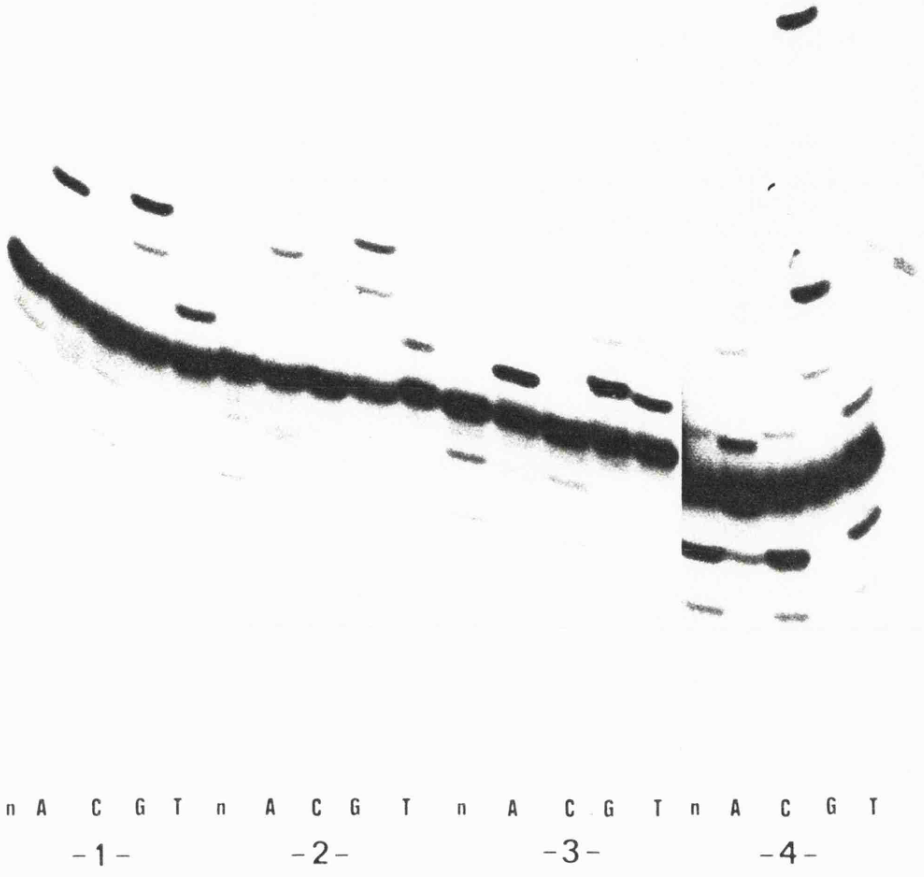


Figure 11.
Misincorporation from universal primer on two templates C6.9 and M-*ras*, and from primer 51 on template M-*ras*.

Sets 1 and 2. Template C6.9: 3'...ACTTAAG
Universal primer: 5'...T

Set 3. Template M-*ras*: 3'...ACGGTT
Universal primer: 5'...T

Set 4. Template M-*ras*: 3'...CCCTCACCTCCTACGG
Primer 51: 5'...G

Misincorporation reactions were with reverse transcriptase for 30 minutes at 37°C. Additions were as follows: n: no added dNTPs, A: dATP, C: dCTP, G: dGTP, T: dTTP.

Table 1. Sequence and position of primers used for misincorporation. Table 1a shows the sequence of synthetic oligonucleotide primers 1-16 and their complementary positions on the template C6.9. Table 1b shows the 3' hydroxyl termini of the primers and the nucleotide that should be incorporated next.

Table 1a

```
Template     3'-TGCAACATTTTGCTGCCGGTCACTTAAGGGCCCCTAGAGTATGAATGGA
Primer 1     5'-ACGTTGTAAAACGACGG
Primer 2        5'-TTGTAAAACGACGGCCA
Primer 3         5'-TGTAAAACGACGGCCAG
Primer 4          5'-GTAAAACGACGGCCAGT
Primer 5           5'-TAAAACGACGGCCAGTG
Primer 6            5'-AAAACGACGGCCAGTGA
Primer 7             5'-AAACGACGGCCAGTGAA
Primer 8              5'-AACGACGGCCAGTGAAT
Primer 9               5'-ACGACGGCCAGTGAATT
Primer 10              5'-CGACGGCCAGTGAATTC
Primer 11                 5'-ACGGCCAGTGAATTCCC
Primer 12                  5'-CGGCCAGTGAATTCCCG
Primer 13                        5'-GTGAATTCCCGGGGATC
Primer 14                          5'-GAATTCCCGGGGATCTC
Primer 15                            5'-ATTCCCGGGGATCTCAT
Primer 16                             5'-TTCCCGGGGATCTCATA
```

Table 1b

3'-OH terminus of primer	T (A)	G (C)	C (G)	A (T)
A	6	16	2	7
C	14	10	11	13
G	5	1	12	3
T	15	9	4	8

next template nucleotide
first incorporated nucleotide

Km of Klenow polymerase is higher for ddNTPs than for dNTPs and this may also be true for reverse transcriptase. The presence of unannealed primer can be inferred from the dideoxysequencing tracks, using Klenow polymerase, in which all primer-templates present were extended but the degraded material probably represented free primer. Using a single primer each of the three potential misincorporation reactions would show new bands of differing intensities, from which reactions were classified as good, poor or absent (See figures 12 and 13) based on the relative intensities of the misincorporated products with the three substrates. Where a band was absent or faint at the length of primer+1, but other bands were present at longer lengths, the level of misincorporation was inferred from the total intensities of the upper bands (which usually represented correct incorporations after misincorporation).

The assay conditions were such that where a non-complementary dNTP was given, a single misincorporation was usually seen; rarely did two misincorporations occur.

The efficiency (i.e. the level of misincorporation expressed as a percentage of the equivalent level of correct incorporation) of a "good" misincorporation could vary between 10% and 80%, according to the primer, but the level of poor misincorporation was usually less than or equal to half that of a good misincorporation. Discrimination between the misincorporation levels reached by different dNTPs with a given primer was clear and reproducible. Where possible, quantitative estimates were derived by use of a gel scanner.

Figure 12.
Samples of misincorporation reactions analysed by PAGE.

Misincorporation reactions were performed with reverse transcriptase as described in Materials and Methods (8.16). The template was C6.9 and reactions using 15 primers are shown (primer 6 can be seen in figure 3). Additions were as follows: n: no dNTP added, N: correct ddNTP added to give a reference correct incorporation, A: dATP, C: dCTP, G: dGTP, T: dTTP, seq: dideoxy sequencing tracks ordered A,C,G,T. The major band at the bottom is unextended primer.

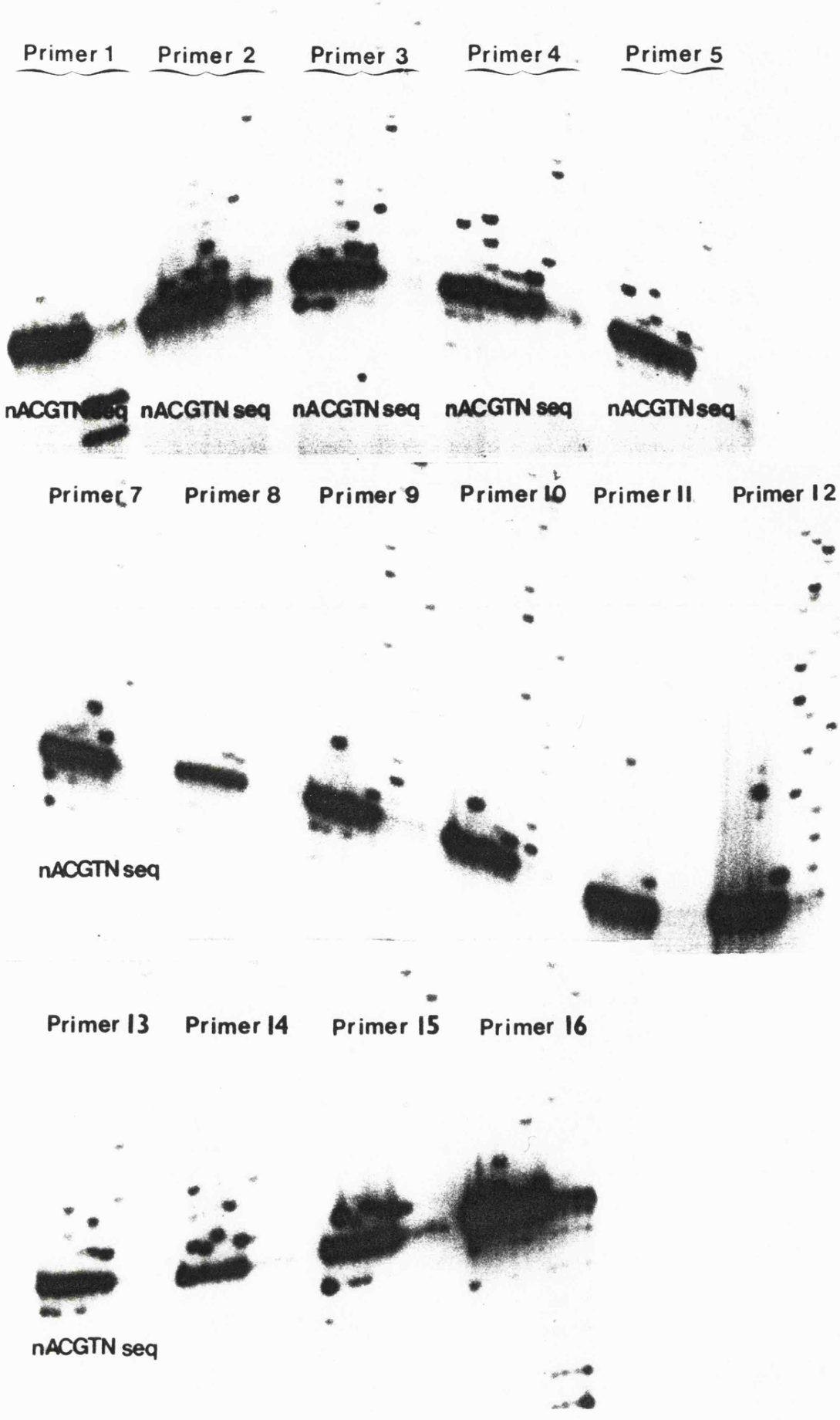
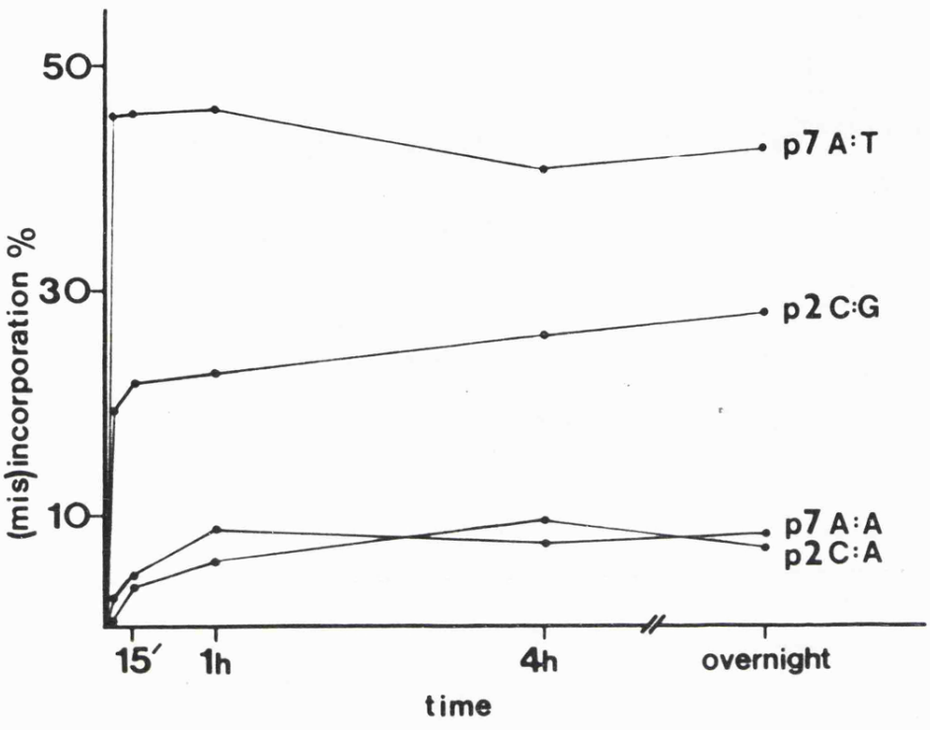
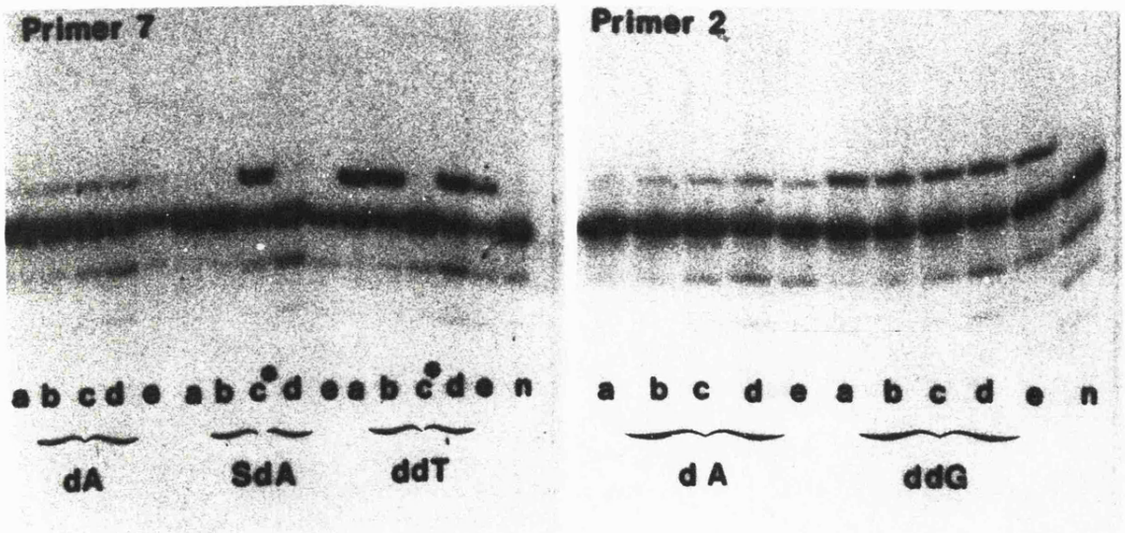

Primer 1
Primer 2
Primer 3
Primer 4
Primer 5
nACGTN seq
nACGTN seq
nACGTN seq
nACGTN seq
nACGTN seq
Primer 7
Primer 8
Primer 9
Primer 10
Primer 11
Primer 12
nACGTN seq
Primer 13
Primer 14
Primer 15
Primer 16
nACGTN seq

Figure 13.
Time course of misincorporation of dATP against a template C (primer 2) and a template A (primer 7).

n: no dNTP added, overnight incubation; a: 5 mins; b: 15 mins; c: 1 hour; d: 4 hours; e: overnight (less sample applied). ddG and ddT show the time course of incorporation of the correct ddNTP.
*: the two samples thus marked (SdA(c) and ddT(c)) were run in the wrong lanes and should be exchanged. Both incubations were for 1 hour. SdA: time course using α-SdATP for misincorporation. The graph shows the level of misincorporation at each time point calculated from the gelscan data as the percentage of the total intensities of each track.


50
30
10
(mis)incorporation %
p7 A:T
p2 C:G
p7 A:A
p2 C:A
15′
1h
4h
overnight
time


Primer 7
a b c d e a b c d e a b c d e n
dA
SdA
ddT
Primer 2
a b c d e a b c d e n
d A
ddG

Quantitative data were used for time course experiments. Each misincorporation was tested four times (with different 3' termini) and was sometimes also seen following an incorporation of the correctly pairing dNTP (denoted with an asterisk in Table 2).

With the 16 different primers a pattern of reverse transcriptase directed misincorporation could be seen that was largely independent of the 3' terminal nucleotide of the primer. This can be seen in Table 2 reading down the columns where, within each block corresponding to the next template nucleotide, the four primers shown have different 3' termini. The patterns of misincorporation (italics) or correct incorporation within the column can be compared. The results can be summarised as follows:

Template nucleotide	Misincorporation of:	
T	C	Always worked; three of six were good, but in one case high efficiency is correlated with the possibility of subsequent correct incorporation ("pull-through" is discussed later).
	G	Always worked to high efficiency.
	T	Twice inefficient. Three cases of high efficiency with subsequent "pull-through"; one other, of poor efficiency, followed a misincorporation.

Table 2.
Observed levels of misincorporation with different combinations of template, primer 3' terminus and substrate nucleotides.

Results of 15 minute misincorporation assays as described in Materials and Methods (8.16). Autoradiographs were exposed for three days. *Italic type* is used for misincorporation reactions which could either be good (***bold italic***) or poor (*standard italic*). Non-italic type is used for correct incorporations. *: these reactions gave a correct incorporation at the first position and any subsequent misincorporation was included in the assessment of preferences for the template nucleotide at which this occurred.

Next template nucleotide		dNTP given in assay	A	C	G	T
Primer 6	T	3' 5'	TTAAG* AA*A*	TTAAG A*C*	TTAAG A*G*	TTAAG A*TTT*
Primer 14		3' 5'	GTATG* C*AA*	GTATG C*C*	GTATG C*G*	GTATG C*TT*
Primer 5		3' 5'	CTTAA* G*AA*	CTTAA G*CC*	CTTAA G*GG*	CTTAA G*T*
Primer 15		3' 5'	ATGAA* T*A*	ATGAA T*CC*	ATGAA T*G*	ATGAA T*T*
Primer 16	G	3' 5'	TGAAT A-	TGAAT* A*CCC*	TGAAT A-	TGAAT A*TTT*
Primer 10		3' 5'	GGGCC C-	GGGCC* C*CC*	GGGCC C-	GGGCC C*T*
Primer 1		3' 5'	CGGTC G-	CGGTC* G*CCC*	CGGTC G-	CGGTC G*T*
Primer 9		3' 5'	AGGGC T-	AGGGC* T*CCC*	AGGGC T-	AGGGC T*T*
Primer 2	C	3' 5'	TCACT A*A*	TCACT A-	TCACT* A*G*	TCACT A*TT*
Primer 11		3' 5'	GCCCC C*A*	GCCCC C-	GCCCCTA* C*GGGGGG*	GCCCC C-
Primer 12		3' 5'	CCCCT G*A*	CCCCT G-	CCCCTA* G*GGGGG*	CCCCT G-
Primer 4		3' 5'	ACTTA T*AAA*	ACTTA T-	ACTTA* T*GGG*	ACTTA T*T*
Primer 7	A	3' 5'	TAAGG A*A*	TAAGGGC A*CCCCC*	TAAGG A*G*	TAAGG* A*TTT*
Primer 13		3' 5'	GAGTA C-	GAGTA C*CC*	GAGTA C*G*	GAGTAG* C*TTT*
Primer 3		3' 5'	CACTT G*A*	CACTT G*C*	CACTT G*GGG*	CACTT* G*T*
Primer 8		3' 5'	AAGGG T-	AAGGGC T*CCC*	AAGGG T*G*	AAGGG* T*T*

Template nucleotide	Misincorporation of:	
G	A	Never worked (five primers).
	G	Never worked.
	T	Failed once, gave poor misincorporation four times, two cases of high efficiency with subsequent "pull-through".
C	A	Good once, poor twice, "pull-through" once
	C	Never worked (six primers).
	T	Failed twice, three poor misincorporations, one good misincorporation with subsequent "pull-through".
A	A	Failed twice, poor four times and one "pull-through".
	C	Three good misincorporations, three were "pull-throughs", one weak second misincorporation.
	G	Once poor, twice good, one "pull-through", one failed, two second misincorporations.

It can be seen that pairing of homologous bases was inefficient except where a subsequent correct incorporation facilitated such a misincorporation. Misincorporation of a dGTP against a template G and of dCTP against a template C had been seen previously (figure 11) but only weakly. The reactions in that experiment were for 30 minutes not 15 minutes and with template M-*ras* which may possibly have

retained some contaminating dNTPs even after purification through Sephadex G100. Purine misincorporation against a purine template was completely dependent on the particular combination: dATP was never misincorporated against G in the template, whereas dGTP could be misincorporated against A in the template. Pyrimidine:pyrimidine mismatches behaved likewise: dTTP was misincorporated poorly against C in the template, whereas dCTP could be misincorporated quite well with T in the template. Such dependence on the exact combination was also observed with purine:pyrimidine misincorporations. Although these were usually more efficient than most purine:purine or pyrimidine:pyrimidine pairs, dTTP was only weakly incorporated against G in the template.

No consistent correlation was observed between the extent of misincorporation and the nucleotide at the primer terminus.

Secondary structure of the template was not analysed systematically, but it was noted that primers 8 and 9 gave generally poor misincorporation which may be influenced by the adjacent template sequence (3'AGGGCCCCT, and 3'GGGCCCCT respectively) potentially folding to give different secondary structure.

The results raised several intriguing questions. Firstly, it was not clear why some misincorporations were less complete than others. Three possible explanations were considered:

1). different rates of forward reaction;

2). a previously undetected or contaminating exonuclease activity which preferred certain misincorporations, thus resulting in a low steady state level of extended primer;

3). mispaired termini may be subject to an enhanced rate of reversal of reaction, i.e. an altered equilibrium level of product. With the high molar ratios of enzyme:primer used here there may even be an equilibrium of enzyme-bound substrate and product.

Secondly, in several cases the effect of the next template nucleotide has been referred to: subsequent correct incorporation appeared to enhance effective misincorporation, described as a "pull-through" reaction. Both these aspects were investigated.

9.1.5 Kinetics of Misincorporation.

Time course assays of poor misincorporation of dATP using primers 2 and 7 were performed with reference to a ddNTP incorporation as mentioned before.

The results shown in figure 13 demonstrate that the level of incorporation reached a plateau in about five minutes or less, whereas misincorporation reached a much lower plateau after an hour. Thus the rate of misincorporation appears to be lower, but longer times do not allow the reaction to approach completion. Incidentally, of the three dNTPs that could be misincorporated with primer 7, dATP was the poorest substrate, whereas with primer 2

dATP was classed as a good substrate (Table 2). This reflects the very different characteristic efficiences of misincorporation of different primers (see Discussion).

The misincorporation reaction was also carried out by using α-SdATP. Phosphorothioate linkages in DNA chains are resistant to cleavage by many nucleases, notably the 3'-5' exonuclease of polymerase I and exonuclease III (Putney *et al.*, 1981; Kunkel *et al.* 1981; Eckstein, 1985). Although the α-SdATP was misincorporated more slowly than dATP (when successful, sometimes it failed to be misincorporated at all), it could be seen that the extent of misincorporation increased with time in the same manner as it did with dATP as substrate . These results indicate that the low levels of misincorporation achieved in comparison to those of correct incorporations are due to lower rates of misincorporation and not to any exonuclease activity preferentially attacking misincorporations, or to vastly altered equilibrium constants for mispaired termini.

However, this does not provide any explanation for the low plateaux for misincorporations. These may reflect enzyme inactivation or depletion of substrate by a contaminating nucleotidase during long time courses. To investigate this, misincorporation of dATP with primer 7 was allowed to proceed for five hours followed by subsequent additions of more substrate, more enzyme, or more annealed primer-template (Table 3 and figure 14). After five hours no further incorporation or misincorporation occurred, and the amount of misincorporation achieved was one third that of

Figure 14.
Identification of limiting factors affecting the level of misincorporation.

A comparison of dATP misincorporation and ddTTP incorporation with primer 7. a and g: 5 hour incubation, 10 units reverse transcriptase (r.t.); b and h: 22 hour incubation, 10 units r.t.; c and i: 22 hour incubation, 20 units r.t.; d and j: 5 hour incubation, 10 units r.t., then 17 hour incubation with additional 10 units r.t.; e and k: 5 hour incubation, 10 units r.t., then 17 hour incubation with extra 0.1mM substrate; f and l: 5 hour incubation, 10 units r.t. with non-labelled primer-template, then 17 hour incubation with extra [^{32}P]-labelled primer-template; m: negative control, 22 hour incubation, 10 units r.t., no dNTP; A: the first template nucleotide.

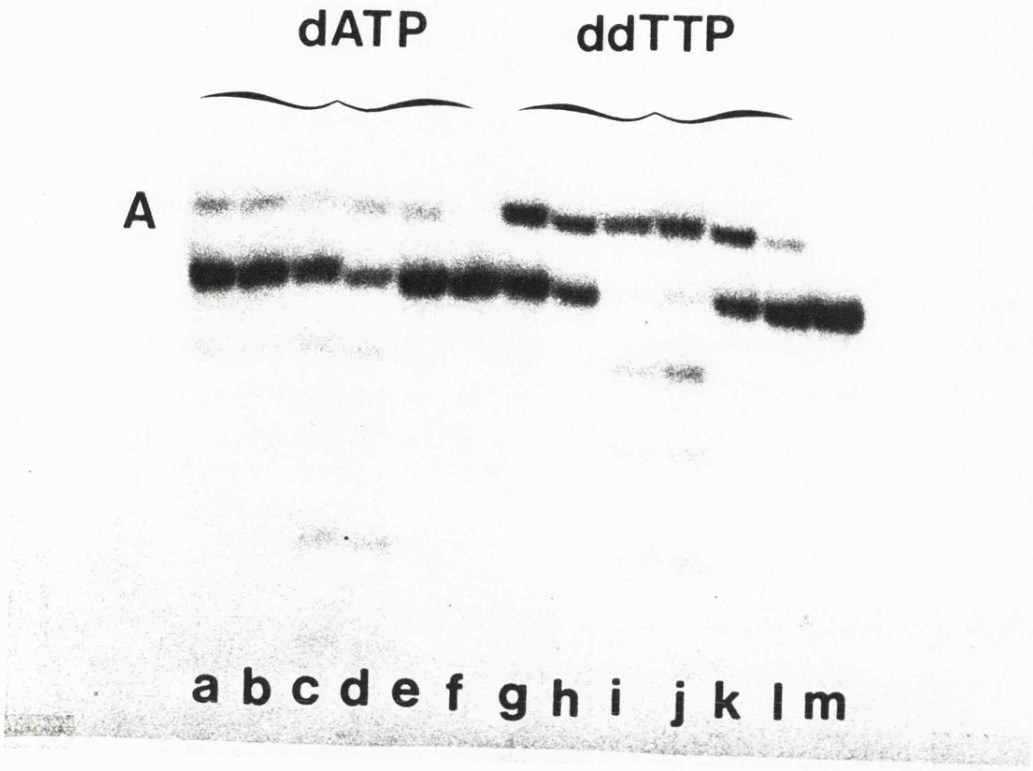


Table 3. Identification of limiting factors affecting the level of misincorporation: analysis of data from Figure 14.

Track	dNTP	Condition	Enzyme	% Intensity of Bands					
				-4	-3	-2	-1	Primer Band	+1
a	dA	5h	10U				4.2	81.9	13.9
b	"	22h	10U				4.3	82.2	13.5
c	"	22h	20U	5.2	12.7	6.4	9.0	59.7	7.0
d	"		10U+10U	6.2	14.1	8.3	11.4	47.5	12.5
e	"	+ substrate	10U				3.5	84.9	11.6
f	"	+ anneal	10U					100	0
g	ddT	5h	10U				2.4	57.6	40.0
h	"	5h	10U				1.2	56.0	42.8
i	"	22h	20U		9.1	8.8	18.6	11.9	51.7
j	"		10U+10U		9.5	9.2	22.0	13.0	46.4
k	"	+ substrate	10U				5.3	54.7	40.0
l	"	+ anneal	10U					82.4	17.6
m	negative	22h	10U				4.8	95.2	0

correct incorporation. Addition of extra substrate at this point and incubation for a further 17 hours gave no additional misincorporation or incorporation, indicating that substrate depletion had not caused the termination of the reactions. However, after allowing the misincorporation reaction to proceed for five hours with unlabelled primer-template, and then adding fresh 5'-labelled primer-template and incubating for a further 17 hours (track f), no misincorporation on the 5' labelled primer occurred, indicating that the reverse transcriptase had become inactive. The parallel experiment with correct incorporation of ddTTP (track 1) showed that after the incubation for 17 hours with fresh primer-template the level achieved was only about 40% of that normally observed (track g); note that normally correct incorporation reactions are complete within five minutes.

The factors responsible for the low plateaux for misincorporations were not elucidated by attempts to increase the efficiency of misincorporation. Adding fresh aliquots of enzyme after five hours (figure 14, tracks d and j) or increasing the initial dose of enzyme (tracks c and i) led to much increased degradation. Under normal conditions, where there is a three-fold molar ratio of enzyme to primer, there is very little nuclease activity even in the absence of dNTPs (track m). There is some evidence for a trace of 3' exonuclease activity in the reverse transcriptase from experiments where the correct 3' terminal nucleotide of an unlabelled primer has been exchanged for a labelled

nucleotide on incubation with reverse transcriptase (figure 8). Misincorporation, unlike ddNTP incorporation, does not protect the primer against the 3' exonuclease activity in the presence of very high levels of enzyme, (compare tracks c, i, and j of figure 14). It may be inferred from track d that the addition of the extra aliquot of enzyme does not enhance misincorporation (the ratio of product to primer is doubled).

It was concluded that the rate of misincorporation is very low in comparison to the rate of incorporation, and that the level of misincorporation is limited by the progressive inactivation of reverse transcriptase. The addition of extra enzyme however, for reasons not entirely understood, greatly increased the degradative activity present.

9.1.6 "Pull-through" misincorporation.

In some cases misincorporations which, on the basis of their template and substrate nucleotides, would be predicted to be poor were nevertheless good, and these occurred when a subsequent correct incorporation was made. It appeared that this correct base pairing facilitated the otherwise unfavourable misincorporation. Examples can be seen with primer 6 plus dTTP, primer 2 plus dTTP, primer 14 plus dATP and primer 14 plus dTTP (Table 2).

Previous experiments with DNA polymerase I have shown a "next-nucleotide" effect where higher misincorporation

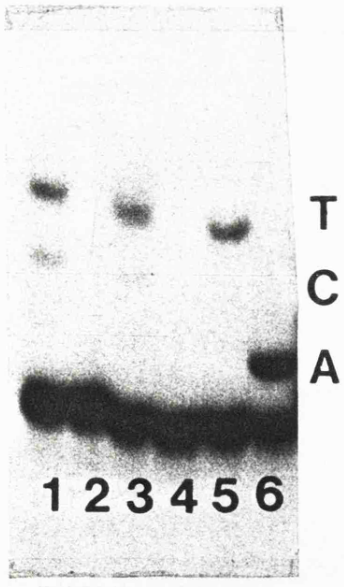
mutation rates were achieved by the additional presence of a nucleotide complementary to the second template position in the misincorporation reaction (Kunkel *et al.*, 1981). The subsequent incorporation protected the misincorporation from the 3' exonuclease proofreading activity that acted preferentially on misincorporations. Reverse transcriptase has no associated proofreading activity (Kunkel *et al.*, 1981) so other possible explanations of "pull-through" have to be considered. These possibilities are: a) that the affinity of reverse transcriptase for the substrate dNTP may be greater if it can simultaneously bind a second dNTP complementary to the second template nucleotide, thus the binding of the second dNTP is sufficient to give "pull-through"; or b) that polymerization between the first and second nucleotides stabilizes misincorporation. To distinguish between these two possibilities dideoxy NTPs were used in a misincorporation reaction to show whether it was necessary to form a phosphodiester link between positions 1 and 2 to allow "pull-through". Figure 15 shows that with primer 3 (template 3'ACT) the presence of dGTP permitted misincorporation of G against the template nucleotide A, but no band appeared at that position as subsequent incorporation and misincorporation gave strong bands of longer lengths. Misincorporation of ddATP or ddGTP would have given a band of primer+1, but in neither case was such a band seen, either with or without the presence of dGTP (complementary to template position 2). As ddGTP alone gave no misincorporation (track 2) whereas dGTP did (track

Figure 15.
Experiments to investigate mechanisms of facilitation of misincorporation by a subsequent correct incorporation ("Pull-through").

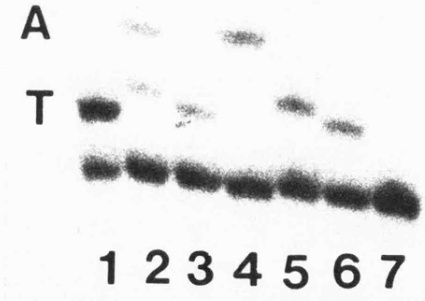
```
Template C6.9: 3'...CACT
     Primer 3: 5'...G

Template C6.9: 3'...GTAT
    Primer 14: 5'...C
```

Additions were as follows:
Primer 3: 1: dGTP, 2: ddGTP, 3: dGTP + ddGTP, 4: ddATP, 5: dGTP + ddATP, 6: ddTTP (incorporation).
Primer 14: 1: ddATP (incorporation), 2: dGTP + ddTTP, 3: ddGTP + dTTP, 4: dGTP + dTTP, 5: dGTP, 6: ddGTP, 7: no dNTP.



Primer 3



Primer 14

1), polymerization had to occur for misincorporation to be enhanced. This is consistent with either protection of the misincorporated nucleotide from a proofreading exonuclease activity or with a mechanism of elongation where insertion of a mismatched base is linked with the formation of a subsequent phosphodiester bond.

A further misincorporation was performed with primer 14 (template 3'TAT) in the combined presence of dTTP and dGTP (figure 15). Each dNTP alone worked well; a subsequent correct incorporation with dTTP permitted the usually unfavourable dTTP:T misincorporation. However, the total percentage of product using dGTP (track 5) equalled that of dGTP + dTTP (track 4) indicating that the presence of dTTP did not improve the dGTP:T misincorporation.

Experiments with primer 6, primer 2 and primer 14 (Table 2) have shown that misincorporations of different predicted efficiences can be facilitated by the next nucleotide. This phenomenon appears to be restricted to situations where the second nucleotide added is the same as the misincorporated nucleotide; several experiments (as with primer 14 in figure 15) have been unable to demonstrate facilitated misincorporation with heterologous dNTPs.

9.2 Generation of random short extensions from an oligonucleotide primer.

A 17 nucleotide primer was annealed to an M13 SS DNA template. This was then extended in a reaction similar to a

dideoxy sequencing reaction using DNA polymerase I (Klenow fragment), three dNTPs at a standard concentration and the fourth dNTP at a limiting concentration. For analytical reasons the limiting dNTP was α-^{32}P-labelled which also facilitated subsequent separation of extended primers from free nucleotides. After the extension reaction the products were analysed on denaturing polyacrylamide gels as shown in figure 16. By comparison of the extended primers to sequencing tracks using ddNTP (N being the nucleotide limited in the extension) it could be seen that many of the products of the extension had their 3' termini at one position prior to an expected N incorportion. The range of lengths of primers was from between 17 nucleotides and 150 nucleotides, achieved by altering the conditions of the short extension reaction (figure 17).

9.2.1 Removal of dNTPs from primer-templates.

Where primers had been extended with dNTPs, the primer-templates had to be purified to remove the remaining contaminating dNTPs so that the ensuing misincorporation reaction could generate specific misincorporaions efficiently. A number of different methods were tried, separating dNTPs from primers of 17 or more nucleotides by virtue of their different molecular weights.

The success of any purification was measured by incubating the product as annealed primer-template with reverse transcriptase in the presence and absence of the

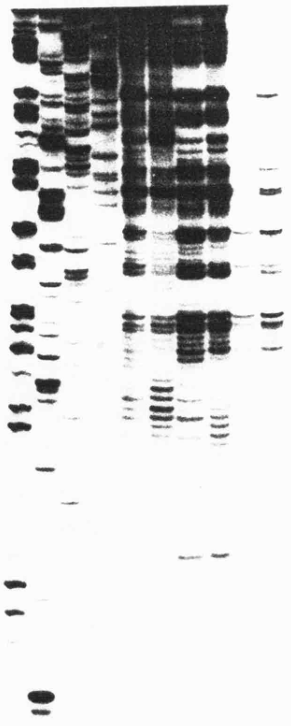
Figure 16.
Generation of limited short extensions of an oligonucleotide primer.

Extensions were from primer 51 annealed to template M-*ras*, for 3 minutes at room temperature with 100nM [^{32}P]-dATP and 40µM each of dCTP, dGTP and dTTP. A: ddATP sequencing track (showing template T positions). The molar ratios of template to primer shown in tracks 1 to 4 were as follows: 1=1:10, 2=1:5, 3=1:2.5, 4=1:1. Track 5: as track 1 but with the unlabelled dNTPs at 4µM each. Track 6: as track 1 but with all the dNTP concentrations reduced by 50% on doubling the reaction volume. Most of the extension products can be seen to have terminated before a template T.

Figure 17.
Conditions for different length primer extensions.

Extensions were from primer 51 annealed to template M-*ras*, for two or five minutes with decreasing concentrations of α-^{32}PdATP, with Klenow polymerase. A,C,G,T: dideoxysequencing reacions. 1: 1.0µM dATP, 2 min.; 2: 1.0µM dATP, 5 min.; 3: 100nM dATP, 2 min.; 4: 100nM dATP, 5 min.; 5: 10nM dATP, 2 min.; 6: 10nM dATP, 5 min.


A 1 2 3 4 5 6


A C G T 1 2 3 4 5 6

appropriate dNTP. When incubated with reverse transcriptase, extended primers free of contaminating dNTPs showed no differences on gel analysis, but when the appropriate dNTP was also present then this was incorporated at the 3' end of each primer which showed as a shifted band on the autoradiograph of the polyacrylamide gel (figures 18 and 19).

Sephadex G100 columns (1ml) were most frequently used to remove free dNTPs whilst keeping the primer and template annealed by running the column in 10mM Tris-HCl pH7.5, 5mM $MgCl_2$. When an α-^{32}P labelled dNTP had been included in the mixes for the extension of the primer, the separation of the primer-template and the free dNTPs could be followed by use of a mini-monitor Geiger counter. Successive fractions could also be collected and analysed in a scintillation counter.

In some experiments there were problems of dNTPs contaminating the primer-template even after filtration through Sephadex G100 (figure 18). Sephadex G50 spun columns were tried because these are much more rapid than conventional columns but contamination proved to be a greater problem than with the G100 and furthermore, the recovery of primer-template was also worse than with conventional G100 columns. NENsorb 20 cartridges were tried but these proved worse than spun G50 columns with contaminating dNTPs still present even after repeated washings. A comparison of DNA purified with a NENsorb 20 cartridge or Sephadex G100, involving generation of misincorporation mutants, is described in section 9.3.2.

Figure 18.
Attempted purification of short primer extensions from free dNTPs.

Primer 24 was annealed to template mp8 and extended in a short reaction with limiting dATP. Most of the products of the extension (E) terminated prior to a template T (seen in the ddATP sequencing track A). After purification through Sephadex G100, incubation of the primer-templates with reverse transcriptase (E RT) gave further extensions, indicating that dNTP impurities were still present. a,c,g,t: misincorporation reactions with dATP, dCTP, dGTP or dTTP respectively.

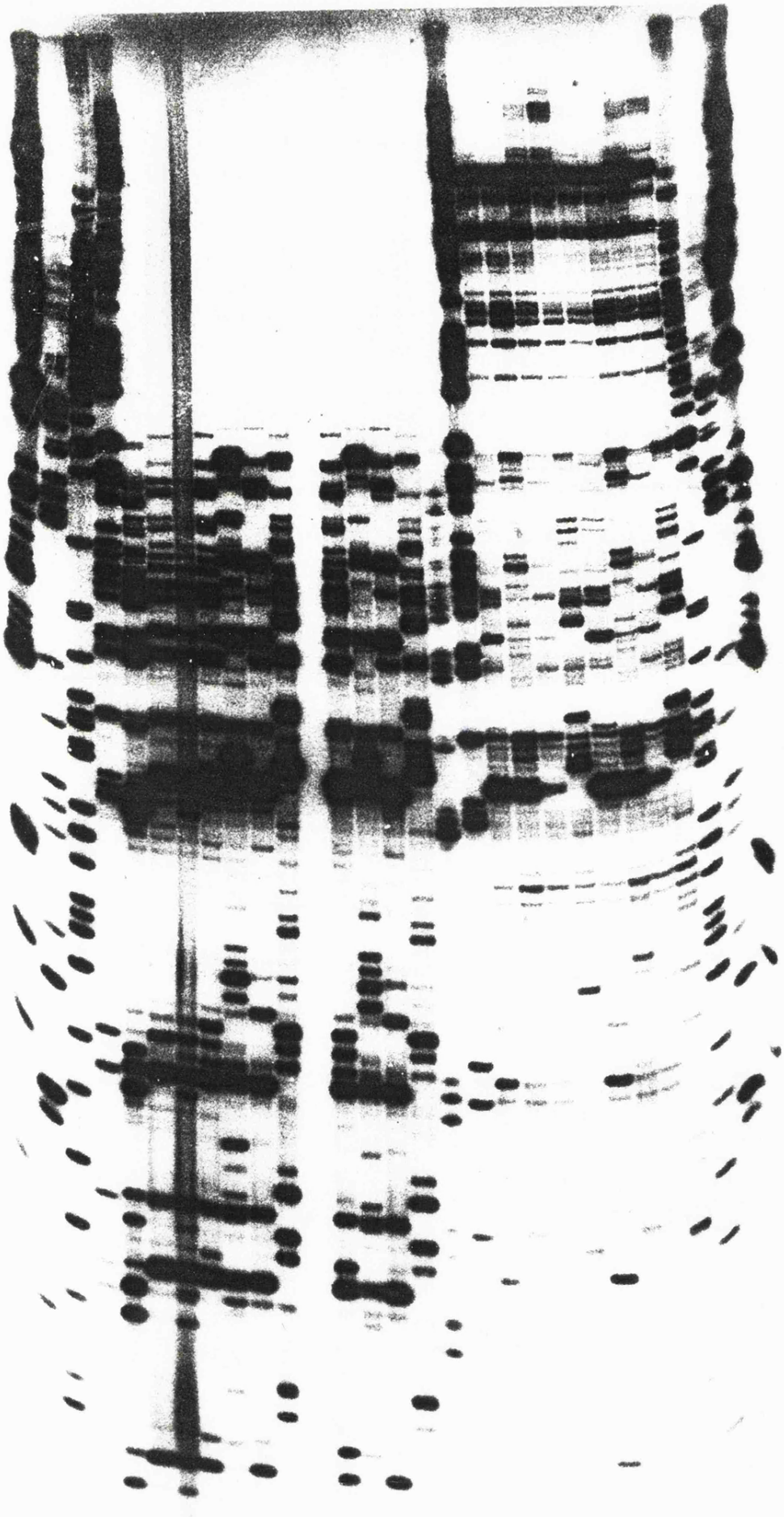
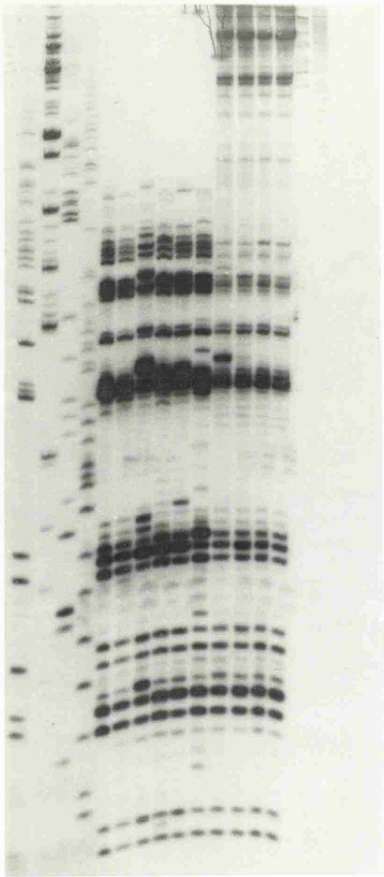

CGTA
E E E a c g t a c g t E A a c g t
R T R T

Figure 19.
Purification of short primer extensions: removal of dNTPs.

Primer 24 was annealed to template mp8 and extended in a short reaction by Klenow polymerase in the presence of all four dNTPs, the concentration of dATP limiting the extensions (E). Purification through Sephadex G100 was successful: incubation of the primer-templates with reverse transcriptase (E RT) showed no further extension. A,C,G,T: dideoxysequencing tracks. a: misincorporation of dATP; c,g,t: attempted misincorporations of dCTP, dGTP, dTTP. a,c,g,t-chase: after the misincorporation reaction the products were extended by a reverse transcriptase-directed chase reaction with all four dNTPs to "seal-in" any mismatches.


A C G T E E a c g t a c g t
RT
chase

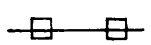
After G100 purification, 20% of the progeny recovered were misincorporation mutants whereas no such mutants were recovered after NENsorb purification suggesting that the remaining contaminating dNTPs prevented misincorporation.

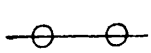
Variations in the procedure for purifying primer-templates through G100 were made in an attempt to discover why contaminating dNTPs sometimes remained. The separation of primers from free dNTPs might be inadequate if the column fractions containing dNTPs overlapped with those containing primer. For this reason the separation of primer-templates, primer and free dNTPs on 1ml G100 columns was investigated. The fractions from 1ml G100 columns which contained annealed primer-template were determined by using 5' ^{32}P-labelled universal primer annealed to SS C6.9 template. The results are shown in figure 20 and demonstrate that annealed primer-templates came off the column when the bromophenol blue dye was between 1/3 and 1/2 of the way through the column, and that free primer came off the column shortly after annealed primer-templates. Free dNTPs came off columns last. Thus, in order to give better separation of primer from free dNTPs it was decided to pass it through the column whilst annealed to a SS DNA template.

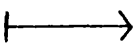
It was considered that the presence of Klenow polymerase, necessary for initial extensions of primers, might be hindering the removal of free dNTPs and so in some cases phenol extraction was performed prior to column loading. Four 1ml G100 columns were run as shown in figure 21 with annealed primer-template plus Klenow

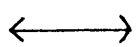
Figure 20.
Sephadex G100 column fractions containing either free primer or primer-template.

Two Sephadex G100 columns were run, samples of three drops each were collected and the radioactivity (c.p.m.) was monitored by Cherenkov counting in a scintillation counter.


Universal primer labelled at the 5' end with γ-^{32}P-ATP.


5' labelled universal primer annealed to SS DNA template (primer-template).


Bromophenol blue came off the columns in these fractions.


Fractions in which DNA was expected *i.e.* when bromophenol blue was between the 1/3 and 1/2 way point.

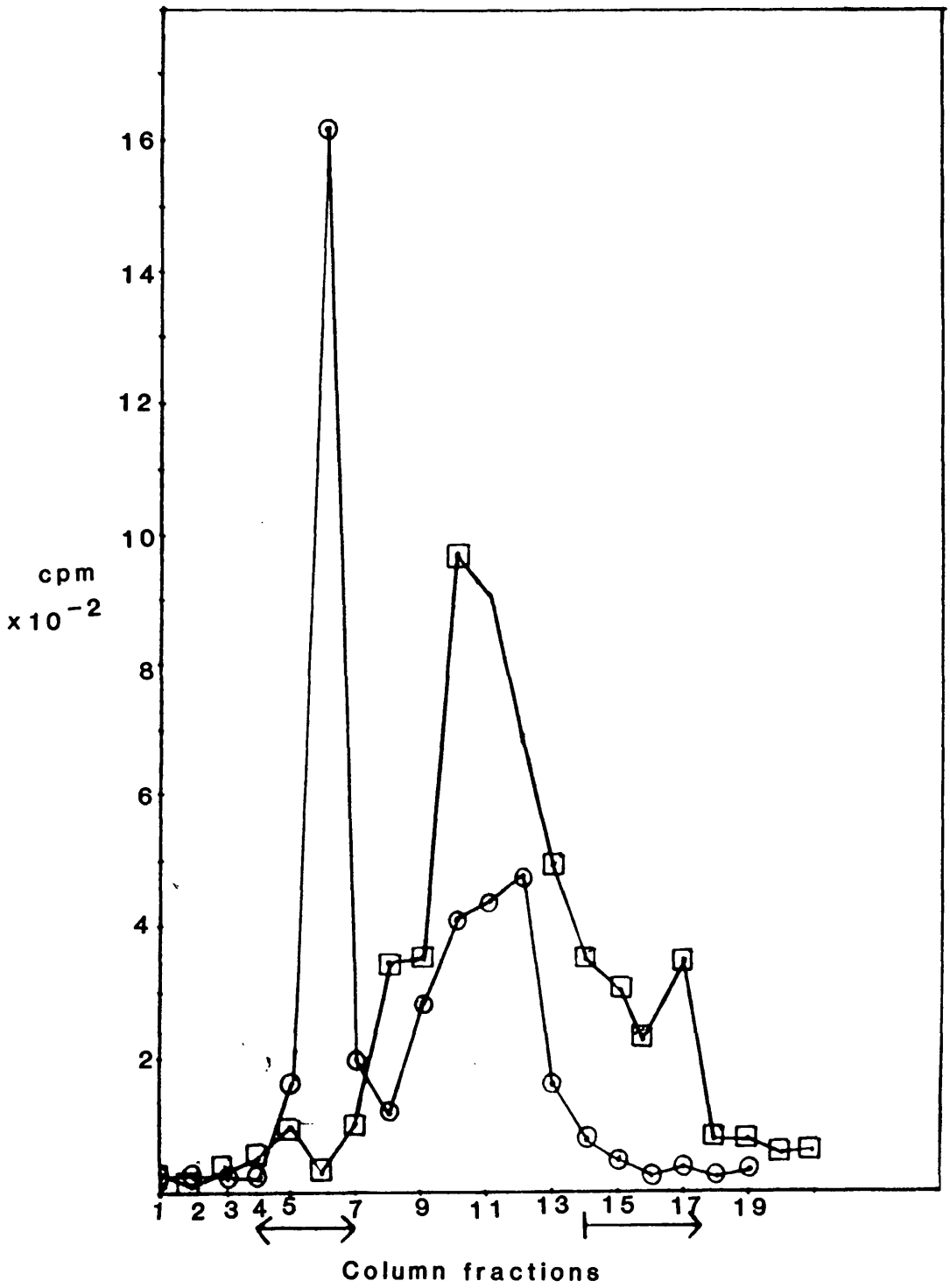

16
14
12
10
8
6
4
2
cpm
x10⁻²
1 2 3 4 5 7 9 11 13 15 17 19
Column fractions

Figure 21.
Sephadex G100 column fractions containing free dNTPs after incubation with primer-template and denatured Klenow polymerase.

Primer-template and Klenow polymerase were incubated together for five minutes and then the polymerase was denatured for ten minutes at 70°C before placing on ice.

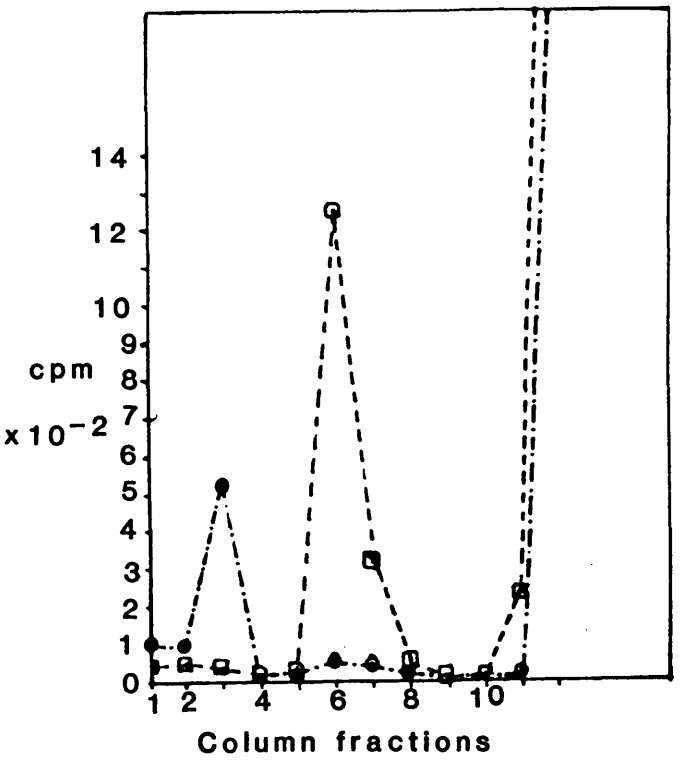
I -□-□- α-^{32}PdNTPs were added, the mixture reannealed and then loaded onto the column.

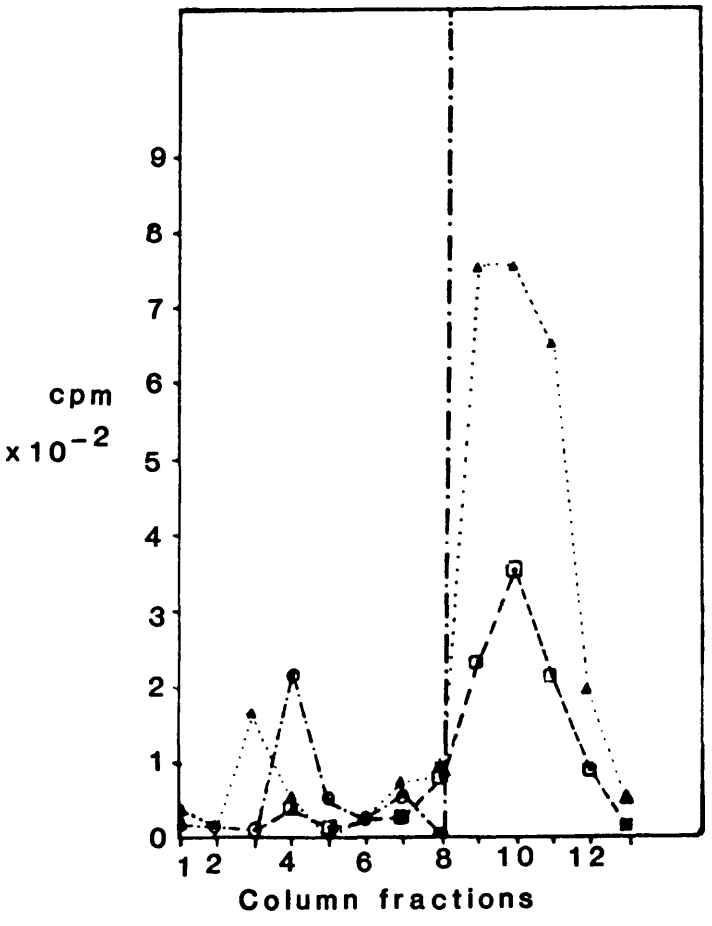
-○-○- The mixture was phenol extracted, α-^{32}PdNTPs added, reannealed and then loaded onto the column.

II -○-○- As in I above.

·▲···▲· α-^{32}PdNTPs were added, the mixture was phenol extracted, reannealed and then loaded onto the column.

-□-□- As ··▲··▲· above but with unlabelled dNTPs added to 0.3mM and a trace of α-^{32}PdNTPs.


cpm
x10⁻²
14
12
10
9
8
7
6
5
4
3
2
1
0
1 2 4 6 8 10
Column fractions


cpm
x10⁻²
9
8
7
6
5
4
3
2
1
0
1 2 4 6 8 10 12
Column fractions

polymerase. Radioactively labelled dNTPs were added after heat denaturation of the polymerase either before or after phenol extraction, so that no incorporation of the dNTPs could occur. Absolute separation of primer-templates and dNTPs would result in the primer-template recovered from the column containing no detectable α-^{32}P-dNTP. The fraction from the columns in which DNA would have been expected showed contamination with dNTPs in all cases. It could be conjectured that, in those cases where no phenol extraction had been performed or where α-^{32}P-dNTPs had been added before phenol extraction, the Klenow polymerase had retained some of its activity and so labelled dNTPs had been polymerised onto the primer. However, when α-^{32}P-dNTPs were added after heat inactivation and phenol extraction this could not apply and so it must be assumed that some free dNTPs were bound up within the "DNA fraction". How significant such levels of contaminating dNTPs might be on subsequent attempts at misincorporation remained unknown as problems with contamination after such G100 purification had been irregular. For this reason, the following experiment was performed in order to examine whether phenol extraction of the template-primer-DNA polymerase mixes was necessary (prior to purification through Sephadex G100) to avoid dNTP contamination in subsequent incorporation or misincorporation reactions.

A 5'-^{32}P-labelled primer was annealed to template DNA and incubated with Klenow polymerase or reverse transcriptase before heat inactivation and addition of

dNTPs. Half of the primer-template was phenol extracted and then both samples were purified through Sephadex G100 columns. The primer-template fractions from the columns were incubated with and without reverse transcriptase and the products run on an analytical polyacrylamide gel. No extension of the labelled primer with contaminating dNTPs was seen whether or not the polymerase had been removed by phenol extraction prior to loading onto the G100 columns. This indicated that the phenol extraction was not necessary in order to eliminate levels of dNTP contamination that would reduce the efficiency of the required misincorporation.

One other method of removing free dNTPs from extended primers has been used. By running the short extension products into a denaturing poyacrylamide gel the fast running dNTPs were separated from primers of 17 or more nucleotides, the positions of which could be established by using ^{32}P-labelled primers or α-^{32}P-dNTPs in the extension reaction. The extended primers were eluted from the polyacrylamide gel slice, ethanol precipitated and were then annealed to fresh template for subsequent misincorporations. This procedure was much lengthier than G100 purification, and resulted in much greater losses. It was only used once, where large amounts of material had been used for the inital short extensions and, before use for misincorporations, the extended primers were annealed to template DNA and assayed within a copying reaction to check the concentration and ability to form CC DNA. The results and analysis of such

purification of primers are shown in figure 22; incubation of the extended primers with reverse transcriptase showed no incorporation implying that they were free of contaminating dNTPs.

9.2.2 Misincorporation on random-extension primers.

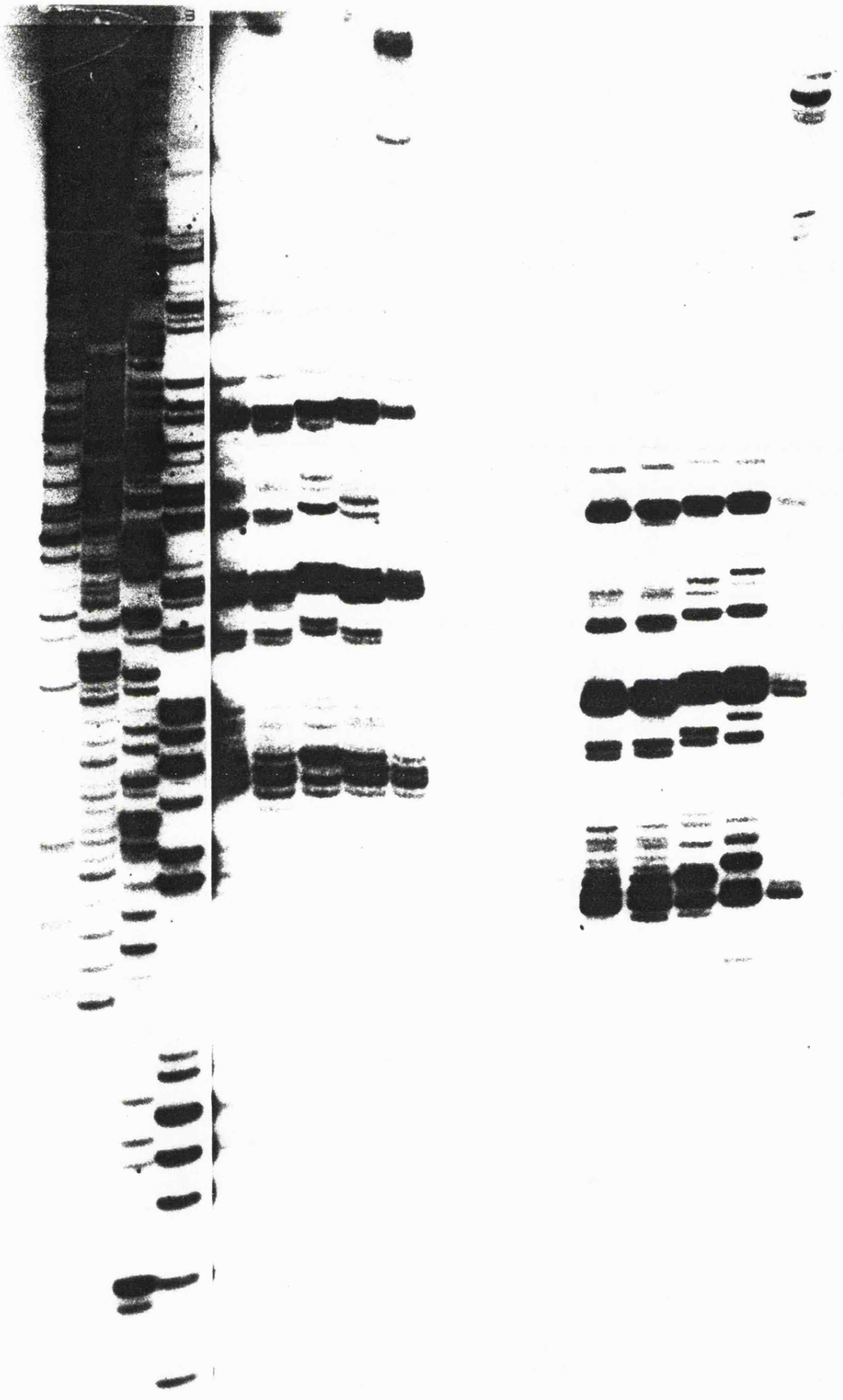
When free dNTPs had been removed from random short extensions of primers, these primers could be used for misincorporation. An example is shown in figure 22 where dTTP and dGTP were misincorporated against template T residues after a dATP-limited initial extensions, and another example is shown in figure 19. Gel analysis of such misincorporations showed a change in the pattern of extended primers seen before and after misincorporation. However, if α-SdNTPs were used for misincorporation then any primers remaining unextended after such a misincorporation reaction could be removed by digestion with exonuclease III. Figure 23 shows that incorporations of α-SdATP and misincorporations of α-SdTTP were resistant to exonuclease III digestion. Contamination of α-SdNTPs with other α-SdNTPs, however, would result in exonuclease III resistant products which might not, in fact, be misincorporations.

9.2.3 "Chase" reactions after misincorporations.

Following a misincorporation reaction, the rest of the template was copied to make double-stranded molecules for

Figure 22..
Misincorporation using a bank of extended primers.

Primer 51 was annealed to template M-*ras* and extended in a dATP-limited reaction. The extended primers were purified by elution from a polyacrylamide gel, and were then reannealed to fresh template M-*ras* for misincorporation reactions. A,C,G,T: dideoxysequencing tracks of M-*ras* from primer 51. 1 = extended primers eluted from the preparative gel; 2 = 1 + reverse transcriptase, shows no contaminating dNTP presence; 3 = 2 + ddATP, correct incorporation; 4 = 2 + dTTP, misincorporation; 5 = 4 + chase; 6 = 1; 7 = 2; 8 = 3; 9 = 7 + dGTP, misincorporation; 10 = 9 + chase.


TGCA 1 2 3 4 5
6 7 8 9 10

transfection into *E. coli*. If an α-SdNTP had been misincorporated, then there could be no *in vitro* repair of the mismatch by the 3'-5' exonuclease activity of Klenow polymerase. (Klenow polymerase was used for second strand synthesis because it is more accurate and causes less damage to SS DNA.) However, a normal nucleotide that had been misincorporated could be repaired on addition of Klenow polymerase. To avoid this problem it was decided to "seal-in" any mismatch by allowing extension with all four dNTPs and reverse transcriptase for a short time so that the mismatch would no longer be accessible to Klenow's 3'-5' exonuclease. Trial reactions were analysed on gels to see the extent of the chase reaction. Examples of this can be seen in figures 23 (gel 3) and 24.

9.3 Generation of mismatch mutants.

9.3.1 Oligonucleotide mutagenesis.

In an attempt to establish correct conditions for second strand synthesis (copyings) of a SS M13 template, a synthetic oligonucleotide primer (primer 24) was used which introduced a one base deletion into the sequence of mp8. This inactivated the β-galactosidase gene and so plaques arising from the second (mutant) strand were white in the presence of BCIG and IPTG in contrast to the blue wild-type plaques.

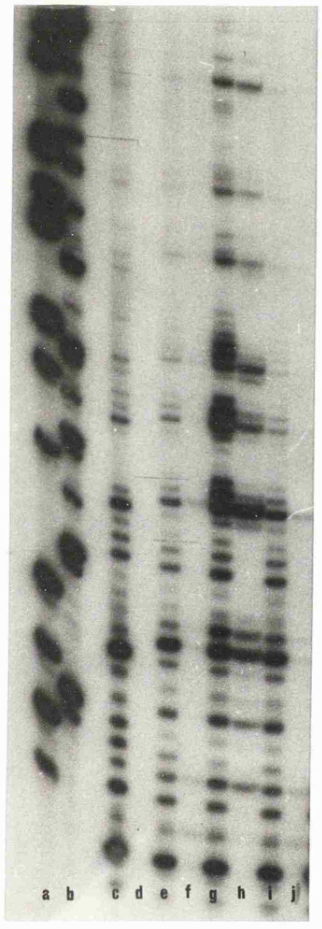
Annealing of primer to template was carried out at a

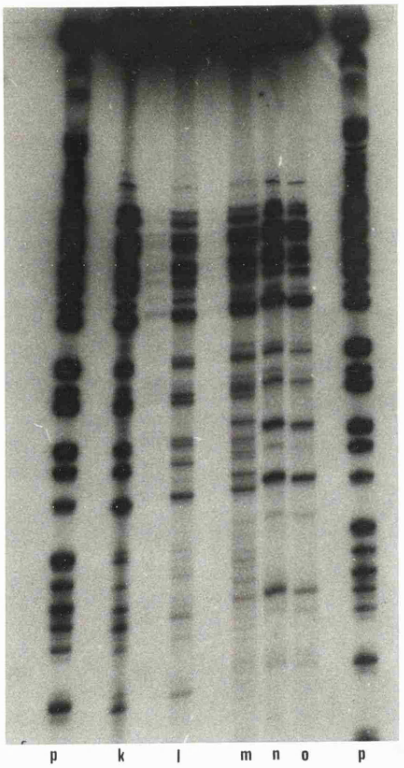
Figure 23.
Incorporation and misincorporation of α-SdNTPs on extended primers: the removal by exonuclease III of primers without phosphorothioate linkages.

Gel 1. a: ddATP sequencing; b: ddCTP sequencing; c: dATP-limited extension + reverse transcriptase; d: c + exonuclease III; e: c + α-SdATP (correct incorporation), 10 min.; f: e + exonuclease III; g: e, 30 min.; h: g + exonuclease III; i and j: e and f with a different batch of reverse transcriptase.

Gel 2. p: ddATP sequencing; k: dATP-limited extension; l: k + reverse transcriptase; m: l + α-SdATP (correct incorporation); n: m + exonuclease III, 10 min.; o: m + exonuclease III, 30 min.

Gel 3. A comparison of exonuclease III treatment of α-SdTTP and dTTP misincorporations. C,G,T,A: dideoxysequencing reactions. Unlabelled tracks left to right: dATP-limited extension (ext), ext + reverse transcriptase (rt), ext + rt + α-SdATP (correct incorporation), ext + rt + α-SdTTP (misincorporation), ext + rt + dTTP, ext + rt + α-SdATP + exonuclease (exo) III, ext + rt + α-SdTTP + exo III, ext + rt + dTTP + exo III, ext + rt + α-SdATP + chase, ext + rt + α-SdTTP + chase, ext + rt + dTTP + chase, blank.


a b c d e f g h i j


p k l m n o p

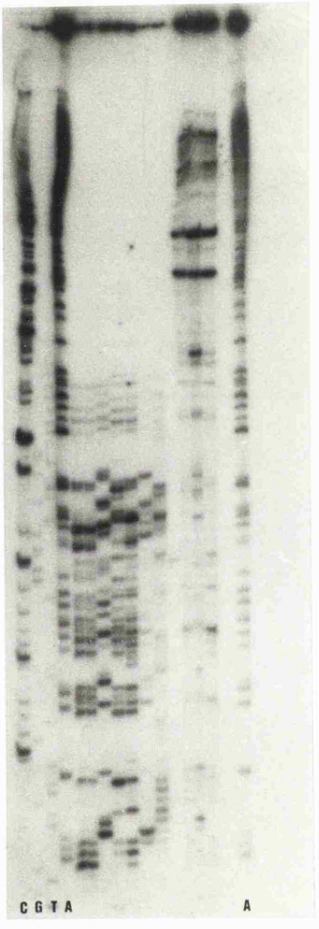
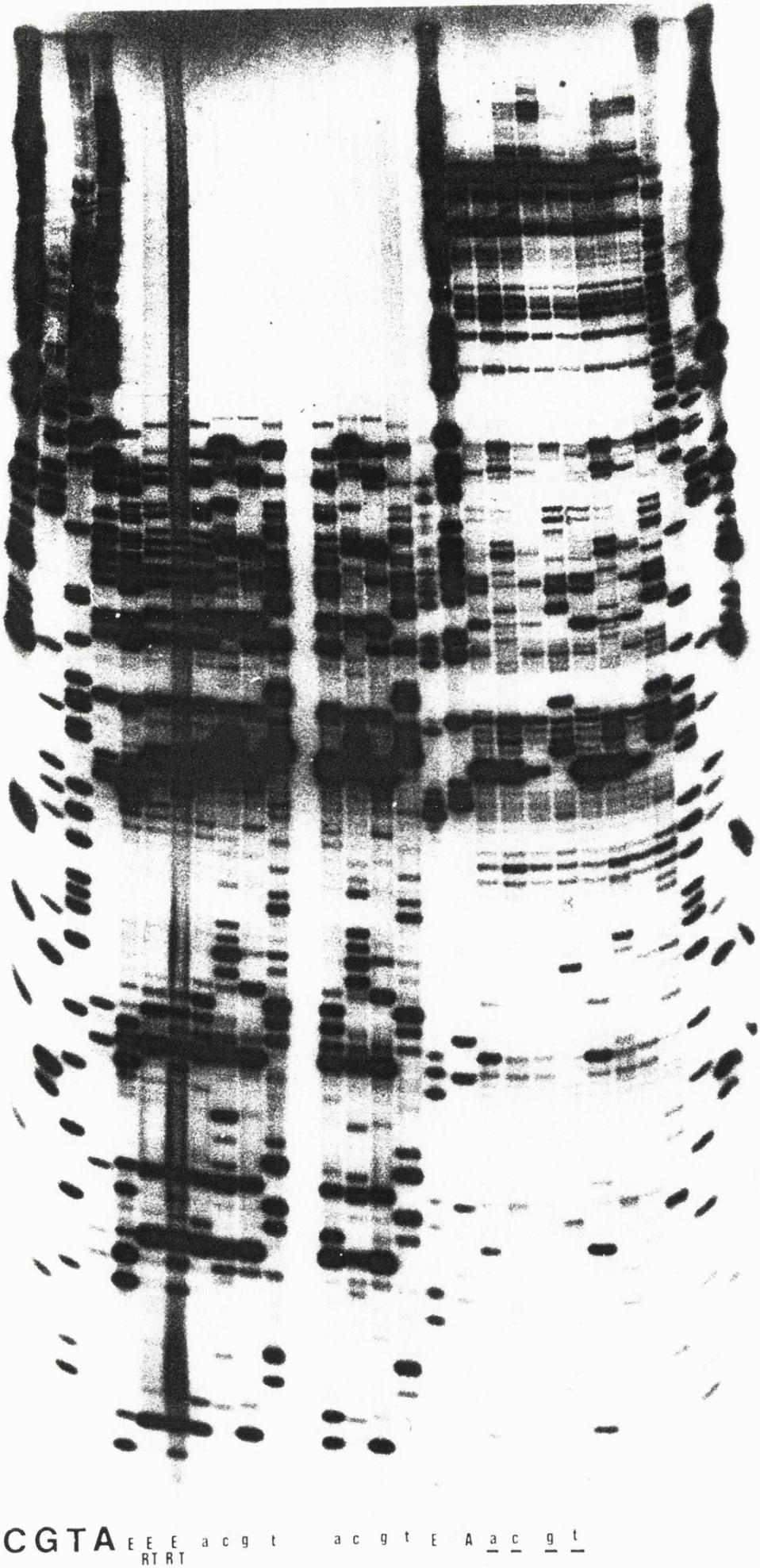

C G T A A

Figure 24.
Chase reactions of misincorporations.

Primer 24 was annealed to template mp8 and extended in a dATP-limited initial extension. a,c,g,t: misincorporation reactions with dATP, dCTP, dGTP and dTTP respectively. E: dATP-limited extension. A: dideoxyATP sequencing track. a , c , g , t : 5 minute chase reactions using products of the misincorporation reactions.


CGTA
E E E a c g t a c g t E A a c g t
R T R T

number of different concentrations of each, in order to titrate the amount of primer required to convert the SS DNA into CC DNA, see figure 25. Copying reactions were routinely performed overnight at 25°C, but 20 minutes at 37°C or two hours at 30°C gave indistinguishable results when analysed by agarose gel electrophoresis.

Mutagenesis with the β-galactosidase-inactivating primer 24 resulted in 10% to 50% mutant white plaques when transfecting CC DNA into *E. coli* JM101. The mutagenesis could easily be verified by single track sequencing using ddATP, the mutant ladder was shifted by one position relative to the wild-type ladder because a single T was missing in the template (figure 26).

9.3.2 Misincorporation mutagenesis

Misincorporation of dTTP against a template C in C6.9 at the 3' end of the universal primer disrupts the coding sequence of the β-galactosidase gene, by giving rise to a UAA termination codon. Thus progeny arising from the mutated DNA form white plaques in the presence of IPTG and BCIG in place of wild-type blue plaques. In this way misincorporation mutants can be screened for by their plaque colour change, allowing a rapid assesment of the efficiency of the misincorporation reaction.

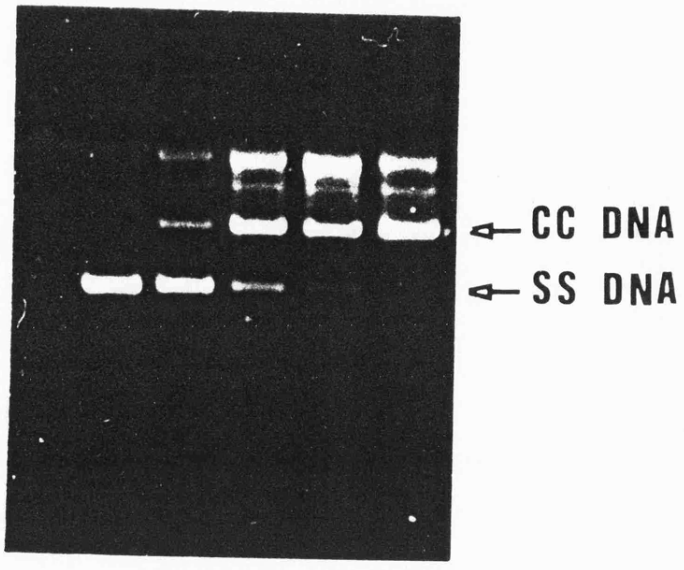
Universal primer was annealed to template C6.9 and incubated with reverse transcriptase and dTTP to give a misincorporation of template C : substrate T. Figure 27

Figure 25.
Titration of an oligonucleotide primer against a template by copying reactions.

Universal primer (5' phosphorylated) was annealed to template C6.9 at different ratios as follows (primer:template): a=0:1, b=0.1:1, c=0.5:1 d=1:1, e=5:1. Copying of the template SS DNA to produce CC DNA was at room temperature for 1 hour followed by 37°C for 20 minutes.

Figure 26.
Single track screening (by dideoxyATP sequencing) of progeny of oligonucleotide mutagenesis.

Primer 24 was annealed to template mp8 and copied overnight at 25°C. CC DNA was isolated from an LMP agarose gel and transfected into *E. coli* JM101. Mutants were identified as white plaques and occurred with a frequency of 10-50%. Single track sequencing of the SS DNA isolated from the phage preparations was performed with universal primer, and analysed on a 6% polyacrylamide gel. Mutants are identified by a shift in the sequence compared to the wild-type mp8 control (W) due to loss of a template T at the position marked A →. Four mutants had an A band missing on sequencing (o), and two had an extra A band (■).


CC DNA
SS DNA

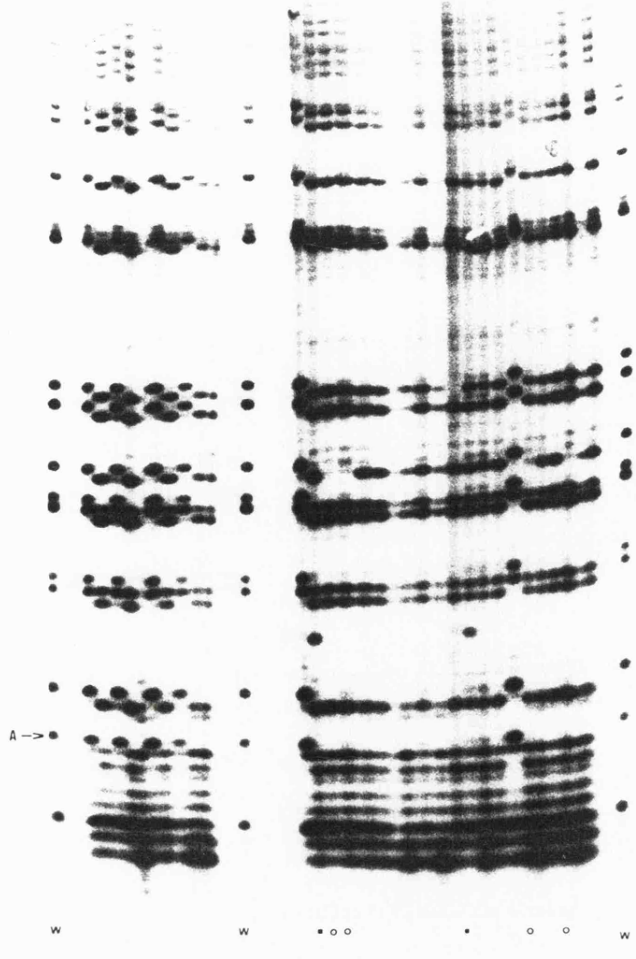
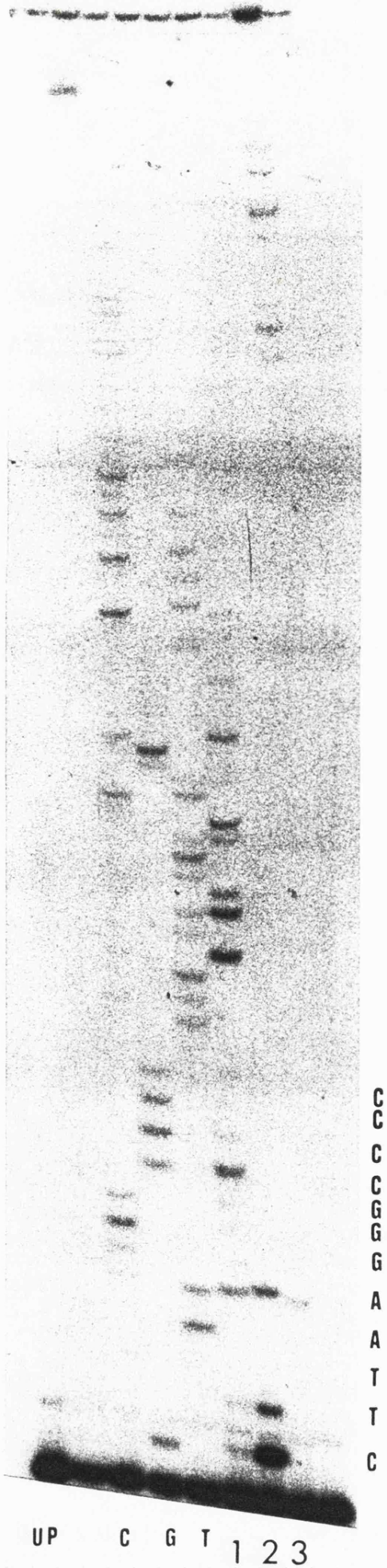

A

Figure 27.
Misincorporation of dTTP and α-SdTTP against template C.

Universal primer (UP) was annealed to template C6.9.
Template: 3'...ACTTAAG
Universal primer: 5'...T

Misincorporations were performed with α-SdTTP (1) or dTTP (2) for 30 minutes at 37°C. Reaction 2 was chased, at room temperature for 5 minutes (3). C,G,T: dideoxysequencing tracks. The template sequence is illustrated at the right hand side of the picture.


C C C G G G A A T T C
UP C G T 1 2 3

illustrates the independent misincorporation of both dTTP and α-SdTTP. dTTP gave a strong band at position 1, and weak bands at positions 2 and 5. α-SdTTP also gave bands at positions 1,2 and 5, but also gave many other bands showing longer extensions indicating that the α-SdTTP contained contaminating dNTPs or α-SdNTPs.

After the misincorporation the reaction was chased with all four dNTPs, extending the second strand well beyond the site of misincorporation (track 3). The second strand synthesis was completed with Klenow polymerase and DNA ligase, and the purified CC DNA was transfected into *E. coli* JM101. This experiment was repeated a number of times. In the first two times the C6.9 SS template was unmethylated having been grown up by Dr. I.C. Eperon in *E.coli* JM101 deficient in the dam methylase gene. Transfection of CC DNA into *E. coli* BMH 71.18 *mutL* gave no higher yields of mutants than did *E. coli* JM101.

In experiment 1 misincorporation of dTTP gave rise to 10% of the progeny being mutants. Misincorporation of α-SdTTP only resulted in 1% of the progeny being mutant. This difference probably reflects the level of contamination of α-SdTTP. The verification of these single misincorporation mutants was carried out by Dr.I.C. Eperon by T track sequencing from a 5'-labelled primer.

Experiment 2 was designed to assess which method of removal of dNTPs would work best when applied to the short initial extension reactions. In this model case the primer was annealed to the unmethylated template (gift of Dr. I.C.

Eperon) and then divided into three portions, two of which were made to 0.1mM of each dNTP. These two portions were then purified by either filtration through a Sephadex G100 column, or through an NENsorb 20 cartridge. All three portions of annealed primer-template were then incubated with dTTP and reverse transcriptase for 30 minutes at 37°C before adding a chase mix of all four dNTPs and then copying overnight with Klenow polymerase and T4 DNA ligase. Transfection into *E. coli* JM101 in the presence of IPTG and BCIG gave the following results:-

Primer-template (universal primer + C6.9Dam-)

	+dTTP	20% white plaques
(+dNTPs +G100)	+dTTP	20% white plaques
(+dNTPs +NENsorb)	+dTTP	0% white plaques

The mutation was verified by T track sequencing of the DNA from isolated white plaques. Clearly the purification through Sephadex G100 removed all the contaminating dNTPs, but the NENsorb cartridge was inadequate, so preventing misincorporation of dTTP.

9.3.2.1 The effect of template primary structure on the production of mutants.

Having established the pattern of misincorporation *in vitro* with different combinations of template nucleotide and

substrate dNTP, experiments were carried out to determine whether or not this pattern was also reflected in the recovery of misincorporation mutants after complete second strand synthesis and transfection into *E. coli* JM101. Four primers only were chosen (primers 4,5,6 and 9), each with a different next template nucleotide, because the *in vitro* results had indicated no effect of the primer 3' terminal nucleotide. Each primer was incubated in four reactions with reverse transcriptase and one dNTP. The progeny arising from transfection of *E. coli* JM101 with the closed circular DNA were grown to give SS DNA preparations for single track sequencing. For each of the 16 different incorporation or misincorporation reactions, 48 plaques were grown up for screening by sequencing. Misincorporation of dTTP onto primer 4 inactivates the β-galactosidase gene resulting in white plaques amongst wild-type blue plaques.

Experiments 3 and 4 used the methylated C6.9 template that had been used for the *in vitro* assays of all 16 primers. In experiment 3 the efficiency of the misincorporation reactions (as determined by the percentage of plaques that were white) could not be assessed by colour discrimination. SS DNA was prepared of 720 clones, 48 from each of the 15 misincorporation and incorporation reactions (excluding primer 4 + dATP) and the progeny of 9 misincorporation reactions were screened by single-track sequencing of 46 clones in each case, a total of over 400 dideoxysequencing tracks. In experiment 4 the mutagenesis reactions with only primer 4 were repeated; 6.4% of the plaques from the dTTP

reaction were white. SS DNA of 192 clones was prepared of which 184 were sequenced.

Primer	Substrate dNTP	Misincorporation mutants recovered by sequencing 46 progeny of each reaction	
		Experiment 3.	Experiment 4.
Primer 4	A	sample lost	3
	C	not tested	0
	G	not tested	2
	T	1	3
Primer 5	C	0	not applicable
	G	1	"
	T	0	"
Primer 7	A	0	"
	C	0	"
	G	0	"
Primer 9	A	0	"
	G	not tested	"
	T	0	"

The single track sequences of the results of experiment 4 are shown in figure 28. The primer/template was as follows:

	1234 (template positions)
template	3'...ACTTA
primer	5'...T

In vitro results were as follows: (see section 9.1.4)

	1234 (template positions)
template	3'...ACTTA
primer + dATP	5'...TAAA
primer + dCTP	5'...T-
primer + dGTP	5'...TGGG
primer + dTTP	5'...TT

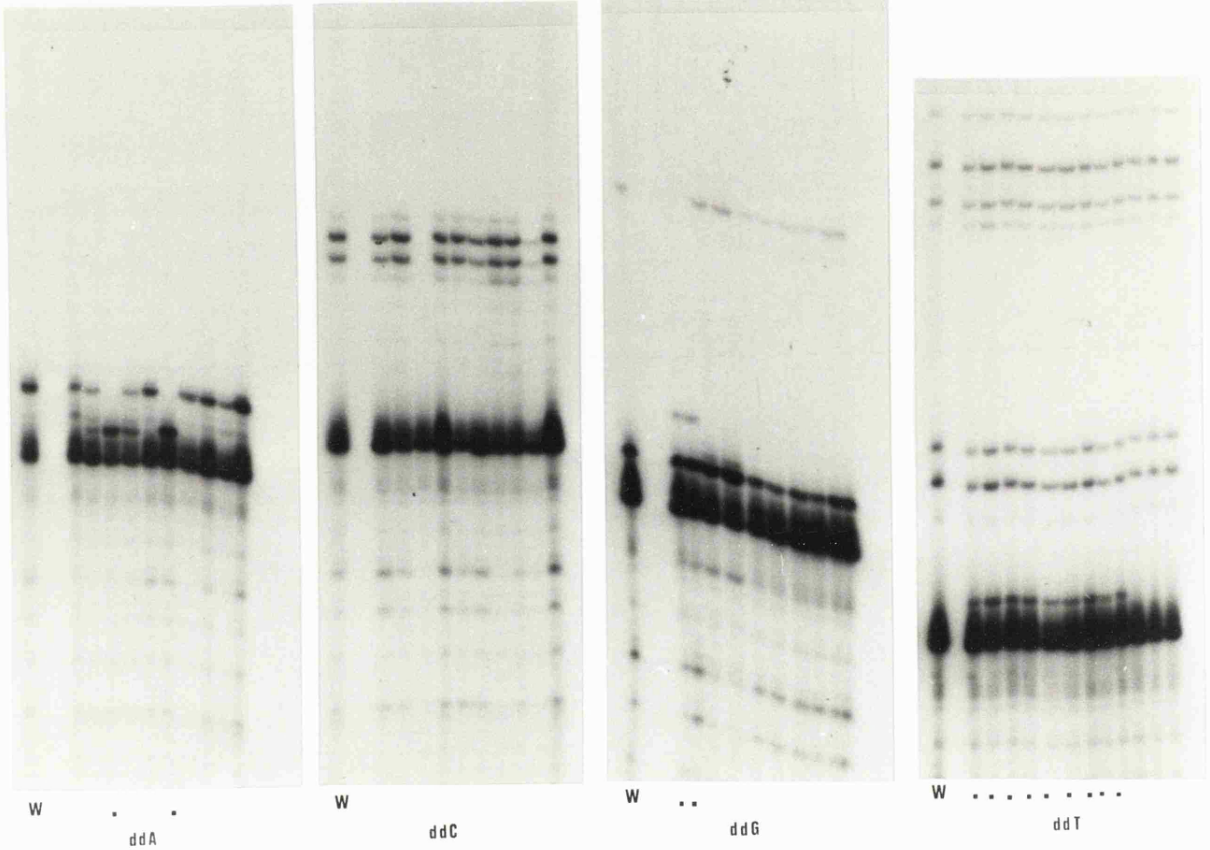
With a template C:substrate dCTP misincorporation none of the 46 plaques screened were mutants. The two mutants found after misincorporation of dGTP both had a correctly incorporated G at position 1 (see figure 28), followed by a misincorporated G at position 2 (template T). Insufficient dGTP was added to the ddGTP sequencing reactions to determine whether or not the third position was also a misincorporated G as seen in the *in vitro* analysis (see section 9.1.4).

A higher overall efficiency of recovery of mutants in experiment 3 (such as the 20% seen with non-methylated SS

Figure 28.
Single track sequencing: screening for misincorporation mutants from primer 4 (universal primer) with template C6.9.

Template: 3'...ACTTAACCCGGGG
Primer 4: 5'...T

Screening was by sequencing 184 phage DNA preparations: 46 from each of the four misincorporation reactions. The single ddNTP track chosen in each case corresponded to the misincorporated nucleotide. Shown here are 12 analytical sequence tracks of each reaction alongside a control sequence of wild-type C6.9 (W). (■): misincorporation mutants.


W
ddA
W
ddC
W
ddG
W
ddT

C6.9 template) would have allowed an analysis of the frequencies of recovery of different misincorporation mutations, and hence a clearer interpretation of the effect of the template primary sequence on the generation of misincorporation mutants.

9.4 Application of misincorporation mutagenesis methodology to a cloned gene

In preparation for the application of the misincorporation mutagenesis methodology to a specific gene, the c-Ha-*ras* gene was subcloned into bacteriophage M13.

9.4.1 Subcloning of c-Ha-*ras* DNA into M13.

Two *ras* clones were made available by Dr. E. Chang: J841 (c-Ha-*ras* 1) the human proto-oncogene, and J180 (pEJ, bladder carcinoma) the activated human oncogene (Tabin *et al.* 1982). Both plasmid DNAs were transformed separately into *E. coli* HB101, and the ampicillin resistant colonies were isolated and grown up to prepare plasmid DNA. Both plasmids contained the 6.6kb *Bam* HI fragment of human c-Ha-*ras* DNA (either the proto- or the activated oncogene) cloned into the *Bam* site of pBR322. All the *ras* coding exons are found on a single *Sac* I fragment of 2.9kb, but an upstream non-coding exon and possible enhancer sequences lie within the preceding 840bp *Sac* I fragment.

It was decided to subclone *Sac* fragments of *ras* into

M13 vectors for ease of single strand preparation and subsequent sequencing. The M13 vectors used, mICE 10 and mICE 11 (Eperon, 1986b), were linearised with *Sst* I, an isoschizomer of *Sac* I, in 1xSeq buffer at 37°C.

9.4.1.1 Partial digestion of *ras* plasmids.

Both *ras* plasmids were partially digested with *Sac* I in 1xSeq. buffer for increasing lengths of time at 37°C; the analytical gel is shown in figure 29. Preparative digests were run on 0.6% LMP agarose gels and the 2.9kb and 3.7kb bands excised and eluted as described previously. Verification of the isolated DNA is illustrated in figure 30 showing the 2.9kb and 3.7kb *Sac* I fragments of J841 and J180 adjacent to partial Sac digests of the plasmids.

9.4.1.2 Ligation of *ras* fragments into M13.

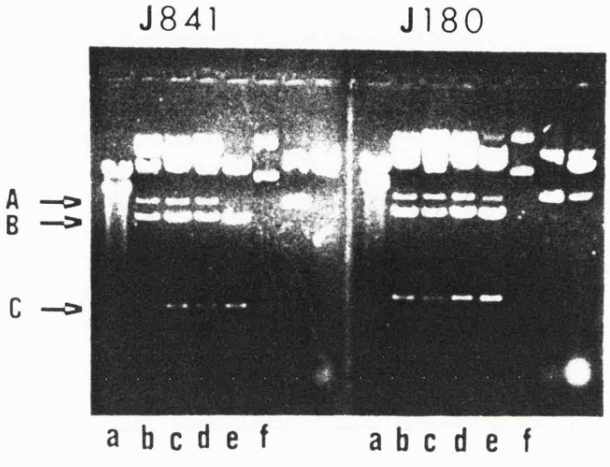
Approximate concentrations of the DNA fragments were calculated from the intensity of bands on agarose gels. Three different molar ratios were used for each set of ligations, 1 vector:1 insert, 1 vector:5 insert, 1 vector:20 insert. Ligations in 10µl 1xC buffer, 1mM rATP, using 1 unit of T4 DNA ligase per 100ng of M13 vector were incubated at room temperature (about 20°C) overnight. Controls were: vector with no ligase and no insert, vector with ligase but with no insert. On transfection into *E. coli* JM101, successful ligations produced 1% recombinants seen as white

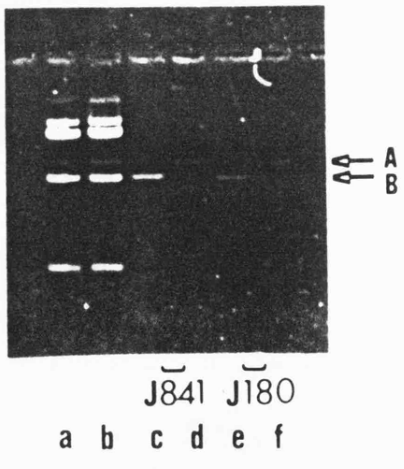
Figure 29.
Partial digestions of *ras* plasmids: analytical gel.

Two *ras* plasmids J841 and J180 were digested with *Bam* HI which produced a 6.6kb fragment containing the *ras* gene: (a). Partial digestions with *Sac* I (b-f) produced fragments of 3.7kb (A), 2.9kb (B) and 840bp (C). b: 1 min., c: 3 min., d: 6 min., e: 15 min., f: 25 min.

Figure 30.
Isolated fragments of *ras* plasmids.

a,b: *Sac* I digestions of J841 and J180,
c,e: 3.7kb *Sac* I fragment of J841 and J180 (A)
d,f: 2.9kb *Sac* I fragment of J841 and J180 (B).


J841
J180
A
B
C
a b c d e f
a b c d e f


A
B
J841 J180
a b c d e f

plaques amongst blue plaques of religated vector. Ligations were also transfected into *E. coli* JM109 in an attempt to recover recombinant phage containing the large 3.7kb fragment. This strain is supposed to allow more stable growth of larger recombinants but transfection of competent *E. coli* JM109 gave poor yields of plaques, and no recombinants were found.

Some ligations were performed with unpurified partial *Sac* digests of *ras* plasmids as insert material. The only recombinant clones from these had inserts of 840bp.

9.4.1.3 Recovery of recombinants.

White plaques recovered after transfection of *E. coli* JM101 were grown for 7 hours in TY broth in microtitre plates. SS phage DNA templates were prepared and samples of each DNA were analysed on agarose gels next to suitably sized marker DNAs. Samples that appeared to be recombinant clones of the correct size often contained contaminating bands of DNA the size of religated vector without insert. This would be expected if mixed plaques had been picked instead of single clones. All recombinant clones were retransfected and thus, purified white plaques were isolated from contaminating blue plaques. The presence of blue plaques in this repurification argues against the vector-sized DNA bands on the agarose gels having arisen from instability of the recombinants. Having repurified the SS phage DNA preparations, these were sequenced to determine

the orientation of the insert. Replicative form DNA was prepared from the recombinant clones and analysed by *Sac* I and *Bam* HI digests to confirm the size of inserts.

9.4.1.4 Recombinant clones. (figure 31)

The 840bp fragment of J180 (pEJ) from 210 to 1054 was subcloned into mICE 11 in both possible orientations: mEJ-*ras* 840:10 and mEJ-*ras* 840:14. The 2.9kb fragment of J180 (pEJ) from 1054 to 3946 was subcloned into mICE 10 in both possible orientations, mEJ-*ras* 2.9:3 and mEJ-*ras* 2.9:10. mEJ-*ras* 2.9:3 was used for misincorporations and extensions and was termed M-*ras*. The 2.9kb fragment of J841 from 1054 to 3946 was also subcloned into mICE 11, two clones were recovered, mcH-*ras* 2.9:1 and mcH-*ras* 2.9:4; 2.9:1 had the same orientation, relative to the universal sequencing primer, as mEJ-*ras* 2.9:3, and the orientation of 2.9:4 was not determined. No clones were recovered containing the *Sac* I partial 3.7kb (2.9kb + 840bp) *ras* fragment from 210 to 3946.

9.4.2 Misincorporations with template M-*ras*.

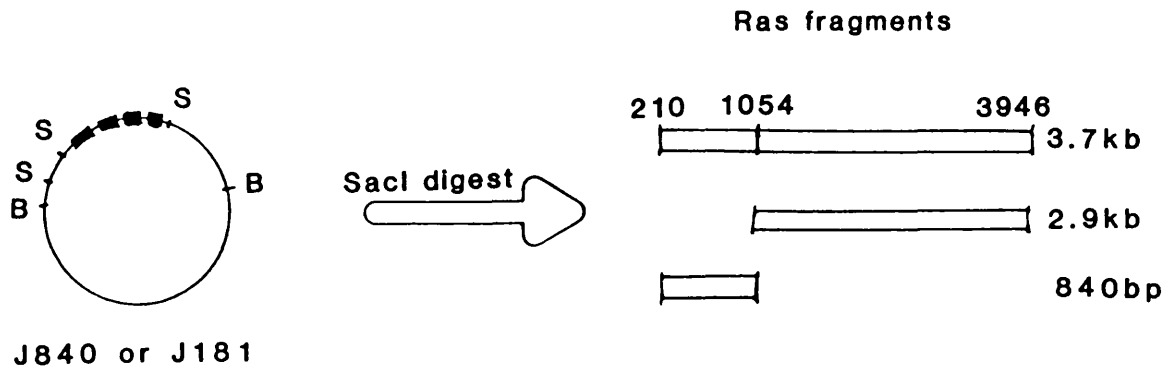
Universal primer and primer 51 were 5' labelled, annealed to SS M-*ras* DNA and subjected to misincorporation of single dNTPs by reverse transcriptase. The results are shown in figure 11.

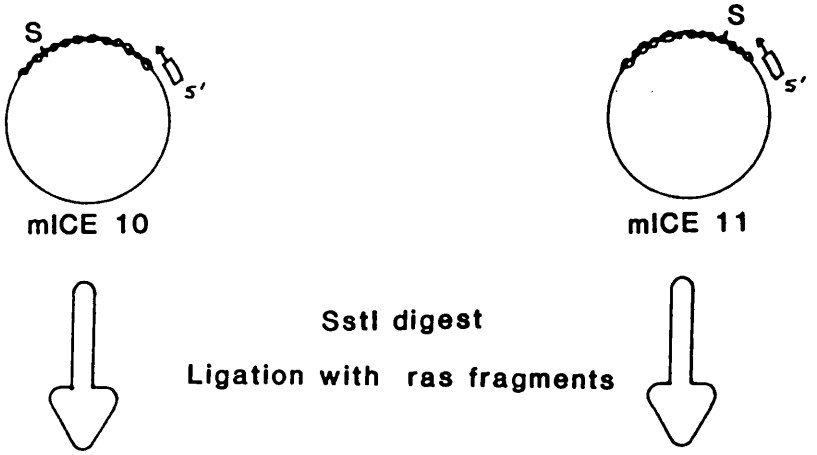
Figure 31.
Recombinant *ras* clones in M13.

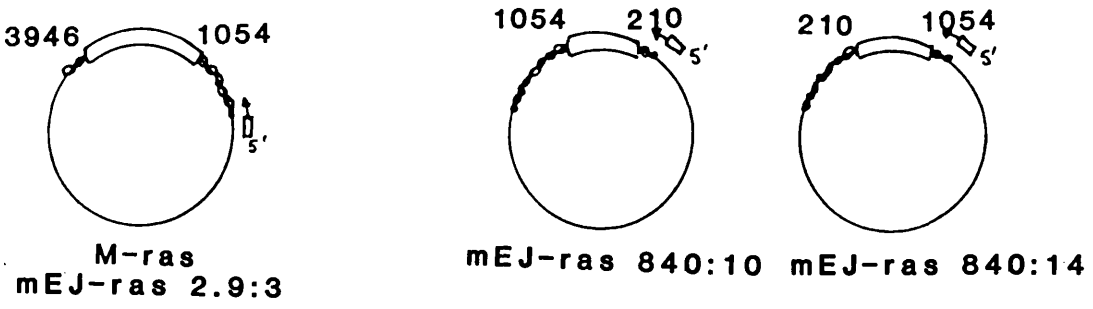
B: *Bam* HI site
S: *Sac* I or *Sst*I site

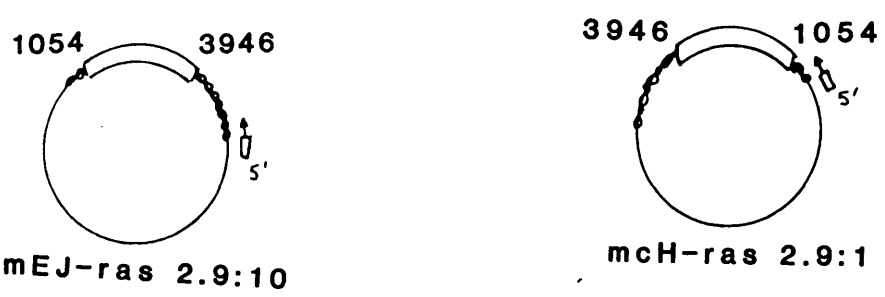
Multiple cloning site

5' Universal priming site and direction of synthesis of second strand


Ras fragments
S
S
S
B
B
J840 or J181
SacI digest
210
1054
3946
3.7kb
2.9kb
840bp


S
5'
mICE 10
S
5'
mICE 11
SstI digest
Ligation with ras fragments


3946
1054
5'
M-ras
mEJ-ras 2.9:3
1054
210
5'
mEJ-ras 840:10
210
1054
5'
mEJ-ras 840:14


1054
3946
5'
mEJ-ras 2.9:10
3946
1054
5'
mcH-ras 2.9:1

9.4.3 Limited extensions of an oligonucleotide primer with template M-*ras*

Primer 51 was annealed to SS M-*ras* DNA and extended in the presence of dCTP, dGTP, dTTP and limiting amounts of dATP including a trace of α^{32}P-dATP. The extended primers ranged from 17 nucleotides to 150 nucleotides and are shown in figures 16 and 17.

Free dNTPs were removed from the extended primers by purification through a denaturing polyacrylamide gel, the primers were reannealed to template SS M-*ras* and tested in copying reactions for production of CC DNA (figure 32).

9.4.4 Misincorporation on extended *ras* primers.

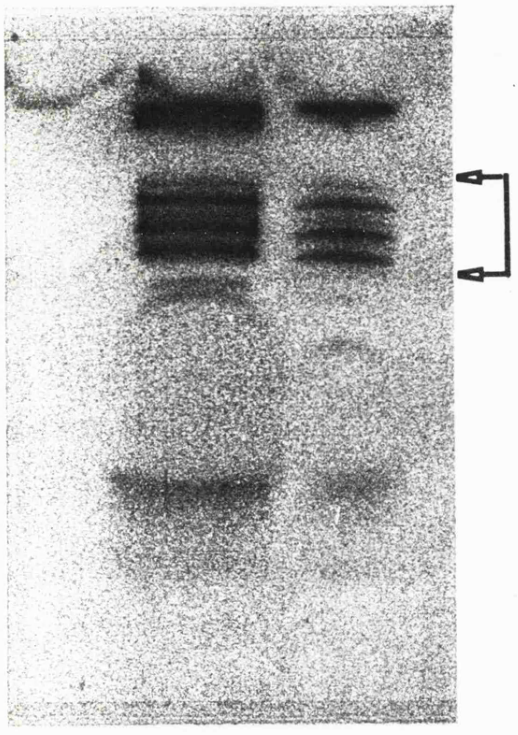
The misincorporation and incorporation of dNTPs onto the extended primers is shown in figure 22. Analysis of the reaction products by PAGE demonstrated the incorporation or misincorporation by a shift in the banding pattern of the primers.

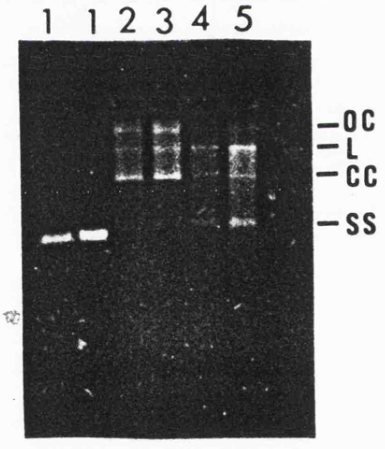
9.4.5 Second strand synthesis after misincorporation.

Reaction products from misincorporations of dCTP, dGTP and dTTP were converted to CC DNA by incubation of the primer-templates in copying reactions for two hours at 30°C. After the short chase reactions to "seal-in" misincorporations, the reverse transcriptase was heat inactivated at

Figure 32.
Purification of primer 51-template M-*ras* short extensions.

Free dNTPs present after the short extension reaction were removed by electrophoresis of the 5'-labelled, extended primers into a polyacrylamide gel and eluting the DNA from the region marked by arrows. The purified primers were annealed to fresh M-*ras* template and analysed in copying reactions:
1: SS template DNA M-*ras*.
2,3: M-*ras* + primer 51
4,5: M-*ras* + extended primer 51.
OC: open circles, L: linear, CC: closed circles, SS: single-stranded.




1 1 2 3 4 5
OC
L
CC
SS

100°C for five minutes and then placed on ice. The primer was reannealed to the template. Agarose gel analysis of copying reactions where the reverse transcriptase had not been denatured gave a background smear of DNA and poor yields of CC DNA. Excess reverse transcriptase in the misincorporation reactions also resulted in poor copyings, possibly due to damage caused to DNA by reverse transcriptase. To avoid any problems occurring in copying reactions caused by reverse transcriptase damage to template DNA, fresh SS template DNA was added to the reactions after the inactivation of reverse transcriptase and prior to reannealing. The presence of fresh template resulted in improved yields of CC DNA.

The CC DNA band was excised from an LMP agarose gel, the gel slice was melted and a fraction transfected into competent *E. coli* JM101.

9.4.6 Analysis of the progeny of misincorporation mutagenesis using a bank of extended primers.

A total of 173 individual plaques were isolated, 58 from dCTP misincorporation reactions, 87 from dGTP misincorporation reactions and 28 from dTTP misincorporation reactions, and SS DNA was prepared in microtitre plates. Preparations of SS DNA were subjected to single-track dideoxysequencing reactions (58 C tracks, 31 G tracks and 28 T tracks) and point mutations were sought in the region spanned by the bank of extended primers. Misincorporations

were expected at positions where the template nucleotide had been a T, because of the dATP-limited random extensions of primer 51. Unfortunately no *ras* mutants were recovered from the screening of 107 plaques.

CHAPTER 10. Discussion

10.1 *In vitro* assay of misincorporation by reverse transcriptase.

The simple assay described here has allowed examination of the efficiency of misincorporation of all three possible dNTPs by AMV reverse transcriptase with all 16 possible combinations of the nucleotide at the 3' terminus of the primer and the template nucleotide. This study was initiated after observing that with any one primer-template all three possible dNTPs were not misincorporated equally.

The principal objective of these misincorporation assays was to establish whether all possible combinations of 3' primer terminus, template nucleotide and substrate dNTP can give rise to efficient misincorporation. Despite the primer-dependent or template-position-dependent variations in efficiency, the preferences shown by each primer for various dNTPs were remarkably consistent. Several surprises arose from the analysis and are discussed below.

10.1.1 Contamination of dNTP stocks.

Initial experiments showed problems of contamination of dNTP stocks with other dNTPs. This problem was highlighted by the widely differing extensions seen when

misincorporating different dNTPs. The nature of the extensions was investigated by Maxam and Gilbert sequencing in order to establish whether the extension was in fact due to misincorporation or to correct incorporation of a contaminating dNTP. The problem was solved by the use of ultrapure dNTP stocks from Pharmacia.

The results of the misincorporation assay showed a very wide range of efficiencies when the levels of misincorporation after a fixed incubation time were compared with a control incorporation. These efficiences were highly dependent upon the primer used. Thus the best possible misincorporation reaction with one primer could be eightfold less efficient than the best reaction with another primer. Since these levels of misincorporation were all compared with the appropriate correct incorporation reaction, it is unlikely that these levels represented artefacts of primer preparation (the oligonucleotide primers were all prepared simultaneously). However, kinetic experiments showed that correct incorporation was usually complete within the assay time used in most of the experiments, and so any correct incorporation that was slower than usual might not have been noticed.

10.1.2 The contribution of the primer in misincorporation.

The 3' terminus of a primer could potentially affect misincorporation onto that primer. Factors playing a role in this effect might include:

1) the 3' terminal nucleotide: specific nucleotides might facilitate certain incorporations and misincorporations more favourably than others;
2) the template position of the primer: the template position might affect the levels of all incorporations and misincorporations;
3) different stacking interactions within the primer 3' terminus might influence which dNTPs can be misincorporated. These factors are discussed below.

10.1.2.1 The effect of the primer 3' terminal nucleotide.

The nucleotide at the 3' terminus of the primer did not appear to affect the preferences among dNTPs for misincorporation in a systematic fashion. This can be seen in Table 2 where no particular 3' terminal nucleotide gave different efficiences of misincorporation.

10.1.2.2 The role of template position.

The effect of template position on misincorporation by DNA polymerase I was investigated by Hillebrand and co-workers (1984) using PAGE analysis of primer extensions in the presence of only three substrate dNTPs. The variations in the propensity for misincorporation at different template positions were in part explained by:

1) the stability of subsequent correct incorporations

(protecting the misincorporated nucleotide from repair by the 3'-5' exonuclease activity);

2) base-stacking in the primer terminus.

The first explanation is inapplicable to the results presented here: AMV reverse transcriptase contains no 3'-5' exonuclease activity. Furthermore, there is no good correlation between the sequence at the 3' terminus of the primer and the levels of misincorporation exhibited with that primer. DNA polymerase I and AMV reverse transcriptase have been shown by Hillebrand and Beattie (1985) to prefer different template sites for pausing and for misincorporation, although the factors influencing this remain unknown and there was no good correlation with secondary structure. It is possible that the poor misincorporation on primer 8 in this study might be attributable to some secondary structure of the template forming adjacent to the 3' terminus of the primer where the sequence was 3'AGGGCCCCT.

10.1.2.3 The role of base-stacking interactions.

Topal and co-workers (1980) considered that stacking interactions were significant in misincorporation with T4 polymerase, although of little consequence when compared with hydrogen bonding in the formation of correct base pairs. In contrast, the results presented here suggest that stacking interactions between the primer 3' terminus and the substrate dNTP are not generally important with reverse

transcriptase. The only cases where stacking interactions might be involved were where dGTP was involved in consecutive misincorporations. In most cases only a single dNTP was misincorporated, although a subsequent incorporation allowed a further misincorporation in a few cases. However, in seven cases two consecutive misincorporations took place, the second usually being weaker, with no subsequent incorporations. In all these cases the preceding base pair was a G:C pair, the misincorporated substrate was dCTP or dGTP and the first misincorporation would have been expected to be favourable. In the only two cases where a preceding A:T base pair allowed consecutive misincorporations a subsequent incorporation might have facilitated this.

The lack of effect of the 3' terminal nucleotide of the primer on misincorporation is important in mutagenesis strategies involving misincorporations at a number of different positions along a template (see 10.6). The substantial effect of the template position of the primer appeared to be mediated by the combination of next-template nucleotides and the substrate dNTP.

10.1.3 The effect of the template nucleotide.

The major determinant of the level of misincorporation of different dNTPs was the template nucleotide. With any one primer a poor misincorporation achieved less than half the level of a good misincorporation, and no apparent reaction

meant that the level was undetectable, i.e. less than 1-10% of a good misincorporation (according to the primer used). The results cannot be explained by a simple scheme based on the strength of the base-pairs, nor even by such a scheme taking into account putative stacking interactions. This is plainly shown by the results of homologous misincorporation: dGTP and dCTP were never misincorporated (regardless of the primer 3' terminus nucleotide) against G or C, respectively, in the template, and yet dATP and dTTP were usually misincorporated against A or T, albeit poorly.

There is a very interesting asymmetry with almost every pair of nucleotides: reactions worked comparatively well with a given combination of dNTP and template nucleotide, but poorly when the same pair were in the converse combinations. With G:T, C:T and G:A pairs (C:A is less clear) the misincorporation was most efficient with dCTP or dGTP substrates and A or T in the template. The difference between pyr:pu and pyr:pyr or pu:pu misincorporations is manifest not in the favoured combinations (dCTP is misincorporated against T as well as against A) but in the unfavourable ones (dATP is more favourably misincorporated against C than is dTTP). When the unfavourable combinations are considered there appears to be a general order of preference, pyr:pu > pyr:pyr > pu:pu.

Thus, an approximate order of preference, written as dNTP:template, is: dGTP:T > dCTP:A = dCTP:T = dGTP:A > dATP:C = dTTP:G > dTTP:T > dTTP:C > dATP:A > dATP:G = dCTP:C = dGTP:G. In all cases the best misincorporation achieved by

a primer was rated as "good", and then the usage of the other dNTPs was assessed relative to this in an attempt to compensate for primer-specific factors. Furthermore, each assessment was based on four primers with different 3' termini.

This order is quite unlike the order of stability of mismatches deduced from measurement of oligonucleotides in solution (Aboul-ela *et al.*, 1985); the only agreement concerns the pre-eminence of G:T mismatches. However, such experiments on stability have addressed mismatches with matched base-pairs on both sides, altering stacking effects; the situation being inherently symmetrical in that the two mismatched nucleotides were both set within a helix. Thus, such experiments could not have shown the substantial assymmetry seen in the results presented here which is apparently a characteristic of the enzyme active site.

A number of features that might account for the order of misincorporation preferences are considered below. These include the symmetrical disposition of the glycosyl bonds with respect to the vector between the sugar C1'..C1' atoms, the dimensions and geometry of a proposed pair of bases, the contributions of rare tautomers, wobble base pairs and base-stacking.

It is apparent that the structural features of a mismatched pair which have been correlated with efficiences of repair (Hunter *et al.*, 1986) do not fit the order of misincorporation preferences presented above at all; the symmetry of the glycosidic linkages would suggest an

order G:A > A:C > G:T (Hunter *et al.*, 1986). The results presented here suggest that stability of the product is only of importance with regard to its function as a primer terminus for any further misincorporations and that the enzyme active site may respond in different ways to the four nucleotides in the template.

The dimensions and geometry of a potential pair of bases have been proposed as factors determining the specificity of DNA polymerase I (Kornberg and Kornberg, 1974). Polymerisation is based on the demand for one of the four base-pairs (A:T, T:A, G:C, C:G) and not on the recognition of the incoming dNTP. When the correct base-pair is in the active site as determined by its dimensions and geometry then the enzyme proceeds with polymerisation. Non-matching dNTPs are rejected because they cannot form base-pairs of the correct dimensions. The differing responses seen with reverse transcriptase indicate that such rejection may vary according to each particular mismatch so that some are polymerised more efficiently than others. Differential responses with DNA polymerase I might be masked by subsequent removal of mismatched termini by the 3'-5' exonuclease activity.

Rare tautomers may be involved in the production of some mismatches but cannot account for the order given above because of the low frequency of tautomerization. Neither can wobble base-pairing account for this order. Stacking effects do not seem to be important and so the criteria which are responsible for the observed order of preferences of

misincorporation with reverse transcriptase remain unknown.

10.2 "Pull-through" effect of next-template nucleotide.

"Pull-through" by reverse transcriptase of an otherwise unfavourable misincorporation has been seen with primers where the template sequence allows a correct incorporation to follow a weak misincorporation. The experiments using ddNTPs and dNTPs (figure 15) have shown that this "pull-through" requires polymerisation between the misincorporated nucleotide and the subsequent correctly incorporated nucleotide, but that proofreading exonuclease activity is not involved: the assays of rates of misincorporation tend to discount active involvement of an exonuclease (steady state levels are not reached quickly and phosphorothioate linkages do not enhance levels). The effect is specific for subsequent incorporation of the same dNTP as that which was misincorporated.

The data presented here are insufficient to formulate an explanation or model for this process, but two models that have been proposed previously as important in the fidelity of replication are discussed below.

10.2.1 "Next-nucleotide" effect.

The "next-nucleotide" effect observed with DNA polymerase I is discussed by Loeb and Kunkel (1982). The

presence in a DNA polymerase I misincorporation reaction of a substrate dNTP complementary to the second template nucleotide allows more misincorporation mutagenesis than when a single dNTP is misincorporated at position 1 and is unable then to be incorporated at position 2. This phenomenon was termed the "next-nucleotide" effect and is consistent with the observation by Bernardi and Ninio (1978) that rapid DNA synthesis by polymerases that contain a proofreading exonuclease activity decreases the fidelity of replication by competition between excision and polymerisation. The "next-nucleotide" effect cannot directly explain the "pull-through" effect observed here because of the lack of a proofreading activity in AMV reverse transcriptase.

10.2.2 Energy relay.

Hopfield (1980) designed a scheme for DNA polymerase fidelity known as energy relay which has no requirement for a 3'-5' exonuclease activity. The mechanism uses the energy derived from the previous incorporation step to increase fidelity of subsequent additions. In this way processivity of the enzyme increases fidelity and this was seen when comparing the error rates of DNA polymerase α (processive: adding about 10 nucleotides per association event) and DNA polymerase β (distributive: dissociating from the template after every nucleotide addition step) with ϕX174 (Kunkel and Loeb, 1981). This scheme contrasts sharply with kinetic

proofreading in which decreased rates of synthesis would increase fidelity.

AMV reverse transcriptase (in its αβ form) is a processive enzyme (Verma, 1977), and might be expected to demonstrate energy relay. Energy relay could not be investigated by the addition of only one of the four possible dNTPs. In the cases where extension beyond the first template position was possible because of incorporation at that first position, the energy relay scheme would suggest that further extension would be more accurate, ie. misincorporation would be reduced. However, in the cases where the first template position permitted a correct incorporation, no reduction in misincorporation at the second position was observed. Thus no energy relay mechanism was implicated from the patterns of misincorporation.

10.3 Degradative activities of reverse transcriptase.

The results (shown in figure 14) demonstrate degradation of the primer on long incubations with large amounts of reverse transcriptase. Proteolytic attack of reverse transcriptase *in vitro* cleaves the β-subunit to α and smaller polypeptides of 20,000-40,000 molecular weights, activating a DNA endonuclease of M.W. 32,000 resembling the virion-associated $p32^{pol}$ endonuclease (Grandgenett *et al.*, 1980). This endonuclease is more heat stable than reverse transcriptase and may have been responsible for digestion of

primers on long incubations. The reverse transcriptase used in these assays was purchased from Anglian Biotechnology Ltd. and is reported to be purified AMV reverse transcriptase containing all its inherent properties but no contaminating exonucleases or endonucleases. Reverse transcriptase produced from cloned genes has recently become available, and therefore may lack the region responsible for the $p32^{pol}$ endonuclease activity.

The only other degradative activity of reverse transcriptase is RNase H (Mölling *et al.*, 1971) an exoribonuclease on the same polypeptide as reverse transcriptase but with a different functional site. The products of RNase H digestion are 4-20 ribonucleotides long. The activity should not have played any part in the assay described here, but it was observed that degradation of the primer often led to the production of 5' labelled oligonucleotides of 4-8 bases long, suggesting some exonuclease or endonuclease digestion which did not liberate mononucleotides.

10.4 Correlation of *in vitro* results with published misincorporation mutagenesis data.

The available data on misincorporation mutants obtained after mutagenesis are very limited. Zakour and co-workers (1984) isolated mutants following misincorporation mutagenesis with AMV reverse transcriptase by selecting for revertants of a ϕX174 amber mutation. They showed that with

a primer 3' terminus T and next template T, misincorporation of individual dNTPs gave rise to mutants with efficiencies dGTP > dCTP >>dTTP. With a 3' terminus C and next template nucleotide A, the efficiencies were dCTP > dATP >> dGTP. The first series follows the results presented here remarkably well, despite the risks inherent in extrapolation from the frequency of mutants recovered to the initial misincorporation efficiency; the second series shows an order with dATP and dGTP that is apparently contrary to the results presented here, perhaps because dATP misincorporation could be substantially enhanced by a subsequent incorporation (see 10.3).

An analysis of mutations achieved by misincorporation with DNA polymerase I produced very different results from those reported here except that substrate dGTP:template G misincorporations were not detected (Shortle and Lin, 1985). The mutations were introduced into the staphylococcal nuclease gene in an attempt to identify residues specifying critical steps in the folding pathway of the protein, and involved misincorporation of α-SdNTPs with Klenow polymerase on random 3' termini generated by limited exonuclease III digestion. Nuclease-minus mutants were isolated and the independent misincorporations recovered were (α-SdNTP: template nucleotide): dCTP:T (15) > dGTP:T (13) > dTTP:T (12) > dTTP:G (9) > dATP:A (8) > dGTP:A (7) > dATP:G (2) = dATP:C (2) = dCTP:A (2). Such results clearly differ from the misincorporation efficiencies with reverse transcriptase described here, but no comparison can be made of the

misincorporation efficiencies among different dNTPs at a given template nucleotide. This is because the mutants were selected (and therefore may not be representative) and because of the random 3' termini used (*cf.* the earlier discussion of the importance of the template position of the primer), and the almost certainly unequal distribution of primer termini at the different sites.

There are several possible explanations for preferential misincorporation of some combinations. For example, if the results shown here represented the maximum level that could be achieved, then equilibria or steady state kinetics might be invoked. On misincorporation, reversal of polymerisation may proceed at a comparatively high rate and equilibrium may be reached at a low product level. Alternatively, a 3' exonuclease acting preferentially on mispaired termini could result in reduced levels of misincorporated intermediate.

Experiments with DNA polymerase I have shown that reversal of polymerisation (pyrophosphorolysis) only removes properly base-paired termini (Brutlag and Kornberg, 1972); mispaired termini are not attacked by pyrophosphate and can only be removed by hydrolysis (in this case by the 3'-5' exonuclease activity). For this reason a high rate of reversal of polymerisation is considered unlikely to be a cause of preferential misincorporation. Furthermore, the results suggest that the time point used for all the dNTP: template nucleotide: primer terminus assays was such that the level of product was increasing linearly, i.e. the

levels seen reflect the rate of misincorporation. Only in the most rapid reactions, those approaching 100% efficiency (relative to incorporation), might the rate of increase of product have been departing greatly from linearity as the available template: primer substrate declined. In this case the ability to discriminate between the most favourable combinations might be reduced.

Attempts were made to find reaction conditions (increased incubation times, extra enzyme) that allowed high levels of misincorporation with unfavourable dNTP:template combinations. No better conditions were established: prolonged incubation times resulted in enzyme inactivation and too great an excess of enzyme caused problems of degradation.

10.5 Correlation of *in vitro* and *in vivo* results from primer 4.

When primer 4 (universal primer) was annealed to template C6.9, the next four template nucleotides were 3'..CTTA. The *in vitro* assay of misincorporations gave no misincorporation with substrate dCTP, poor misincorporation with substrate dTTP, good misincorporation of substrate dATP followed by two correct incorporations, and correct incorporation of substrate dGTP followed by two good misincorporations. On analysis of mutants recovered from these misincorporations, no dCTP:C mutants were present.

However, the numbers of dATP:C, dGTP:T, and dTTP:C mutants were very similar to one another. Although the numbers of mutants recovered was low, only two or three per 46 plaques, the similarity of the recoveries of both dATP:C and dTTP:C mutants contrasted with the differing misincorporation efficiences observed *in vitro*. A dATP:C misincorporation creates a transition mismatch which can be repaired efficiently *in vivo* (Kramer *et al.*, 1984b), but a transversion mismatch of dTTP:C may not be repaired at all (Kramer *et al.*, 1984b). Thus the differing misincorporation efficiences of these two situations which were highlighted *in vitro* appear to be masked by subsequent repair *in vivo*.

Transfection of the dTTP:C misincorporation into a *mutL* strain of *E. coli* gave no increase in the recovery of white plaques and so the assumption was made that the strain might have lost its *mutL* phenotype, and no longer be capable of producing higher yields of mutants than wild-type *E. coli* JM101. Such an assumption may have been erroneous, having regrettably been based upon the results of a dTTP:C misincorporation. The recovery of other mutations was not attempted in the *mutL* strain.

Further comparisons of different misincorporation efficiences *in vitro* and *in vivo* would be interesting in order to establish what differential repair mechanisms might exist both in the presence or absence of the Mut repair system.

The efficiency of the misincorporation reactions may have been lower than in other experiments where a higher

proportion of the plaques from the dTTP:C misincorporation were white.

In primer 4 experiments recoveries of mutants of 6.4% and 10 to 20% were observed. In the latter case the increased yields were possibly due to the use of an unmethylated template strand. An unmethylated template destroys the strand bias of Mut repair system. If this were the reason for the increased recovery of mutants, then it could be argued that the Mut repair system is able to recognise and (to some extent) repair dTTP:C transversions. Had unmethylated template been used for all the misincorporations of A,C,G and T, losses of the mutants due to selective repair of the newly synthesized strand might have been reduced.

10.6 The design of mutagenesis strategies involving AMV reverse transcriptase.

The results obtained from this study of misincorporation by reverse transcriptase have considerable implications for design of mutagenesis strategies.

It can be seen that single misincorporation is usual; where a second misincorporation could follow, most molecules will still carry just one misincorporated nucleotide. Using just one dNTP at a time, the initial misincorporation will be followed depending on the template sequence, either by a subsequent incorporation or, very rarely, by a misincorporation.

The subsequent incorporations show that a mispaired 3' terminus is no barrier to extension before transfection. It is believed that DNA polymerase I requires a correctly paired 3' terminus for extension and so mismatched termini are removed by the 3'-5' exonuclease activity giving very low frequencies of misincorporation. However, reverse transcriptase cannot proofread any such mistakes but can clearly extend beyond them as seen by PAGE analysis of misincorporations to which a chase mix of four dNTPs had been added. The situation with reverse transcriptase being unable to proofread repair misincorporations is analogous to DNA polymerase I misincorporating α-SdNTPs.

It is apparent that unless the levels reached by poor misincorporation reactions can be raised, then it will be necessary to optimise the combination of dNTP and template nucleotide for each case with reference to the observed template nucleotide preferences. The template nucleotide following a 3' primer terminus can be specified easily during the formation of a ladder of termini. The termini can be formed by preliminary extensions from another primer where all four dNTPs are present but one is at a much reduced concentration.

Conditions have been presented here which give a spread of primer termini as described above, almost all giving a next template nucleotide of T where dATP was the dNTP limiting in the short extension. The primer lengths ranged from 17 bases to 150 bases and were capable of extension by reverse transcriptase and dATP. Misincorporations were

observed by PAGE analysis when one of the other dNTPs was used instead of dATP. Subsequent chase reactions extended the primers well beyond the mismatch.

An alternative method of generating a range of primers with specified next-template-nucleotides would be to incubate a pre-existing ladder in a "minus" reaction (Sanger and Coulson, 1975), where the 3'-5' exonuclease and the polymerase activity of DNA polymerase I would expose the template nucleotide complementary to the missing dNTP. Given knowledge of the template nucleotide and the order of preference of misincorporation, an efficient combination can be chosen.

Since the 3' terminal nucleotide of the primer has little or no effect on preferences for misincorporation, it will not be necessary to adjust the 3' termini of a ladder of primers such that they conform to an optimal 3' terminal nucleotide. This could have been performed with a "plus" reaction (Sanger and Coulson, 1975): on incubation with DNA polymerase I and a single dNTP the 3' exonuclease activity degrades the primer to a position where the 3' terminal nucleotide is determined as the dNTP present in the reaction. However, with the constraints of the template nucleotide preferences this would have greatly reduced the number of positions at which mutants might be expected.

That mutants can be isolated by reverse transcriptase-directed misincorporation has been demonstrated here. Even after quite poor reactions (Primer 4, dTTP:template C), mutants were isolated at levels of 6-20% of recovered

plaques. Further developments could perhaps enhance such levels by raising the efficiency of poor reactions but as the factors involved are not yet understood, they cannot be controlled. Removal of primers that have not been used as substrates for misincorporations would increase the yield of mutant heteroduplex molecules. This could be done, for example, by using pure α-SdNTPs for misincorporations and then digesting with exonuclease III. Only those primers extended with thionucleotides and which, therefore, have been used as substrates for misincorporation, will survive the exonuclease III digestion. Such digestions have been shown here to be effective, the limitation on this method at the moment being the purity of the α-SdNTPs available.

In summary then, it is apparent that misincorporation mutagenesis could be used to generate a library of random single-base mutations. At present the reaction conditions are such that the library would be biased in favour of certain mutations, but alternative reaction conditions to eliminate this bias could be tested using the assays developed here.

The results presented here also have implications for our understanding of the mechanism of the DNA-directed DNA polymerase activity of AMV reverse transcriptase.

The potential uses of misincorporation mutagenesis in the growing field of protein engineering are illustrated in Chapter 11.

CHAPTER 11 The application of *in vitro* mutagenesis to protein engineering.

Over the last decade, recombinant DNA technology has advanced rapidly and many genes of interest have been cloned and sequenced. Mutagenesis techniques have allowed both gross and subtle changes in the sequence of the cloned DNA, altering the subsequent expression of the gene or the structure and function of the encoded protein. Thus existing proteins can be "engineered" by introducing changes in the DNA sequence which alter the eventual structure, properties and/or function of the protein.

The possible applications of protein engineering are diverse but can be divided into two categories: those of commercial interest and those of academic interest. The two categories cannot be divorced from each other. The first category involves commercial exploitation of protein engineering techniques in particular with regard to the industrial production of enzymes or hormones. Useful developments in protein engineering could include making specific alterations to improve the stability or activity of a protein, or tailoring the substrate specificities of enzymes to reduce production costs in an industrial process by the use of cheaper raw materials.

The second category of applications of protein engineering could be titled "academic interest" and the results will be indispensible to the commercial activities. Such applications will involve the testing of predicted

roles of DNA sequences, amino acid interactions and hydrogen bonding in protein folding, and the testing of hypotheses of protein functions.

Functional domains of a protein may include amino-acids from different regions of the protein's primary sequence by virtue of the folding pattern of that protein. The overall structure of the protein can be determined by X-ray crystallography but this is a very time consuming process, averaging between five and ten man-years and to date only about 200 proteins have had their structures determined in this manner (Blundell and Sternberg, 1985).

The ability to make specific structural changes to a protein allows the systematic study of the structure and function of such a protein, and the interdependence of one with the other. Naturally occurring mutants and mutants induced *in vivo*, facilitate initial investigatio examiningns the effects of such changes. However, the isolation and recovery of such mutants is usually dependent upon a phenotypic change occurring as a result of the mutation. Whilst many such phenotypic changes are due to alterations in the expression of a gene, the location of the active-site(s) of a protein can be pinpointed by such investigations. Furthermore, the introduction of random or specific alterations into a gene *in vitro* will,on expression, produce engineered proteins that will allow the elucidation of interactions of other regions of the protein with the active site. In the absence of X-ray crystallography the effects of the *in vitro* mutageneses

may suggest details of the structure of the protein, but in conjunction with X-ray crystallography the interactions of different proteins domains can be determined.

11.1 Structural analysis of proteins.

11.1.1 Protein crystallography.

X-ray crystallography of proteins requires crystals of the pure protein in millimeter dimensions thus requiring tens of milligrams of pure protein. Such large amounts of pure protein limited crystallographers initially to proteins naturally produced in large amounts. With the advent of recombinant cloning techniques, eukaryotic genes can be cloned in vectors allowing their expression in prokaryotes. Use of multi-copy expression vectors facilitates the isolation and purification of proteins normally produced in low amounts. The use of yeast vectors is becoming increasingly important for the expression of cloned genes because they permit the glycosylation and secretion of proteins.

X-ray crystallography of crystals of purified protein generates electron densitiy maps which, together with the amino acid sequence, allow a three dimensional model to be built up. Where the resolution is 3Å or better, then the positions of nearly all the hydrogen bonds can be determined. One interesting example is that of tyrosyl tRNA synthetase of *B. stearothermophilus*, for which the structure has been solved after crystallization by Bhat and co-workers

(1982). Many mutagenesis experiments have been carried out to test predictions of the structure and function of this enzyme (see 11.2.1). The addition of tyrosine and ATP to the crystals identified the active site of the enzyme so indicating possible hydrogen bonds between the enzyme and the substrate.

Crystallization has only really been successful when using single polypeptides or multimers of single polypeptides. Thus it is not yet possible to study protein-protein interactions by X-ray crystallography because it is not possible to co-purify or co-crystallize the two or more proteins involved.

11.1.2 Computer graphics

Substitutions of one or a few amino acids which do not affect the gross folding pattern of the protein might change side-chain contacts between the enzyme and its substrate or distort the polypeptide chain within the active site. When a three dimensional model of the enzyme has been developed then computer graphics can be used to observe the effects of such substitutions, predicted on the basis of a knowledge of of protein folding. These predictions can be tested by a computer-assisted simple Fourier difference analysis of the mutant and wild-type crystals, but substitutions which drastically affect the protein folding pattern require *de novo* crystallization and structure determination.

11.1.3 Nuclear magnetic resonance.

Nuclear magnetic resonance can provide information on protein dynamics. It is not necessary to crystallize proteins, and resonances from the protons of the protein in solution can be assigned to individual amino acids (Searle *et al.*, 1986). Studies on dihydrofolate reductase have provided detailed structural information on the complexes of the enzyme with small molecules such as NADP+ and thioNADP+ (Feeney *et al.*, 1983). The three dimensional co-ordinates of the molecule can be established from the assigned resonances but as yet the research is at an early stage and can only be applied to small polypeptides. However, combined with information from protein crystallography and molecular graphics, a detailed picture of the structure, function and dynamics of a protein can be built up (Birdsall *et al.*, 1984) which in turn can be examined further by the introduction of site specific mutations *in vitro.*

The combination of the techniques of protein crystallography, computer graphics and nuclear magnetic resonance in the analysis of wild-type and engineered proteins increases our understanding of the structure and function of proteins, and in turn increases our understanding of the relationships between the amino acid sequence and the folding patterns of proteins. Correlation of the functional differences of proteins and the structural

differences observed experimentally should ultimately allow the development of predictive rules for protein engineering.

11.1.4 Protein folding

The tertiary structure of a protein is the result of the folding of an amino acid chain. Examination of the folding patterns of different proteins should allow predictions to be made about the structure of related proteins so developing an understanding of the interactions between different sequences of amino acids. At the moment we are unable to predict reliably three dimensional structures from a knowledge of amino acid sequences due to the inaccuracy of the energy functions used in these predictions. Busetta (1986) has used protein homologies within large families of distantly related proteins to acquire average formation energies of nuclei potentially involved in folding. This has allowed improvements in the predictions of protein structures. Rational modification of peptides is now possible based upon computer-aided model building and predictions of secondary structures. A synthetic somatostatin peptide has been made that is more biologically active than natural somatostatin, retains the secondary structure of the natural peptide, but is smaller on account of a single proline replacing nine amino acid residues (Veber *et al.*, 1981). Thus as our understanding of protein folding pathways improves (reviewed by King, 1986) we will be able to apply protein engineering techniques more

widely and design proteins *de novo* (see review by Van Brunt, 1986).

11.2 Functional analysis of proteins

Described below are two examples illustrating how protein engineering has deepened our understanding of how proteins function.

11.2.1 Tyrosyl t-RNA synthetase

With a detailed knowledge of the structure of the enzyme tyrosyl tRNA synthetase, the groups of Fersht and Winter have used *in vitro* mutagenesis to alter residues within the enzyme. (Winter *et al.*, 1982; Leatherbarrow *et al.*, 1985). They have been able to study and quantify the role of binding energy in catalysis, measuring the energetic value of hydrogen bonding, and have shown that many small binding-energy interactions play a role delocalising the catalysis around the active site (reviewed by Fersht *et al.*, 1986). It was possible to show that the mechanism of catalysis involved only binding energies of side-chain interactions which stabilise the transition state. Furthermore, it has been found that linear free energy relationships (LFER) exist not only for simple organic reactions, but also for the reactions of mutant tyrosyl tRNA synthetases with modified structures. Some 15 mutations fitted an LFER which showed how closely the transition state

resembled either the starting materials or the product of the reaction (Fersht *et al.*, 1986).

The use of LFER states will remove the need for the crystallisation of closely related mutant proteins in investigations into the relationship between structure and function.

Thus protein engineering has been used as a tool in the systematic anlaysis of the mechanism of action of an enzyme.

11.2.2 Mammalian *ras*

11.2.2.1 *Ras* mutations isolated *in vivo*

An activated c-Ha-*ras* oncogene was isolated from a human bladder carcinoma cell line T24 (Goldfarb *et al.*, 1982; Shih and Weinberg, 1982) and was identified by its homology to the oncogene of the transforming retrovirus Harvey Rat Sarcoma virus. The activation of c-Ha-*ras* in T24 and EJ bladder carcinomas was found to be a consequence of a single amino acid substitution of valine for glycine at position 12 (Tabin *et al.*, 1982; Reddy *et al.*, 1982; Taparowsky *et al.*, 1982).

Other activating mutations have been discovered by the isolation of oncogenic *ras* genes and have been shown to be due to amino acid substitutions of valine, cysteine, arginine or aspartic acid at position 12 (Capon *et al.*, 1983b; Shimizu *et al.*, 1983; Santos *et al.*, 1984; Tainsky et al., 1984) or substitutions at postions 13 or 61 (Yuasa *et al.*, 1983; Bos *et al.* , 1984; Bos *et al.*, 1985).

The *ras* protein p21 binds GTP and GDP, (Papageorge *et al.*, 1982; Finkel *et al.*, 1984) and natural activating mutations of *ras* have been shown to reduce the intrinsic GTPase activity of the protein providing a link between mutations and the function of the protein (McGrath *et al.*, 1984; Sweet *et al.*, 1984; Gibbs *et al.*, 1984). Sequence analyses have highlighted regions of homology between *ras* p21 and G proteins (Gay and Walker, 1983; Leberman and Egner, 1984; Hurley *et al.*, 1984), and these have allowed comparisons of *ras* p21 and structural analogues.

Halliday (1984) has alligned the sequence of p21 with that of the bacterial elongation factor EF-Tu for which there exists a crystal structure, showing that 42% of the p21 amino acids have identical or conservative equivalents in EF-Tu. From this information McCormick and co-workers (1985) have proposed a model for the tertiary structure of p21.

11.2.2.2 Mutations of *ras* introduced *in vitro*

In vitro mutagenesis techniques have been used to study the *ras* oncogene in an attempt to understand the links between the predicted activities of the protein, its structure, function and its transforming potential. The introduction of changes that alter an activity of the protein, *eg.* GTP binding, enable links to be made between the predicted protein structure and its transforming activity.

Fasano and co-workers (1984) subjected the cloned wild-type human c-Ha-*ras* gene to modification by sodium bisulphite. Two regions were treated separately, the first encoding amino acids 1-37 and the second encoding amino acids 38-110. The protein (p21) consists of 189 amino acids and is a phosphoprotein of molecular weight 21,000. Most of the bisuphite induced mutations did not activate the transforming potential of *ras*, but substitutions of amino acids 12, 13, 59 or 63 did activate *ras*. This information along with that of *in vivo* activating mutations at positions 12, 13 and 61 suggested two main sites within p21 where alteration of the amino acid residues converted the cellular proto-oncogene into the activated oncogene.

Seeburg and co-workers (1984) used site-directed mutagenesis to produce 20 different Ha-*ras* proteins each with a different amino acid at position 12. All except two of the substitutions resulted in activation of p21, the two exceptions being proline and wild-type glycine. Structural analyses of other proteins have shown that both proline and glycine residues tend to break polypeptide α-helices suggesting that activation of p21 could be due to a conformational change at this site.

A similar study of *in vitro* mutagenesis of *ras* changed the wild-type amino acid glutamine at position 61 for 17 different amino acids. Only two of the substitutions failed to activate the transforming potential of the resulting p21 and these were proline or glutamic acid. The other 15 mutants had widely differing transforming activities all of

which were were greater than the wild-type p21 (Der *et al.*, 1986).

The function of the *ras* protein in mammalian cells has not yet been discovered although yeast RAS proteins are known to be required for the stimulation of membrane-bound adenylate cyclase and the mammalian *ras* protein can substitute for yeast RAS in yeast cells. Application of reverse transcriptase-directed misincorporation mutagenesis to the activated *ras* oncogene, using a bank of extended primers to define sites for misincorporation, would generate many mutants. Analysis of such mutants with altered activities of p21 (but not fully inactivated) will facilitate study of the interactions of different domains of the protein and, combined with the crystal structure of the protein (not yet available), will lead to a greater understanding of the structural basis for the action of the *ras* protein.

BIBLIOGRAPHY

Abarzúa, P. and Marians, K.J. (1984). Enzymatic techniques for the isolation of random single-base substitutions in vitro at high frequency. Proc. Natl. Acad. Sci. USA. 81, 2030-2034.

Aboul-ela, F., Koh, D., Tinoco Jr, I. and Martin, F. (1985). Base-base mismatches. Thermodynamics of double helix formation for dCA_3XA_3G and dCT_3YT_3G (X,Y=A,C,G,T). Nucl. Acids Res. 13, 4811-4824.

Baas, P.D., and Jansz, H.S. (1972). Asymmetric information transfer during ØX174 DNA replication. J. Mol. Biol. 63, 557-568.

Baltimore, D. (1970). RNA-dependent DNA polymerase in virions of RNA tumour viruses. Nature (London) 226, 1209-1211.

Battula, N. and Loeb, L.A. (1974). The infidelity of avian myeloblastosis virus deoxyribonucleic acid polymerase in polynucleotide replication. J. Biol. Chem. 249, 4086-4093.

Battula, N. and Loeb, L.A. (1976). On the fidelity of DNA replication. Lack of exodeoxyribonculease activity and error-correcting function in avian myeloblastosis virus DNA polymerase. J. Biol. Chem. 251, 982-986.

Bauer, J., Krämmer, G. and Knippers, R. (1981). Asymmetric repair of bacteriophage T7 heteroduplex DNA. Mol. Gen. Genet. 181, 541-547.

Bernardi, F. and Ninio, J. (1978). The accuracy of DNA replication. Biochimie 60: 1083-1095.

Bhat, T.N. Blow, D.M., Brick, P. and Nyborg, J. (1982). Tyrosyl-tRNA synthetase forms a mononucleotide - binding fold. J. Mol. Biol. 158,. 699-709.

Birdsall, B., Feeney, J., Pascual, C., Roberts, G.C.K., Kompis, I., Then, R.L., Muller, K. and Kroehn, A. (1984). A ^1H NMR study of the interactions and conformations of rationally designed brodimoprim analogues in complexes with Lactobacillus casei dihydrofolate reductase. J. Med. Chem. 27: 1672-1676.

Blundell, T. and Sternberg, M.J.E. (1985). Computer aided design in protein engineering. Trends Biotech. 3, 228-235

Bos, J.L., Toksov, D., Marshall, C.J., Verlaan de Vries, M., Veeneman, G.H., van der Eb, A.J., van Boom, J.H., Jannsen, J.W.G. and Steenvoorden, A.C.M. (1985). Amino-acid substitutions at codon 13 of the N-ras oncogene in human acute myeloid leukaemia. Nature 315, 726-730.

Bos, J.L., Verlaan-de Vries, M., Jansen, A.M., Veeneman, G.H., van Boom, J.H., van der Eb, A.J. (1984). Three different mutations in codon 61 of the human N-ras gene detected by synthetic olionucleotide hybridisation. Nucl. Acids Res. 12, 9155-9163.

BIBLIOGRAPHY

Abarzúa, P. and Marians, K.J. (1984). Enzymatic techniques for the isolation of random single-base substitutions in vitro at high frequency. Proc. Natl. Acad. Sci. USA. 81, 2030-2034.

Aboul-ela, F., Koh, D., Tinoco Jr, I. and Martin, F. (1985). Base-base mismatches. Thermodynamics of double helix formation for dCA_3XA_3G and dCT_3YT_3G (X,Y=A,C,G,T). Nucl. Acids Res. 13, 4811-4824.

Baas, P.D., and Jansz, H.S. (1972). Asymmetric information transfer during ØX174 DNA replication. J. Mol. Biol. 63, 557-568.

Baltimore, D. (1970). RNA-dependent DNA polymerase in virions of RNA tumour viruses. Nature (London) 226, 1209-1211.

Battula, N. and Loeb, L.A. (1974). The infidelity of avian myeloblastosis virus deoxyribonucleic acid polymerase in polynucleotide replication. J. Biol. Chem. 249, 4086-4093.

Battula, N. and Loeb, L.A. (1976). On the fidelity of DNA replication. Lack of exodeoxyribonculease activity and error-correcting function in avian myeloblastosis virus DNA polymerase. J. Biol. Chem. 251, 982-986.

Bauer, J., Krämmer, G. and Knippers, R. (1981). Asymmetric repair of bacteriophage T7 heteroduplex DNA. Mol. Gen. Genet. 181, 541-547.

Bernardi, F. and Ninio, J. (1978). The accuracy of DNA replication. Biochimie 60: 1083-1095.

Bhat, T.N. Blow, D.M., Brick, P. and Nyborg, J. (1982). Tyrosyl-tRNA synthetase forms a mononucleotide - binding fold. J. Mol. Biol. 158,. 699-709.

Birdsall, B., Feeney, J., Pascual, C., Roberts, G.C.K., Kompis, I., Then, R.L., Muller, K. and Kroehn, A. (1984). A ^1H NMR study of the interactions and conformations of rationally designed brodimoprim analogues in complexes with Lactobacillus casei dihydrofolate reductase. J. Med. Chem. 27: 1672-1676.

Blundell, T. and Sternberg, M.J.E. (1985). Computer aided design in protein engineering. Trends Biotech. 3, 228-235

Bos, J.L., Toksov, D., Marshall, C.J., Verlaan de Vries, M., Veeneman, G.H., van der Eb, A.J., van Boom, J.H., Jannsen, J.W.G. and Steenvoorden, A.C.M. (1985). Amino-acid substitutions at codon 13 of the N-ras oncogene in human acute myeloid leukaemia. Nature 315, 726-730.

Bos, J.L., Verlaan-de Vries, M., Jansen, A.M., Veeneman, G.H., van Boom, J.H., van der Eb, A.J. (1984). Three different mutations in codon 61 of the human N-ras gene detected by synthetic olionucleotide hybridisation. Nucl. Acids Res. 12, 9155-9163.

Botstein, D. and Shortle, D. (1985). Strategies and applications of *in vitro* mutagenesis. Science *229*, 1193-1201.

Breathnach, R., Mandel, J.L. and Chambon, P. (1977). Ovalbumin gene is split in chicken DNA. Nature *270*, 314-319.

Brutlag, D. and Kornberg, A. (1972). Enzymatic synthesis of deoxyribonucleic acid. J. Biol. Chem. *247*: 241-248.

Busby, S., Irani, M. and de Crombrugghe, B. (1982). Isolation of mutant promoters in the *Escherichia coli* galactose operon using local mutagenesis on cloned DNA fragments. J. Mol. Biol. *154*, 197-209.

Busetta, B. (1986). Examination of folding patterns for predicting protein topologies. Biochim. Biophys. Acta. *870*, 327-338.

Capon, D.J., Chen, E.Y., Levinson, A.D., Seeburg, P.H. and Goeddel, D.V. (1983) Complete nucleotide sequence of the T24 human bladder carcinoma oncogene and its normal homologue. Nature. *302*, 33-37.

Carter, P., Bedouelle, H. and Winter, G. (1985). Improved oligonucleotide site-directed mutagenesis using M13 vectors. Nucl. Acids Res. *13*: 4431-4443.

Claverys, J-P. and Lacks, S.A. (1986). Heteroduplex deoxyribonucleic acid base mismatch repair in bacteria. Microb. Rev. *50*, 133-165.

Cohen, N.N., Chang, A.C.Y. and Hsu, L. (1972). Non-chromosomal antibiotic resistance in bacteria: genetic transformation of *Escherichia coli* by R factor DNA. Proc. Natl. Acad. Sci. USA. *69*: 2110-2114.

Cox, E.C. (1976). Bacterial mutator genes and the control of spontaneous mutation. Ann. Rev. Genet. *10*, 135-156.

Der, C.J., Finkel, T. and Cooper G.M. (1986). Biological and biochemical properties of human ras^H genes mutated at codon 61. Cell. *44*, 167-176.

Dodson, L.A., Foote. R.S., Mitra, S. and Masker, W.E. (1982). Mutagenesis of bacteriophage T7 *in vitro* by incorporation of O^6-methylguanine during DNA synthesis. Proc. Natl. Acad. Sci. USA *79*, 7440-7444.

Dunn, D.B. and Smith, J.D. (1958). The occurrence of 6-methylaminopurine in deoxyribonucleic acids. Biochem. J. *68*: 627-636.

Eadie, J.S., Conrad, M., Toorchen, D. and Topal, M.D. (1984). Mechanism of mutagenesis by O^6-methylguanine. Nature *308*, 201-203.

Eckstein, F. (1985). Nucleoside phosphorothioates. Ann. Rev. Biochem. *54*, 367-402.

Enea, V., Vovis, G.F. and Zinder, N.D. (1975). Genetic studies with heteroduplex DNA of bacteriophage F1. Asymmetric segregation, base correction and implications for the mechanism of genetic recombination. J. Mol. Biol. 96, 495-509.

Eperon, I.C. (1986a). Rapid preparation of bacteriophage DNA for sequence analysis in sets of 96 clones, using filtration. Analytical Biochemistry 156, 406-412.

Eperon, I.C. (1986b). M13 vectors with T7 polymerase promoters: transcription limited by oligonucleotides. Nucl. Acid Res. 14, 2830.

Fasano, O., Aldrich, T., Tamanoi, F., Taparowsky, E., Furth, M. and Wigler, M. (1984). Analysis of the transforming potential of the human H-ras gene by random mutagenesis. Proc. Natl. Acad. Sci. USA. 81, 4008-4012.

Feeney, J., Birdsall, B., Roberts, G.C.K. and Burgen, A.S.V. (1983). Use of transferred nuclear overhauser effect measurements to compare binding of co-enzyme analogues to dihydrofolate reductase. Biochemistry 22, 628-633.

Fersht, A.R., Leatherbarrow, R.J. and Wells, T.N.C. (1986). Quantitative analysis of structure-activity relationships in engineered proteins by linear free-energy relationships. Nature 322: 284-286.

Fersht, A.R., Shi, J-P., and Tsui, W.C. (1983). Kinetics of base misinsertion by DNA polymerase I of Escherichia coli. J. Mol. Biol. 165, 655-667.

Finkel, T., Der, C.J. and Cooper, G.M. (1984). Activation of ras genes in human tumours does not affect localization, modification or nucleotide binding properties of p21. Cell 37, 151-158.

Fischer, S.G. and Lerman, L.S. (1983). DNA fragments differing by single base-pair substitutions are separated in denaturing gradient gels: correspondence with melting theory. Proc. Natl. Acad. Sci. USA. 80: 1579-1583.

Fritz, H-J. (1985). The oligonucleotide-directed construction of mutations in recombinant filamentous phage. In: DNA cloning - a practical approach. Vol 1.(Ed. Glover, D.M.) IRL pages 151-164.

Gay, N.J., Walker, J.E. (1983). Homology between human bladder carcinoma oncogene product and mitochondrial ATP-synthase. Nature 301, 262-264.

Gibbs, J.B., Sigal, I.S., Poe, M. and Scolnick, E.M. (1984). Intrinsic GTPase activity distinguishes normal and oncogenic ras p21 molecules. Proc. Natl. Acad. Sci. USA. 81, 5704-5708.

Gillam, S. and Smith, M. (1979). Site specific mutagenesis using synthetic oligodeoxyribonucleotide primers. I. Optimum conditions and minimum oligodeoxyribonucleotide length. Gene 8, 81-97.

Glickman, B.W. and Radman, M. (1980). Escherichia coli mutator mutants deficient in methylation-instructed DNA mismatch correction. Proc. Natl. Acad. Sci. USA. 77, 1063-1067.

Glickman, B., van den Elsen, P., Radman, M. (1978). Induced mutagenesis in dam⁻ mutants of Escherichia coli: a role for 6-methyladenine residues in mutation avoidance. Mol. Gen. Genet. 163, 307-312.

Gold, M., Hurwitz, J. and Anders, M. (1963). The enzymatic methylation of RNA and DNA. Biochem. Biophys. Res. Commun. 11: 107-114.

Goldfarb, M., Shimizu, K., Perucho, M. and Wigler, M. (1982). Isolation and preliminary characterization of a human transforming gene from T24 bladder carcinoma cells. Nature. 296, 404-409.

Goodman, N.C. and Spiegelman, S. (1971). Distinguishing reverse transcriptase of an RNA tumour virus from other known DNA polymerases. Proc. Natl. Acad. Sci. USA. 68, 2203-2206.

Grandgenett, D.P., Gerard, G.F. and Green, M. (1973). A single subunit from avian myeloblastosis virus with both RNA-directed DNA polymerase and ribonuclease H activity. Proc. Natl. Acad. Sci. USA. 70, 230-234.

Grandgenett, D.P., Golomb, M., and Vora, A.C. (1980). Activation of an Mg^{2+}-dependent DNA endonuclease of avian myeloblastosis virus αβDNA polymerase by in vitro proteolytic cleavage. J. Virol. 33, 264-271.

Grandgenett, D.P., Vora, A.C. and Schiff, R.D. (1978). A 32,000 dalton nucleic acid-binding protein from avian retrovirus cores possesses DNA endonuclease activity. Virology. 89, 119-132.

Gray, H.B., Ostrander, D.A., Hodnett, J.L., Legerski, R.J. and Robberson, D.L. (1975). Extracellular nucleases of Pseudomonas BAL 31. 1. Characterization of single strand-specific deoxyribonuclease and double-strand deoxyribonuclease activities. Nucl. Acids Res. 2: 1459-1492.

Grossberger, D., Clough, W. (1981) Incorporation into DNA of the base analogue 2-aminopurine by Epstein-Barr virus induced DNA polymerase in vivo and in vitro. Proc. Natl. Acad. Sci. USA. 78, 7271-7275.

Guo, L-H., Wu, R. (1983). Exonuclease III: use for DNA sequence analysis and in specific deletions of nucleotides. Methods Enzymol. 100, 60-96.

Hall, Z.W. and Lehman, I.R. (1968). An in vitro transversion by a mutationally altered T4-induced DNA polymerase. J. Mol. Biol. 36, 321-333.

Halliday, K. (1984). Regional homology in GTP-binding proto-oncogene products and elongation factors. J. Cyclic. Nucleotide Protein Phosphorylation Res. 9, 435-448.

Henikoff, S. (1984). Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene 28, 351-359.

Hillebrand, G.G. and Beattie, K.L. (1985). Influence of template primary and secondary structure on the rate and fidelity of DNA sythesis. J. Biol. Chem. 260, 3116-3125.

Hillebrand, G.G., McCluskey, A.H., Abbott, K.A., Revich, G.G. and Beattie, K.L. (1984). Misincorporation during DNA synthesis, analysed by gel electrophoresis. Nucl. Acids Res. 12, 3155-3171.

Hopfield, J.J. (1980). The energy relay: a proofreading scheme based on dynamic co-operativity and lacking all characteristic symptoms of kinetic proofreading in DNA replication and protein synthesis. Proc. Natl. Acad. Sci. USA. 77: 5248-5252.

Hunter, W.N., Brown, T., Anand, N.N. and Kennard, O. (1986). Structure of an adenine-cytosine base pair in DNA and its implications for mismatch repair. Nature. 320, 552-555.

Hurley, J.B., Simon, M.I., Teplow, D.B., Robishaw. J.D. and Gilman, A.G. (1984). Homologies between signal transducing G proteins and ras gene products. Science. 226, 860-862.

Hurwitz, J. and Leis, J.P. (1972). RNA-dependent DNA polymerase activity of RNA tumour viruses. J. Virol. 9, 116-129.

Hutchinson, C.A., Phillips, S., Edgell, M.H., Gillam, S., Jahnke, P. and Smith, M. (1978). Mutagenesis at a specific position in a DNA sequence. J. Biol. Chem. 253, 6551-6560.

Jones, M., Wagner, R. (1981). N-Methyl-N^1-nitro-N-nitrosoguanidine sensitivity of E. coli mutants deficient in DNA methylation and mismatch repair. Mol. Gen. Genet. 184, 562-563.

Kacian, D.L., Spiegelman, S., Bank, A., Terada, M., Metafora, S., Dow, L. and Marks, P.A. (1972). In vitro synthesis of DNA components of human genes for globins. Nature New Biol. 235, 167-169.

Kadonaga, J.T. and Knowles, J.R. (1985). A simple and efficient method for chemical mutagenesis of DNA. Nucl. Acids Res. 13, 1733-1745.

King, J. (1986). Genetic analysis of protein folding pathways. Bio/Technology 4, 297-303.

Kornberg, T. and Kornberg, A. (1974). Bacterial DNA polymerases. Pages 119-144. In: The Enzymes Vol X, Academic Press. N.Y.

Kramer, W., Schughart, K. and Fritz, H-J. (1982). Directed mutagenesis of DNA cloned in filamentous phage: influence of hemimethylated GATC sites on marker recovery from restriction fragments. Nucl. Acids Res. 10, 6475-6485.

Kramer, W., Drutsa, V., Jansen H-W., Kramer, B., Pflugfelder, M. and Fritz, H-J. (1984a) The gapped duplex DNA approach to oligonucleotide-directed mutation construction. Nucl. Acids Res. 12, 9441-9456.

Kramer, B., Kramer, W. and Fritz, H-J. (1984b) Different base/base mismatches are corrected with different efficiences by the methyl-directed DNA mismatch repair systems of E. Coli. Cell 38, 879-887.

Kunkel, T.A., (1985). Rapid and efficient site specific mutagenesis without phenotypic selection. Proc. Natl. Acad. Sci. USA. 82, 488-492.

Kunkel, T.A., Eckstein, F., Mildvan, A.S., Koplitz, R.M. and Loeb, L.A. (1981). Deoyribonucleoside [1-thio] triphospates prevent proofreading during in vitro DNA synthesis. Proc. Natl. Acad. Sci. USA. 78, 6734-6738.

Kunkel, T.A., and Loeb, L.A. (1979). On the fidelity of DNA replication. Effect of divalent metal ion activators and deoxyribonucleoside triphophate pools on in vitro mutagenesis. J. Biol. Chem. 254, 5718-5725.

Kunkel, T.A. and Loeb, L.A. (1981). Fidelity of mammlian DNA polymerases. Science 213: 765-767.

Kunkel, T., Meyer, R.R. and Loeb, L.A. (1979). Single-strand binding protein enhances fidelity of DNA synthesis in vitro. Proc. Natl. Acad. Sci. USA. 76, 6331-6335.

Lacks, S.A., Dunn, J.J. and Greenberg, B. (1982). Identification of base mismatches recognized by the heteroduplex-DNA-repair system of Strepococcus pneumoniae. Cell 31, 327-336.

Lathe. R.F., Lecocq, J.P. and Everett, R. (1983). DNA engineering: the use of enzymes, chemicals and oligonucleotides to restructure DNA sequences in vitro. In: Genetic Engineering Vol 4 p 1-56. (Ed. Robert Williamson, Academic Press).

Leatherbarrow, R.J., Fersht, A.R. and Winter, G. (1985). Transition-state stabilization in the mechanism of tyrosyl-tRNA synthetase revealed by protein engineering. Proc. Natl. Acad. Sci. USA. 82, 7840-7844.

Leberman, R. and Egner, U. (1984). Homologies in the primary structure of GTP-binding proteins: the nucleotide-binding site of EF-Tu and p21. EMBO J. 3: 339-341.

Loeb, L.A. and Kunkel, T.A. (1982) Fidelity of DNA synthesis. Ann. Rev. Biochem. 52, 429-457.

Lu, A.L., Clark, S. and Modrich, P. (1983). Methyl-directed repair of DNA base-pair mismatches in vitro. Proc. Natl. Acad. Sci. USA. 80, 4639-4643.

Maniatis,T., Fritsch, E.F., Sambrook, J. (1982). Synthesis of the second cDNA strand. In: Molecular Cloning, Cold Spring Harbor Laboratory, p214.

Marinus, M.G. (1984). Methylation of prokaryotic DNA. In: DNA Methylation. A. Razin, H.Cedar and A.D.Riggs (Eds) Springer-Verlag, New York. p81-109.

Marinus, M.G. and Norris, N.R. (1974). Biological function for 6-methyladenine residues in the DNA Escherichia coli K12. J. Mol. Biol. 85, 309-322.

Marmenout, A., Remaut, E., van Boom, J. and Fiers, W. (1984). Oligonucleotide directed mutagenesis: selection of mutants by hemimethylation of GATC-sequences. Mol. Gen. Genet. 195, 126-133.

Matteucci, M.D. and Heynecker, H.L. (1983). Targeted random mutagenesis: the use of ambiguously synthesized oligonucleotides to mutagenize sequences immediately 5′ of an ATG initiation codon. Nucl. Acids Res. 11, 3113-3121.

Matthes, H.W.D, Zenke, W. M., Grundström, T., Staub, A., Wintzerith, M. and Chambon, P. (1984). Simultaneous rapid chemical synthesis of over one hundred oligonucleotides on a microscale. EMBO J. 3, 801-805.

Maxam, A. and Gilbert, W. (1980) Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. 65, 499-560.

McCormick, F., Clark, B.F.C., Lacour, T.F.M., Kjeldgaard, M., Norskov-Lauritsen, L., Nyborg, J. (1985). A model for the tertiary structure of p21, the product of the ras oncogene. Science 230, 78-82.

McGrath, J.P., Capon, D.J., Goeddel, D.V. and Levinson, A.D. (1984). Comparative biochemcal properties of normal and activated human ras p21 protein. Nature 310, 644-649.

Messing, J. (1983). New M13 vectors for cloning. Methods Enzymol. 101, 20-78.

Mölling, K., Bolognesi, D.P., Bauer, W.B., Büsen, W., Plassmann, H.W. and Hausen, P. (1971). Association of viral reverse transcriptase with an enzyme degrading the RNA moiety of RNA-DNA hybrids. Nature New Biol. 234, 240-243.

Mott, J.E., Van Arsdell, J. and Platt, T. (1984). Targeted mutagenesis in vitro: lac repressor mutations generated using AMV reverse transcriptase and dBrUTP. Nucl. Acids Res. 12, 4139-4152.

Müller, W., Weber, H., Meyer, F. and Weissmann, C. (1978). Site directed mutagenesis in DNA: generation of point mutations in cloned β-globin complementary DNA at the positions corresponding to amino acids 121 to 123. J. Mol. Biol. 124, 343-358.

Myers, R.M., Fischer, S.G., Lerman, L.S. and Maniatis, T. (1985a). Nearly all single base substitutions in DNA fragments joined to a GC-clamp can be detected by denaturing gradient gel electrophoresis. Nucl. Acids Res. 12, 3131-3145.

Myers, R.M., Fischer, S.G., Maniatis, T. and Lerman, L.S. (1985b). Modification of the melting properties of a duplex DNA by attachment of a GC-rich DNA sequence as determined by denaturing gradient gel electrophoresis. Nucl. Acids Res. 13,3111-3129.

Myers, R.M., Lerman, L.S. and Maniatis, T. (1985c) A general method for saturation mutagenesis of cloned DNA fragments. Science 229, 242-247.

Papageorge, A., Lowy, D. and Scolnick, E.M. (1982). Comparative biochemical properties of p21 ras molecules coded for by viral and cellular ras genes. J. Virol. 44, 509-519.

Pukkila, P.J., Peterson, J., Herman, G., Modrich, P. and Meselson, M. (1983). Effects of high levels of DNA adenine methylation in methyl-directed mismatch repair in Escherichia coli. Genetics. 104, 571-582.

Putney, S.D., Benkowitz, S.J. and Schimmel, P.R. (1981). A DNA fragment with an α-phosphorothioate nucleotide at one end is asymmetrically blocked from digestion by exonuclease III and can be replicated in vivo. Proc. Natl. Acad. Sci. USA. 78, 7350-7354.

Radman, M., Wagner, R.E., Glickman, B.W. and Meselson, M. (1980). DNA methylation, mismatch correction and genetic stability. In: Progress in environmental mutagenesis. M.Alacevic (Ed). Elsevier/North-Holland, Amsterdam. p121-130.

Radman, M., Dohet, C., Bourguignon, M-F., Doubleday, D.P. and Lecombe, P. (1981). High Fidelity devices in the reproduction of DNA. In: Chromosome damage and repair. E. Seeberg and K. Kleppe (Eds). Plenum Press, New York. p431-445.

Razin, A., Hirose, T., Itakura, K. and Riggs, A.D. (1978). Efficient correction of a mutation by use of a chemically synthesized DNA. Proc. Natl. Acad. Sci. USA. 75, 4268-4270.

Reddy, E.P., Reynolds, R.K., Santos, E. and Barbacid, M. (1982). A point mutation is responsible for the acquisition of transforming properties by the T24 human bladder carcinoma oncogene. Nature. 300, 149-152.

Robert, M.S., Smith, R.G., Gallo, R.C., Sarin, P.S. and Abrill, J.W. (1972). Viral and cellular DNA polymerase: comparison of activities with synthetic and natural RNA templates. Science 176, 798-800.

Ross, J., Aviv, H., Scolnick, E. and Leder, P. (1972). In vitro synthesis of DNA complementary to purified rabbit globin mRNA. Proc. Natl. Acad. Sci. USA. 69, 264-268.

Rydberg, B. (1977). Bromouracil mutagenesis in Escherichia coli: evidence for involvment of mismatch repair. Mol. Gen. Genet. 152, 19-28.

Rydberg, B. (1978). Bromouracil mutagenesis and mismatch repair in mutator strains of Escherichia coli. Mutat. Res. 52, 11-24.

Sanger, F., Nicklen, S. and Coulson, A.R. (1977). DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA. 74: 5463-5467.

Sanger, F. and Couson, A.R. (1975). A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. J. Mol. Biol. 94, 441-448.

Santos, E., Martin-Zanca, D., Reddy, E.P., Pierotti, M.A., Della-Porta, G. and Barbacid, M. (1984). Malignant activation of a K-ras oncogene in lung carcinoma but not in normal tissue of the same patient. Science 223, 661-664.

Schaaper, R.M., Kunkel, T.A. and Loeb, L.A. (1983) Infidelity of DNA synthesis associated with bypass of apurinic sites. Proc. Natl. Acad. Sci. USA. 80, 487-491.

Schiff, R.D. and Grandgenett, D.P. (1980). Partial phosphorylation in vivo of the avian retrovirus pp32 DNA endonuclease. J.Virol. 36, 889-893.

Searle, M.S., Hammond, S.J., Birdsall, B., Roberts. G.C.K., Feeney, J., King, R.W. and Griffiths, D.V. (1986). Identification of the ^{1}H resonances of valine and leucine residues in dihydrofolate reductase by using a combination of selective deuteration and two-dimensional correlation spectroscopy. FEBS Lett. 194, 165-170.

Seeburg, P.H., Colby, W.W., Capon, D.J., Goeddel, D.V. and Levinson, A.D. (1984). Biological properties of human c-Ha-ras 1 genes mutated at codon 12. Nature 312, 71-75.

Shenk, T.E., Carbon, J., Berg, P. (1976). Construction and analysis of viable deletion mutants of simian virus 40. J.Virol. 18, 664-671.

Shih, C. and Weinberg, R.A. (1982). Isolation of a transforming sequence from a human bladder cell line. Cell 29, 161-169.

Shimizu, K., Birnbaum, D., Ruley, M.A., Fasano, O., Suard, Y., Edlund, L., Taparowsky, E., Goldfarb, M. and Wigler, M. (1983). Structure of the Ki-ras gene of the human lung carcinoma cell line Calu-1. Nature 304, 497-500.

Shortle, D. and Botstein, D. (1983). Directed mutagenesis with sodium bisufite. Methods Enzymol. 100, 457-468.

Shortle, D., DiMaio, D. and Nathans, D. (1981). Directed mutagenesis. Ann. Rev. Genet. 15: 265-294.

Shortle, D., Grisafi, P., Benkovic, S. and Botstein, D. (1982). Gap misrepair mutagenesis: efficient site-directed induction of transition, transversion and frameshift mutations in vivo. Proc. Natl. Acad. Sci. USA. 79, 1588-1592.

Shortle, D. and Lin, B. (1985). Genetic analysis of staphylococcal nuclease: identification of three intragenic "global" suppressors of nuclease-minus mutations. Genetics 110, 539-555.

Shortle, D. and Nathans, D. (1978). Local mutagenesis: a method for generating viral mutations with base substitutions in preselected regions of the viral genome. Proc. Natl. Acad. Sci. USA. 75, 2170-2174.

Sirover, M.A. and Loeb, L.A. (1976). Infidelity of DNA synthesis in vitro: screening for potential metal mutagens or carcinogens. Science 194, 1434-1436.

Sirover, M.A. and Loeb, L.A. (1977). On the fidelity of DNA replication: effect of metal activators during synthesis with avian myeloblastosis virus DNA polymerase. J. Biol. Chem. 252, 3605-3610.

Skinner, J.A., Öhrlein, A. and Grummt, I. (1984). In vitro mutagenesis and transcriptional analysis of a mouse ribosomal promoter element. Proc. Natl. Acad. Sci. USA. 81, 2137-2141.

Smith, M. (1985). In vitro mutagenesis. Ann. Rev. Genet. 19, 423-462.

Smith, M. and Gillam, S. (1981) In: Genetic Engineering: Principles and Methods. Vol 3. J.K.Setlow and A. Hollaender, (Eds). Plenum Press, New York. p1-32.

Srinivasan, P.R. and Borek, E. (1964). Enzymatic alteration of nucleic acid structure. Science 145: 548-553.

Sweet, R.W., Yokoyama, S., Kamata, T., Feramisco, J.R., Rosenberg, M. and Gross, M. (1984). The product of ras is a GTPase and the T24 oncogenic mutant is deficient in this activity. Nature 311, 273-275.

Tabin, C.J., Bradley, S.M., Bargmann, C.I., Weinberg, R.A., Papageorge, A.G., Scolnick, E.M., Dhar, R., Lowy, D.R. and Chang, E.H. (1982). Mechanisms of activation of a human oncogene. Nature 300, 143-149.

Tainsky, M.A., Cooper, C.S., Giovanella, B.C. and Vande-Woude, G.F. (1984). An activated ras^N gene: detected in late but not early passage human PA1 teratocarcinoma cells. Science 225, 643-645.

Taparowsky, E., Suard, Y., Fasano, O., Shimizu, K., Goldfarb, M.P. and Wigler, M.P. (1982). Activation of the T24 bladder carcinoma transforming gene is linked to a single amino acid change. Nature 300, 762-765.

Temin, H.M. and Baltimore, D. (1972). RNA-directed DNA synthesis and RNA tumor viruses. Advan. Virus Res. 17, 129-186.

Temin, H.M. and Mizutani, S. (1970). RNA dependent DNA Polymerase in virions of Rous Sarcoma virus. Nature (London) 226, 1211-1213.

Topal, M.D., DiGuiseppi, S.R. and Sinha, N.K. (1980). Molecular basis for subsitution mutations. J. Biol. Chem. 255, 11717-11724.

Traboni, C., Ciliberto, G. and Cortese, R. (1984). Mutations in box-B of the promoter of a eukaryotic $tRNA^{Pro}$ gene affect rate of transcription, processing, and stability of the transcripts. Cell 36, 179-187.

Van Brunt, J. (1986). Protein architecture: Designing from the ground up. Bio/Technology 4, 277-283.

Veber, D.F., Freidinger, R.M., Schwenk, P.D., Palaveda Jr, W.J., Holly, F.W., Strachan, R.G., Nutt, R.F., Arison, B.H., Homnick, C., Randall, W.C., Glitzer, M.S., Saperstein, R. and Hirschmann, R. (1981). A potent cyclic hexapeptide analogue of somatostatin. Nature 292, 55-58.

Verma, I.M. (1977). The reverse transcriptase. Biochim. Biophys. Acta. 473, 1-37.

Verma, I.M. (1981). Reverse Transcriptase; pp 87-103, Chapter 6. In: The Enzymes Vol XIV. Academic Press. New York.

Verma, I.M. Temple, G.F., Fan, H. and Baltimore, D., (1972). In vitro synthesis of DNA complementary to rabbit reticulocyte 10S RNA. Nature New Biol. 235, 163-166.

Volckaert, G. and Fiers, W. (1977). A micromethod for base analysis of ^{32}P-labeled oligonucleotides. Analytical Biochemistry 83, 222-227.

Wagner, R., Dohet, C., Jones, M., Doutriaux, M.P., Hutchinson, F. and Radman, M. (1984). Involvement of Escherichia coli mismatch repair in DNA replication and recombination. Cold Spring Harbor Symp. Quant. Biol. 49: 611-615.

Wagner, R. Jr. and Meselson, M. (1976). Repair tracts in mismatched DNA heteroduplexes. Proc. Natl. Acad. Sci. USA. 73, 4135-4139.

Wallace, R.B., Johnson, P.E., Tanaka, S., Schöld, M., Itakura, K. and Abelson, J. (1980). Directed deletion of a yeast transfer-RNA intervening sequence. Science 209, 1396-1400.

Wells, R.D., FlÜgel, R.M., Larson, J.E., Schendel, P.F. and Sweet, R.W. (1972). Comparison of some reactions catalyzied by deoyribonucleic acid polymerase from avian myeloblastosis virus, Escherichia coli and Micrococcus luteus. Biochemistry 11, 621-629.

Wildenberg, J. and Meselson, M. (1975). Mismatch repair in heteroduplex DNA. Proc. Natl. Acad. Sci. USA. 72, 2202-2206.

Winter, G., Fersht, A.R., Wilkinson, A.J., Zoller, M. and Smith, M. (1982). Redesigning enzyme structure by site-directed mutagenesis: tyrosyl tRNA synthetase and ATP binding. Nature 299, 756-758.

Yanisch-Perron, C., Vieira, J. and Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and puc 19 vectors. Gene 33, 103-119.

Yanofsky, C., Cox, E.C. and Horn, V. (1966). The unusual mutagenic specificity of an Escherichia coli mutator gene. Proc. Natl. Acad. Sci. USA. 55: 274-281.

Yuasa, Y., Srivastava, S.K., Dunn, C.Y., Rhim, J.S., Reddy, E.P. and Aaronson, S.A. (1983). Acquisition of transforming properties by alternative point mutations within c-bas/has human proto-oncogene. Nature 303, 775-779.

Zakour, R.A., James, E.A. and Loeb, L.A. (1984). Site specific mutagenesis: insertion of single noncomplementary nucleotides at specified sites by error-directed DNA polymerization. Nucl. Acids Res. 12, 6615-6628.

Zakour, R.A. and Loeb, L.A. (1982). Site specific mutagenesis by error-directed DNA synthesis. Nature 295, 708-710.

Zoller, M.J. and Smith, M. (1982). Oligonucleotide-directed mutagenesis using M13-derived vectors: an efficient and general procedure for the production of point mutations in any fragment of DNA. Nucl. Acids Res. 10, 6487-6500.

Zoller, M.J. and Smith, M. (1983). Oligonucleotide-directed mutagenesis of DNA fragments cloned in M13 vectors. Methods Enzymol. 100, 468-500.

Misincorporation by AMV-reverse transcriptase and its potential for mutagenesis.

Judith A. Skinner

This thesis describes a systematic investigation of the efficiency of misincorporation by Avian Myeloblastosis Virus reverse transcriptase with all possible combinations of dNTP substrate, template nucleotide, and the nucleotide at the 3' terminus of the primer. Each of a series of 16 synthetic oligonucleotide primers was annealed to single-stranded M13 DNA templates, and a single dNTP was misincorporated at the primer 3' end using AMV reverse transcriptase. The proportion and pattern of misincorporation and incorporation in all 64 situations were assayed using 5'-labelled primers, and the products were separated on denaturing polyacrylamide gels.

Correct incorporations occurred more readily than misincorporations. The efficiency of misincorporation depended on the individual primer but, comparing primers, a clear dependence on the template nucleotide was observed for the preferential misincorporation of different dNTPs. The exact combination of template nucleotide and dNTP was important; although purine:pyrimidine (dNTP substrate: template nucleotide) and pyrimidine:purine misincorporations occurred comparatively readily, some pyrimidine:pyrimidine and purine:purine reactions were equally efficient and yet others were never seen to occur. Some misincorporations were facilitated by subsequent correct incorporations, but despite this the results suggest that the level of misincorporation is limited by the rate of reaction and enzyme inactivation rather than by exonuclease activity.

The recovery of point mutants arising from reverse transcriptase-directed misincorporation of single dNTPs onto single oligonucleotide primers is described and discussed.

Misincorporation of dNTPs is a form of *in vitro* mutagenesis which facilitates the generation of a library of point mutations spread throughout a gene. Conditions have been established in this study for the production of a bank of primers with 3' termini distributed over a region of a gene to be mutated. The misincorporation of single dNTPs onto the termini of such a bank of primers should allow the generation of a library of point mutants.

LEICESTER UNIVERSITY LIBRARY

19 JUN 1987