DESIGN, CONSTRUCTION AND PROPERTIES OF NOVEL PROTEIN MOLECULES

PROCEEDINGS OF
A ROYAL SOCIETY DISCUSSION MEETING
HELD ON 5 AND 6 JUNE 1985

ORGANIZED AND EDITED BY
D. M. BLOW, F.R.S., A. R. FERSHT, F.R.S., AND
G. WINTER

LONDON
THE ROYAL SOCIETY
1986

Printed in Great Britain for the Royal Society
by the
University Press, Cambridge

ISBN 0 85403 271 1

First published in *Philosophical Transactions of the Royal Society of London*,
series A, volume 317 (no. 1540), pages 291–451

Published by the Royal Society
6 Carlton House Terrace, London SW1Y 5AG

PREFACE

Until the present decade, the evolution of protein molecules had proceeded, since life began, through random mutation and natural selection. The techniques of nucleotide synthesis and directed mutagenesis have now provided a new mechanism. Genes, and the proteins which they encode, may be constructed or modified at will. Sensitive techniques exist to probe the properties of these factitious proteins. At least in certain cases, our insight into protein structure–function relations is enough to enable rudimentary designs to be made, aimed to create molecules with specific properties.

Together, these advances make up the new technology of protein engineering. In this volume we present a cross-section of activity in this new and burgeoning subject, presented at the Royal Society in June 1985. The work spans the scientific spectrum from cell biology through chemical catalysis and crystallography to computer science. These articles herald the new achievements that protein engineering is likely to bring in industrial chemistry, in materials science, and in medicine.

D. M. BLOW
A. R. FERSHT
G. WINTER

December 1985

CONTENTS

[One plate]

Phil. Trans. R. Soc. Lond. A **317**, 293–294 (1986)
Printed in Great Britain

Introduction

By D. M. Blow, F.R.S.
Biophysics Section, Blackett Laboratory, Imperial College, London SW7 2BZ, U.K.

Our understanding of the function of protein molecules was revolutionized in the 1960s by the use of X-ray crystallography to give a three-dimensional picture of their structures at atomic resolution. The structure of myoglobin was rapidly followed by the structure of several hydrolytic enzymes such as lysozyme, carboxypeptidase, ribonuclease, chymotrypsin, and subtilisin; and, not long after, by the much more complicated structure of haemoglobin, composed of four myoglobin-like molecules interacting with each other.

The first hydrolytic enzyme structures showed us how enzymes perform biological catalysis by immobilizing their substrates at the enzyme active site, and gave us definite ideas about the specific functions of different parts of the protein molecules. These ideas had to be treated as hypotheses, because there was no direct method to check them. A few particular points could be proved by cunning but tedious experiments.

In the 1970s Max Perutz pursued his pioneering studies on haemoglobin by exploring the wealth of human haemoglobin mutants which exist in the world's population. Some of these mutants cause important diseases in the third world (sickle-cell anaemia and β-thalassaemia), while others account for specific instances of anaemia and other blood disorders in individual families. Perutz found that many of these mutants allowed him to pinpoint the functions of specific amino acids in the haemoglobin molecule, and to correlate their altered properties with the structural changes they cause. These experiments were possible because of the relative ease of detecting the presence of an unusual haemoglobin molecule, by techniques available in hospital haematology departments all over the world.

In the 1970s new methods of reading the chemical sequence of genes were being developed, and these led to techniques which allowed genes to be transferred from one organism to another, and for the level of their expression to be controlled. These techniques provide a basic technology for protein engineering in the 1980s but further developments were needed in important directions.

First, methods of introducing specific mutations into a gene had to be found. These would allow genetic changes to be made at will and in a totally controlled fashion, instead of waiting for a chance mutational event which would need to be identified and exploited.

Second, our much deeper knowledge of protein function (especially enzyme function) had to be used to design accurate and sensitive methods to measure and to compare the functional properties of different mutants. Such changes can be expressed directly in terms of the energetics of the enzyme's interaction with its substrate.

Because of the complication of protein structures, computers have become essential aids in understanding them. Graphical displays which provide an accurate three-dimensional perception of the structure are invaluable. The third development in hand is to find how to

use the amazing computing power now accessible to us, to guide us in protein design by providing accurate predictions about a structure which has not yet been made.

In this Meeting the latest advances in all these techniques will be presented. A number of specific applications will be discussed, in which the interplay of these three strands of technical advance is exploited to achieve particular ends.

Phil. Trans. R. Soc. Lond. A **317**, 295–304 (1986)
Printed in Great Britain

Site-directed mutagenesis

By M. Smith

Department of Biochemistry, Faculty of Medicine, University of British Columbia, 2146 Health Sciences Mall, Vancouver, British Columbia, Canada V6T 1W5

The development of genetic engineering, whereby a specific gene or cDNA (c is copy or complementary) can be isolated as part of a minichromosome that can replicate and be expressed (by transcription and translation) in a living cell, has made possible *in vitro* techniques for micromanipulation (i.e. site-directed mutagenesis) of a cloned gene, to make defined changes in the portion of the gene that encodes its protein product. The methods by which this micromanipulation of a structural gene are effected fall under three broad headings: (i) the production of random single base-pair substitutions by chemical or enzymatic means; (ii) the construction of heteroduplex DNA by annealing single-strand target DNA with a chemically synthesized mutagenic oligonucleotide and (iii) the total or partial synthesis of mutant duplex DNA from chemically synthesized oligonucleotides. As a consequence it is now possible to modify a gene so that any amino acid in its product protein can be replaced by any other amino acid.

Introduction

Definition of the functional roles of the structural components of proteins has benefitted enormously from the comparison of amino-acid sequences of proteins with identical functions that have been isolated from different species. This is especially true when the three-dimensional structure of one or more of the proteins is available in atomic detail. The classical example of the application of this principle is provided by cytochrome *c* where the amino-acid sequences of at least sixty different eukaryote cytochromes *c* are known (Ferguson-Miller *et al.* 1979) and the detailed structures of tuna ferrocytochrome *c* and ferricytochrome *c* are established (Takano & Dickerson 1981 *a*, *b*). Thus, it is possible to establish the precise locations of the amino acids, approximately twenty-eight in number, that are conserved in all known eukaryote cytochromes *c* and which, presumably, have essential functional or structural roles. Comparison of the properties of different cytochromes *c* provides considerable insight into the roles of many of the conserved amino-acid residues (Ferguson-Miller *et al.* 1979). However, interpretation is always complicated by the multiplicity of differences in the amino-acid sequences of cytochromes *c* and this demonstrates the need for mutants obtained by site-directed mutagenesis (Pielak *et al.* 1985). A similar situation exists with haemoglobins (Perutz 1979, 1983).

More detailed understanding of the role(s) of amino-acid side chains in protein structure and function is possible when many point mutants are available in the structural gene of a protein for an individual species. The structural variants of the 146 amino-acid human β-globins, of which there are more than 200, provide an excellent example (McKusick 1983). In micro-organisms, fine-structure genetic analysis can provide a large number of mutants resulting from single amino-acid replacements and which provide detailed insights into protein function. Here, the archetypal example is the *lac* repressor (Miller 1978). Understanding of the function of this and other bacterial proteins has been greatly extended by suppressing nonsense mutations with

a battery of tRNA suppressors which allow insertion of several different amino acids at the site of the nonsense mutation (Miller 1984).

The advent of recombinant DNA technology, which makes possible the molecular cloning of individual genes and cDNA together with the reinsertion of the cloned gene into a living cell, has opened up the possibility of site-directed mutagenesis; the construction of specific mutations *in vitro* followed by *in vivo* analysis of the mutated gene (Shortle *et al.* 1981; Smith 1985; Watson *et al.* 1983). The methods for site-directed mutagenesis that are applicable to problems of protein structure and function are the subject of this article. Emphasis will be on the strategies which are most appropriate for the modification of protein structural genes to produce single amino-acid replacements in the protein of interest. A general recent review of *in vitro* mutagenesis has been prepared by Smith (1985). Clearly, the achievement of a specific codon replacement or of a family of codon replacements is only the first stage in obtaining a mutant protein. For this reason, most success to date has been with genes or cDNA that can be expressed and that give rise to an intact functional protein in *Escherichia coli* (Fersht *et al.* 1984; Inouye 1983; Shortle 1983; Shortle & Lin 1985; Villafranca *et al.* 1983; Winter & Fersht 1984). However, there is an increasing amount of work with homologous and heterologous mutant protein structural genes expressed in *Saccharomyces cerevisiae* (Brake *et al.* 1984; Pielak *et al.* 1985; Rosenberg *et al.* 1984) and in higher eukaryotes (Craik *et al.* 1985; Gil *et al.* 1985; Weinmaster *et al.* 1984). Studies on mutant proteins produced in eukaryote cells in culture are bound to increase in number and diversity because of their previous inaccessibility for structure and function analysis.

The manipulations of a structural gene which constitute an experiment in site-directed mutagenesis involve several discrete but essential steps. First, the DNA as an isolated fragment or as a component of a single- or double-strand DNA vector is modified by chemical or enzymatic reagents. Next, the modified DNA, incorporated into a DNA replicating vector, is used to transform bacterial cells, usually *E. coli*, and individual clones are isolated. Then, the mutated vector-gene recombinant DNA is isolated from a number of clones and the sequence of the mutant gene(s) is(are) determined. Finally, a characterized, mutant gene is used to transform host cells in which the gene is expressed (transcribed and translated). In many cases the host cell will also perform post-translational modification of the initial polypeptide product to achieve a mature protein. Characterization of the mutant protein can involve definition of its amino-acid composition and sequence, and its enzymatic characteristics, as well as biophysical studies on its absorption, optical, nuclear magnetic and electron spin resonance spectra, definition of its stability to thermal and chemical denaturation, its thermodynamic properties (by microcalorimetry) and its crystallographic structure.

Methods that have been successfully used for site-directed mutagenesis of protein structural genes (or which have the potential to be useful in this kind of study) fall into two broad classes. These are the production of random point mutations within the target gene, which result in random single amino-acid substitutions, and the production of specific, predetermined mutations designed to produce a specific amino-acid substitution.

Random point mutations

The objective in producing random point mutations in the structural gene of a protein is to define, without prejudice, the role(s) of all, or most of, the constituent amino acids in the

structure and function of the target protein. The protein to which this principle has been applied most extensively is the endonuclease from *Staphylococcus aureus*, the gene for which can be expressed in *E. coli* and where the activity of the product nuclease can be determined by a simple plate assay (Shortle 1983). Over 150 unique mutations in the nuclease gene have been generated and defined by DNA sequence determination. As well as creating mutants which clearly relate to the function and the calcium-binding site of the enzyme, the strategy has discovered a new class of amino-acid replacement, which acts as a global suppressor of second-site mutations (Shortle & Lin 1985).

Two features of methods for production of random point mutations have to be considered. These are the procedures, chemical or enzymatic, used to obtain the mutation and the strategy for optimizing the production of single hits (to avoid or minimize the generation of multiple mutations while producing single base-pair substitutions at a sufficiently high frequency that screening by DNA sequence determination is feasible).

Chemical methods for producing point mutations can be divided into two classes: those which modify bases so that their hydrogen-bond pairing with a complementary base is changed, and those which so damage a base that it can no longer form hydrogen bonds with a complementary base. The classical reagent for changing the base-pairing specificity of a target nucleotide is sodium bisulphite. Reaction with DNA causes the hydrolytic deamination of deoxycytidine residues with the formation of deoxyuridine, which pairs with deoxyadenosine. *In vivo* replication of bisulphite-treated DNA results in CG to TA mutations (Shortle & Botstein 1983; Shortle & Nathans 1978). Bisulphite has the additional important property of being a reagent that reacts, perferentially, with single-strand DNA; this is an aid to targeting. Other chemical reagents which react preferentially, but not exclusively, with deoxycytidine residues, and which have been used to produce random point mutations are hydrazine (Myers *et al.* 1985*c*; hydrazine reacts equally well with deoxythymidine residues), hydroxylamine (Busby *et al.* 1982), methoxylamine (Borrias *et al.* 1976; Kadonaga & Knowles, 1985) and nitrous acid (Hirose *et al.* 1982; Myers *et al.* 1985*c*; Warburton *et al.* 1983). None of these reagents shows a notable preference for single-strand DNA.

Chemical reagents that damage bases in DNA so that they can no longer form a Watson–Crick base-pair are typified by acids which induce depurination (Gilbert 1981; Kunkel *et al.* 1981*b*; Loeb 1985; Myers *et al.* 1985*c*). Their use to produce random point mutations involves the treatment of single-stranded DNA with the reagent, followed by *in vitro* enzymatic synthesis of the complementary strand by using a DNA polymerase (such as avian retrovirus reverse transcriptase) that misincorporates a nucleotide opposite the damaged or missing base in the template DNA strand (Kunkel *et al.* 1981*a*; Loeb 1985; Myers *et al.* 1985*c*).

Totally enzymatic methods for producing random point mutations use a DNA polymerase to insert a mutagenic nucleoside analogue in a random way in a DNA strand produced *in vitro* by template-dependent synthesis, or they deliberately induce misincorporation of normal nucleotides into an *in vitro* synthesized DNA strand. The nucleoside analogue that has been used most extensively as a substrate for *in vitro* mutagenesis is N^6-hydroxydeoxycytidine (Mueller *et al.* 1978; Weissmann *et al.* 1979). Its appeal lies in the fact that it substitutes with almost equal efficiency for both deoxycytidine and deoxythymidine. However, in normal host strains the analogue is a good substrate for *in vivo* mismatch (or damage) repair and an additional *in vitro* step is required for efficient mutagenesis (Dierks *et al.* 1983; Wieringa *et al.* 1983). Other nucleoside analogues, which have been used as site-specific mutagens after *in vitro*

incorporation into DNA as deoxynucleoside-5′-triphosphates, are O^6-methyldeoxyguanosine (Dodson *et al.* 1982), 2-aminopurinedeoxyriboside (Grossberger & Clough 1981), and also 5-bromodeoxyuridine (Mott *et al.* 1984). None of these analogues is as efficient in producing mutants as N^6-hydroxydeoxycytidine and a phenotypic screen is essential for their use.

Enzymatic misincorporation of nucleotides involves the use of strategies which subvert enzymatic mechanisms for editing misincorporation. The first type of strategy creates randomly distributed short single-strand gaps in duplex DNA where the single-strand segment is only three or four nucleotides in length. Here, the gap is filled enzymatically by using a DNA polymerase with only three of the four deoxyribonucleoside-5′-triphosphates as substrates. Misincorporation occurs when the template demands the insertion of the missing nucleotide and the reaction is driven by enzymatic ligation of the newly synthesized gap-filling sequence to the adjacent segment of the same DNA strand (Shortle *et al.* 1982). The second strategy for misincorporation of normal nucleotides involves two enzymatic synthetic steps. In the first step a DNA primer, annealed to the target template strand, is extended by using only one nucleotide as substrate (four different reactions can take place in parallel), in conjunction with a non-editing DNA polymerase, such as a reverse transcriptase. The primer is then extended with all four nucleotides used as substrates (Traboni *et al.* 1983, 1984; Zakour & Loeb 1982; Zakour *et al.* 1984).

Nucleoside-5′-α-thiotriphosphates are substrates for DNA polymerases; however, the incorporated nucleoside-5′-α-thiophosphate is not removed by the editing 3′-exonuclease activity of *E. coli* DNA polymerase I (Kunkel *et al.* 1981). This property makes these nucleotide analogues ideal substrates for the first, misincorporation, step of the two-step strategy outlined above (Shortle *et al.* 1982) and their use has been very effective in obtaining random point mutants (Shortle & Lin 1985).

As indicated earlier in this section, the efficient achievement of single-hit mutation is an essential component of site-directed mutagenesis. The chemical damage methods have used two procedures to maximize single hits. When the reagent is single-strand specific, such as sodium bisulphite, the target single strand can be a very short gap in an otherwise duplex molecule (Shortle & Botstein 1983; Shortle & Nathans 1978). When the strategy involves a long single-strand or double-strand region as a target (there are many such strategies: see Smith 1985), single hits result from control of the chemical reaction (Myers *et al.* 1985*c*; Shortle & Botstein 1983). The problem with limiting a reaction to no more than a single hit on a relatively long DNA target is that 90% or more of the target DNA molecules are not damaged. Two procedures have been devised for isolating mutants at high efficiency under these circumstances. The first of these involves treatment of each of the separated DNA strands of the target with the reagent under one-hit conditions, after which the two mutagenized DNA strands are annealed and cloned (Kadonaga & Knowles 1985). The second procedure takes advantage of the exquisite sensitivity of the initiation of denaturation of double-strand DNA fragments to local DNA sequences and of the fact that partly denatured double-strand DNA migrates very slowly during gel electrophoresis (Myers *et al.* 1985*a*, *b*). These properties allow the separation of mutant duplex DNA fragments, up to several hundred nucleotides in length, from the parent wild-type fragment (Myers *et al.* 1985*c*).

In chemical mutagenesis, where the reagent does not select a target region, such as a single-strand gap in a duplex molecule, mutations (although the reaction may be base-specific) will be distributed randomly throughout the target region. When the target is a single-strand

gap, it is necessary to ensure that the gaps (one per target molecule) is distributed randomly. This is achieved by use of superhelical double-strand circular plasmid DNA as target for a non-specific endonuclease in the presence of ethidium bromide. This limits the exonuclease to one hit per molecule (Greenfeld *et al.* 1975), and the resultant nick is converted to a short gap by limited exonuclease treatment (Shortle & Botstein 1983).

For nucleotide misincorporation, the reaction tends to misincorporate only one nucleotide (Shortle *et al.* 1982; Shortle & Lin 1985). Hence misincorporation by *in vitro* enzyme-catalysed synthesis is intrinsically efficient in producing single hits. A variety of approaches have been developed for obtaining a completely random set of priming sites throughout the target DNA. These include the use of randomly distributed short gaps (Shortle & Botstein 1983; Shortle & Lin 1985), random priming sites produced by exonuclease digestion from one end of the target (Abarzua & Marians 1984) and random priming sites generated by enzymatic extension of a single primer fragment (Brown & Smith 1977; Traboni *et al.* 1983; Zakour *et al.* 1984; A. Spence & M. Smith, unpublished results).

The use of synthetic oligodeoxyribonucleotides as mutagens does not appear to be an attractive approach to random mutant production. However, synthetic oligonucleotides have been used in this way to a limited extent and it is likely that this use will increase. There are several reasons for this. First, methods for the automated synthesis of oligonucleotides are becoming more widely available and improvements in the synthetic chemistry are making possible the synthesis of very long oligodeoxyribonucleotides (Sinha *et al.* 1984). Secondly, the segmented-disc–solid-support method allows the parallel synthesis of very many oligodeoxyribonucleotides (Frank *et al.* 1983; Ott & Eckstein 1984). Thirdly, very efficient procedures are now available for producing mutants by using oligonucleotides and, lastly, oligonucleotide mutagenesis is the only method available for deleting or inserting specific codons in the structural gene for a protein or for making codon substitutions which involve more than a single base-pair change.

The prime example of the use of synthetic oligonucleotides to produce a considerable number of replacement or deletion mutants by oligonucleotide mutagenesis of a structural gene is the study of the signal peptide of the outer membrane lipoprotein of *E. coli* (Inouye *et al.* 1982). The basic strategy in these experiments is that which is used for producing a single defined mutant. Two other strategies, specifically designed to produce random mutants, use synthetic oligonucleotides. The first of these uses a long (50 nucleotides or more) synthetic strand in which a low random or semi-random level of the three nucleotides, other than the correct nucleotide, is incorporated at each position, as well as the correct nucleotide. This oligonucleotide is incorporated into the target DNA and clones of mutant DNAs are isolated after transformation of the host-cell line (J. B. McNeil & M. Smith, unpublished results). The method has two attractive features. First, the target is precisely defined by the ends of the synthetic oligonucleotide. Secondly, the misincorporations can be chosen so that all the codon changes generate amino-acid substitutions with no nonsense mutations. The second strategy, where a single oligodeoxyribonucleotide is used to produce random mutations, is the structural gene for a protein involved in frame insertions of a short symmetrical oligonucleotide (containing $3n$ nucleotides: see Stone *et al.* 1984). This is a very useful technique for identifying or disrupting functional domains.

Oligonucleotide heteroduplex mutagenesis

The principle of this method is to use heteroduplex DNA, formed by annealing a vector containing the target gene as a single strand with a complementary synthetic oligodeoxyribonucleotide containing the desired codon change (which, usually, is extended and integrated into duplex DNA by using DNA polymerase and ligase), to transform a cell line in which the vector can replicate (Hutchison *et al.* 1978; Razin *et al.* 1978). The progeny recombinant DNA contains either the wild-type gene or the gene with the mutation specified by the synthetic oligodeoxyribonucleotide (Smith & Gillam 1981; Smith 1985). A considerable amount of work has been done to make this basic procedure as technically straightforward as possible (Zoller & Smith 1984; Inouye & Inouye 1985), and to improve the efficiency with which the heteroduplex DNA produces progeny from the mutant-containing strand. Apart from care with the enzymatic manipulations which are used to lengthen the mutagenic oligonucleotide before transformation (Gillam & Smith 1979*a*; Kunkel 1985; Zoller & Smith 1983), procedures which improve yields of desired mutants include genotypic selection with the use of the mutagenic oligonucleotide (Gillam & Smith 1979*b*), prevention of *in vivo* DNA mismatch repair that is biased towards the sequence of the template (Kramer *et al.* 1982; Marmenout *et al.* 1984), by using mismatch repair of the template strand to convert it to the mutant sequence (Kunkel 1985), *in vitro* enzymatic synthesis of mutant recombinant DNA (Abarzua & Marians 1984) and covalent linkage of the mutagenic oligonucleotide to a selectable genetic marker (Carter *et al.* 1985; Kramer *et al.* 1984). Use of these principles has resulted in the generation of a very large number of mutant proteins for structure–function studies (Smith 1985) and in the production of new proteins with useful biochemical properties (Courtenay *et al.* 1985; Craik *et al.* 1985; Perry & Wetzel 1984; Rosenberg *et al.* 1984; Wang *et al.* 1984; Wilkinson *et al.* 1984). An impressive application has been the production of a large number of mutants of tyrosyl-tRNA synthetase (Fersht *et al.* 1984).

One of the most attractive features of oligonucleotide heteroduplex mutagenesis is the use of the mutagenic oligonucleotide as a probe for mutant DNA recognition. This procedure takes advantage of the very substantial decrease in duplex stability for a mismatched oligonucleotide relative to the stability of a perfectly matched oligonucleotide (Gillam *et al.* 1975; Smith 1983). It has been applied to nitrocellulose dot-blots of recombinant phage DNA (Zoller & Smith 1982, 1983, 1984), to plaques (Zoller & Smith 1984), and to colonies containing recombinant plasmids (Wallace *et al.* 1981).

Mutagenesis by partial (cassette) or total gene synthesis

A major factor in choosing a mutagenic strategy for producing a desired change or set of changes in the gene encoding a protein of interest is the efficiency with which the mutant(s) is(are) produced. Modification in the heteroduplex procedures outlined above, and analogous procedures for mutant enrichment which have been applied to random *in vitro* mutagenesis (Abarzua & Marians 1984; Kunkel 1985; Myers *et al.* 1985; Traboni *et al.* 1983), have greatly increased efficiencies and reduced the labour of genotypic screening. However, a mutagenic procedure which is 100% efficient has a great appeal. The most obvious approach to this objective is the synthesis of both strands of the desired mutant by enzymatic ligation of chemically synthesized oligodeoxyribonucleotides. A simple application of this principle is

possible if the target segment of the gene lies between two closely spaced, unique, restriction endonuclease cleavage sites. The intervening segment is excised and replaced by a totally chemically synthesized mutant duplex (Liu *et al.* 1984; Smith 1985). Oligonucleotides which contain a mixture of substitutions at a particular codon can be used in this procedure to generate a large family of amino-acid substitutions (Wells *et al.* 1985). The ready availability of synthetic oligodeoxyribonucleotides of lengths of more than 20 nucleotides has made total gene synthesis, formerly a formidable task (Khorana 1979), an attractive option for mutant production (Edge *et al.* 1981). The attraction of partial or total gene synthesis for genes that are to be expressed in a heterologous host is that codons can be used to specify the amino acids which are most generally used in the new host-cell line (Edge *et al.* 1981; Buell *et al.* 1985). In addition, the fact that the chemical synthesis yields relatively large amounts of the oligonucleotides allows the repeated use of subsets as building blocks for different mutants of the same gene (Lo *et al.* 1984).

Conclusions

This article has described the three basic strategies that are presently available, and in use, for the systematic *in vitro* mutagenesis of protein structural genes. These are: the production of random base-pair substitutions in the target gene by chemical or enzymatic means; the production of a mutant programmed by a chemically synthesized oligodeoxyribonucleotide annealed as a heteroduplex with the wild-type target DNA; and the partial or total chemical synthesis of both strands of a mutant gene. All three approaches are exceedingly powerful in their ability to produce mutant genes. For this reason, it is not clear that one method is superior to another. In fact, it is quite possible that studies directed at producing mutants in a particular protein will have occasion to use more than one or all of the approaches. What is impressive is that an experimental approach that was first described less than ten years ago has developed so rapidly that it is now essential to the work of a number of laboratories. Even more surprising is the rapidity with which oligodeoxyribonucleotide synthesis has become one of the most efficient and rapid of chemistries, with the consequence that molecular genetics cannot be performed in the most efficient way without the routine use of synthetic oligodeoxyribonucleotides. It is clear that the combination of site-directed mutagenic methods described in this article has opened up a completely new era in protein and enzyme chemistry.

Research in the author's laboratory was supported by the Medical Research Council and the National Institute of Canada. The author is a Career Investigator of the Medical Research Council of Canada.

References

Abarzua, P. & Marians, K. J. 1984 Enzymatic techniques for the isolation of random single-base substitutions *in vitro* at high frequency. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2030–2034.

Borrias, W. E., Wilschut, I. J. C., Vereijken, J. M., Weisbeek, P. J. & van Arkel, G. A. 1976 Induction and isolation of mutants in a specific region of gene A of bacteriophage ϕ174. *Virology* **70**, 195–197.

Brake, A. J., Merryweather, J. P., Coit, D. G., Heberlen, V. A., Masiarz, F. R., Mullenbach, G. T., Urdea, M. S., Valenzuela, P. & Barr, P. J. 1984 α-Factor-directed synthesis and secretion of mature foreign proteins in *Saccharomyces cerevisiae*. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4642–4646.

Brown, N. L. & Smith, M. 1977 The sequence of a region of bacteriophage ϕX174 DNA coding for parts of genes A and B. *J. molec. Biol.* **116**, 1–28.

Buell, G., Schulz, M.-F., Selzer, G., Chollet, A., Movva, N. R., Semon, D., Escanez, S. & Kawashima, E. 1985 Optimizing the expression by *E. coli* of a synthetic gene encoding somatamedin-C (IGF-I). *Nucl. Acids Res.* **13**, 1923–1938.

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

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

Courtenay, M., Jallat, S., Tessier, L.-H., Benavente, A., Crystall, R. G. & Lecocq, J. P. 1985 Synthesis in *E. coli* of α_1-antitrypsin varients of therapeutic potential for emphysema and thrombosis. *Nature, Lond.* **313**, 149–151.

Craik, C. S., Largman, C., Fletcher, T., Rocznaik, S., Barr, P. J., Fletterick, R. & Rutter, W. J. 1985 Redesigning trypsin: alteration of substrate specificity. *Science, Wash.* **228**, 291–297.

Dierks, P., van Ooyen, A., Cochran, J. D., Dobkin, C., Reiser, J. & Weissman, C. 1983 Three regions upstream from the cap site are required for efficient and accurate transcription of the rabbit β-globin gene in mouse 3T6 cells. *Cell* **22**, 697–706.

Dodson, L. A., Foote, R. S., Mitra, S. & Maskier, W. E. 1982 Mutagenesis of bacteriophage T7 *in vitro* by incorporation of O^6-methylguanine during DNA synthesis. *Proc. natn. Acad. Sci. U.S.A.* **79**, 7440–7444.

Edge, N. D., Greene, A. R., Heathcliffe, G. R., Meacock, P. A., Schuch, W., Scanlon, D. B., Atkinson, T. C., Newton, C. R. & Markham, A. F. 1981 Total synthesis of a human leukcyte interferon gene. *Nature, Lond.* **292**, 756–762.

Ferguson-Miller, S., Brautigan, D. L. & Margoliash, E. 1979 The electron transfer function of cytochrome *c*. In *The porphyrins* (ed. D. Dolphin), vol. 7, pp. 149–239. New York: Academic Press.

Fersht, A. R., Shi, J.-P., Wilkinson, A. J., Blow, D. M., Carter, P., Waye, M. M. Y. & Winter, G. P. 1984 Analysis of enzyme structure and activity by protein engineering. *Angew. Chem. int. Edn Engl.* **23**, 467–473.

Frank, R., Heikens, W., Heisberberg-Moutsis, G. & Blocker, H. 1983 A new general approach for the simultaneous chemical synthesis of large numbers of oligonucleotides: segmented solid supports. *Nucl. Acids Res.* **11**, 4366–4377.

Gil, G., Faust, J. R., Chin, D. J., Goldstein, J. L. & Brown, M. S. 1985 Membrane-bound domain of HMG CoA reductase is required for sterol enhanced degradation of the enzyme. *Cell* **41**, 249–258.

Gilbert, W. 1981 DNA sequencing and gene structure. *Science, Wash.* **214**, 1305–1312.

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

Gillam, S. & Smith, M. 1979*b* Site-specific mutagenesis using synthetic oligodeoxyribonucleotide primers. II. *In vitro* selection of mutant DNA. *Gene* **8**, 99–106.

Gillam, S., Waterman, K. & Smith, M. 1975 The base-pairing specificity of cellulose-pT_9. *Nucl. Acids Res.* **2**, 624–634.

Greenfeld, D., Simpson, L. & Kaplan, D. 1975 Conversion of closed circular DNA melecules to single nicked molecules by digestion with DNaseI in the presence of ethidium bromide. *Biochim. biophys. Acta* **407**, 365–375.

Grossberger, D. & Clough, W. 1981 Incorporation into DNA of the base analog 2-aminopurine by Epstein–Barr virus-induced DNA polymerase *in vivo* and *in vitro*. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7271–7275.

Hirose, S., Takeuchi, K. & Suzuki, Y. 1982 *In vitro* characterization of the fibroin gene promoter by the use of single base substitution mutants. *Proc. natn. Sci. U.S.A.* **79**, 7258–7262.

Hutchison, C. A., III, Phillips, S., Edgell, M. H., Gillam, S., Jahnke, P. & Smith, M. 1978 Mutagenesis at a specific position in a DNA sequence. *J. biol. Chem.* **253**, 6551–6560.

Inouye, S. & Inouye, M. 1985 Oligonucleotide-directed site-specific mutagenesis using double-stranded plasmid DNA. In *DNA and RNA synthesis* (ed. S. Narang). New York: Academic Press.

Inouye, S., Soberon, X., Franceschini, T., Nakamura, K., Itakura, K. & Inouye, M. 1982 Role of positive charge on the amino-terminal region of the signal peptide in protein secretion across the membrane. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3438–3641.

Kadonaga, J. Y. & Knowles, J. R. 1985 A simple and efficient method for chemical mutagenesis of DNA. *Nucl. Acids Res.* **13**, 1733–1745.

Khorana, H. G. 1979 Total synthesis of a gene. *Science, Wash.* **203**, 614–625.

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

Kramer, W., Schughart, K. & 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.

Kunkel, T. A. 1985 Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. natn. Acad. Sci. U.S.A.* **82**, 488–492.

Kunkel, T. A., Eckstein, F., Mildvan, A. S., Koplitz, R. M. & Loeb, L. A. 1981*a* Deoxyribonucleoside [1-thio]triphosphates prevent proofreading during *in vitro* DNA synthesis. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6734–6738.

Kunkel, T. A., Shearman, C. W. & Loeb, L. A. 1981*b* Mutagenesis *in vitro* by depurination of φX174 DNA. *Nature, Lond.* **291**, 349–351.

Liu, C.-C., Simonsen, C. C., & Levison, A. D. 1984 Initiation of translation at internal AUG codons in mammalian cells. *Nature, Lond.* **309**, 82–85.

Lo, K.-M., Jones, S. S., Hackett, N. R. & Khorana, H. G. 1984 Specific amino acid substitutions in bacterioopsin: replacement of a restriction fragment in the structural gene by synthetic DNA fragments containing altered codons. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2285–2289.

Loeb, L. A. 1985 Apurinic sites as mutagenic intermediates. *Cell* **40**, 483–484.

Marmenout, A., Remaut, E., van Boom, J. & Fiers, W. 1984 Oligonucleotide-directed mutagenesis: selection of mutants by hemimethylation of GATC sequence. *Molec. gen. Genet.* **196**, 126–133.

McKusick, V. A. 1983 *Mendelian inheritance in man* (6th edn), pp. 556–562. Baltimore: Johns Hopkins University Press.

Miller, J. H. 1978 The *lacI* gene: its role in lac operon control and its use as a genetic system. In *The operon* (ed. J. H. Miller & W. S. Reznikoff), pp. 31–88. Cold Spring Harbor: Cold Spring Laboratory.

Miller, J. H. 1984 Genetic studies of the *lac* repressor. XII. Amino acid replacements in the DNA binding domain of the *Escherichia coli lac* repressor. *J. molec. Biol.* **1980**, 205–212.

Mott, J. E., van Arsdell, J. & Platt, T. 1984 Targeted mutagenesis *in vitro*: *lac* repressor mutations generated using AMV transcriptase and dBrUTP. *Nucl. Acids Res.* **12**, 4139–4152.

Mueller, W., Weber, H., Meyer, F. & 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. molec. Biol.* **124**, 343–358.

Myers, R. M., Fischer, S. G., Lerman, L. S. & Maniatis, T. 1985*a* Nearly all single base substitutions in DNA fragments joined in GC-clamp can be detected by denaturing gradient gel electrophoresis. *Nucl. Acids Res.* **13**, 3131–3145.

Myers, R. M., Fischer, S. G., Maniatis, T. & Lerman, L. S. 1985*b* Modification of the melting properties of duplex DNA by attachment of a GC-rich DNA sequence as determined by denaturing gel electrophoresis. *Nucl. Acids Res.* **13**, 3147–3159.

Myers, R. M., Lerman, L. S. & Maniatis, T. 1985*c* A general method for saturation mutagenesis of cloned DNA fragments. *Science, Wash.* **229**, 242–247.

Ott, J. & Eckstein, F. 1984 Filter disc supported oligonucleotide synthesis by the phosphite method. *Nucl. Acids Res.* **12**, 9137–9142.

Perry, J. L. & Wetzel, R. 1984 Disulphide bond engineered into T4 lysozyme: stabilization of the protein towards thermal inactivation. *Science, Wash.* **226**, 555–557.

Perutz, M. F. 1979 Regulation of oxygen affinity of hemoglobin: influence of structure of the globin on the heme iron. *A. Rev. Biochem.* **48**, 327–386.

Perutz, M. F. 1983 Species adaption in a protein molecule. *Molec. Biol. Evol.* **1**, 1–28.

Pielak, G. J., Mauk, A. G. & Smith, M. 1985 Site-directed mutagenesis of cytochrome c shows that an invarient Phe is not essential for function. *Nature, Lond.* **313**, 152–154.

Razin, A., Hirose, T., Itakura, K. & Riggs, A. D. 1978 Efficient correction of a mutation by use of chemically synthesized DNA. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4268–4270.

Rosenberg, S., Barr, P. J., Najarian, R. C. & Hallewell, R. A. 1984 Synthesis in yeast of functional oxidation-resistant mutant of human α_1-antitrypsin. *Nature, Lond.* **312**, 77–80.

Shortle, D. 1983 A genetic system for analysis of staphylococcal nuclease. *Gene* **32**, 181–189.

Shortle, D. & Botstein, D. 1983 Directed mutagenesis with sodium bisulfite. *Meth. Enzymol.* **100**, 457–468.

Shortle, D., Di Maio, D. & Nathans, D. 1981 Directed mutagenesis. *A. Rev. Genet.* **15**, 265–294.

Shortle, D., Grisafi, P., Benkovic, S. & Botstein, D. 1982 Gap misrepair mutagenesis: efficient site-directed induction of transition, transversion and frameshift mutations *in vitro*. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1588–1592.

Shortle, D. & Lin, B. 1985 Genetic analysis of *Staphylococcus* nuclease: identification of three intragenic 'global' suppressors of nuclease-minus mutations. *Genetics* **110**, 539–555.

Shortle, D. & Nathans, 1978 Local mutagenesis: a method for generating viral mutants with base substitutions in preselected regions of the viral genomes. *Proc natn. Acad. Sci. U.S.A.* **75**, 2170–2174.

Sinha, N. D., Biernat, J., McManus, J. & Koster, H. 1984 Polymer support oligonucleotide synthesis. XVII. use of β-cyanoethyl-*N*,*N*-dialkyl amino/*N*-morpholino phosphoramidite of deoxynucleosides for the synthesis of DNA fragments simplifying deprotection and isolation of the final products. *Nucl. Acids Res.* **11**, 4539–4557.

Smith, M. 1983 Synthetic oligodeoxyribonucleotides as probes for nucleic acids and as primers in sequence determination. In *Methods of DNA and RNA sequencing* (ed. S. M. Weissman), pp. 23–68. New York: Praeger.

Smith, M. 1985 *In vitro* mutagenesis. *A. Rev. Genet.* **19**, 423–462.

Smith, M. & Gillam, S. 1981 Constructed mutants using synthetic oligodeoxyribonucleotides as site-specific mutagens. In *Genetic engineering. Principles and methods* (ed. J. K. Setlow & A. Hollaender), vol. 3, pp. 1–32. London & New York: Plenum Press.

Stone, J. C., Atkinson, T., Smith M. & Pawson, R. 1984 Identification of functional regions in the transforming protein of Fujinami sarcoma virus by in-phase insertion mutagenesis. *Cell* **37**, 549–558.

Takano, T. & Dickerson, R. E. 1981*a* Conformation change of cytochrome *c*. I. Ferocytochrome *c* structure refined at 1.5 Å resolution. *J. molec. Biol.* **153**, 79–94.

Takano, T. & Dickerson, R. E. 1981*b* Conformation change of cytochrome *c*. II. Ferricytochrome *c* refinement at 1.8 Å and comparison with the ferrocytochrome *c* structure. *J. Molec. Biol.* **153**, 95–115.

Traboni, C., Cortese, T., Ciliberto, G. & Cesareni, G. 1983 A general method to select for M13 clones carrying base pair substitution mutants constructed *in vitro*. *Nucl. Acids Res.* **11**, 4229–4239.

Traboni, C., Cilberto, G. & Cortese, R. 1984 Mutations in box B of the promoter of a eucaryote $tRNA^{Pro}$ gene affecting rate of transcription, processing and stability of transcripts. *Cell* **30**, 179–187.

Villafranca, J. E., Howell, E. E., Boet, D. H., Strobel, M. S., Ogden, R. C., Abelson, J. N. & Kraut, J. 1983 Direct mutagenesis of dihydrofolate reductase. *Science, Wash.* **222**, 782–788.

Wallace, R. B., Schold, M., Johnson, M. J., Dembek, P. & Itakura, K. 1981 Oligonucleotide-directed mutagenesis of the human globin gene: a general method for producing specific point mutations. *Nucl. Acids Res.* **9**, 3647–3656.

Wang, A., Lu, S.-D. & Mark, D. F. 1984 Site-specific mutagenesis of the human interluekin-2 gene: structure function analysis of the cysteine residues. *Science, Wash.* **224**, 1431–1433.

Warburton, N., Bosely, P. G. & Porter, A. G. 1983 Increased expression of a cloned gene by local mutagenesis of its promoter and ribosome binding site. *Nucl. Acids Res.* **11**, 5837–5854.

Watson, J. D., Tooze, J. & Kurtz, D. T. 1983 *Recombinant DNA. A short course*, pp. 106–116. New York: Scientific American Books.

Weinmaster, G., Zoller, M. J., Smith, M., Hinze, D. & Pawson, T. 1984 Mutagenesis of Fujinami sarcoma virus: evidence that tyrosine phosphorylation of $P130^{gag\text{-}fps}$ modulates its biological activity. *Cell* **37**, 559–568.

Weissmann, C., Nagata, S., Taniguchi, T., Wever, H. & Meyer, F. 1979 The use of site-directed mutagenesis in reverse genetics. In *Genetic engineering: Principles and methods* (ed. J. K. Setlow & A. Hollaender), vol. 1, pp. 133–150. New York & London: Plenum Press.

Wells, J. A., Vasser, M. & Powers, D. B. 1985 Cassette mutagenesis: an efficient method for generation of multiple mutations at defined sites. *Gene* **34**, 315–323.

Wieringa, B., Meyer, R., Reiser, J. & Weissmann, C. 1983 Unusual splice sites revealed by mutagenic inactivation of an authentic splice site of the rabbit β-globin gene. *Nature, Lond.* **301**, 38–43.

Winter, G. & Fersht, A. R. 1984 Engineering enzymes. *Trends biotechnol.* **2**, 115–119.

Zakour, R. A., James, E. A. & 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.

Zoller, M. J. & Smith, M. 1982 Oligonucleotide-directed mutagenesis using M13-derived vectors: an efficient and general procedure for production of point mutations in any fragment of DNA. *Nucl. Acids Res.* **10**, 6487–6500.

Zoller, M. J. & Smith, M. 1983 Oligonucleotide-directed mutagenesis of DNA fragments cloned in M13-derived vectors. *Meth. Enzymol.* **100**, 468–500.

Zoller, M. J. & Smith, M. 1984 Oligonucleotide-directed mutagenesis: a simple method using two oligonucleotide primers and a single-stranded DNA template. *DNA* **3**, 479–488.