SYNTHETIC DNA AND BIOLOGY

Nobel Lecture, December 8, 1993

by

MICHAEL SMITH

Biotechnology Laboratory and Department of Biochemistry, University of British Columbia, Vancouver, B.C. V6T lW5, Canada

> *"Most of the significant work has been summarized in a number of reviews and articles. In these there was, of necessity, a good deal of simplification and omission of detail. . . With the passage of time, even I find myself accepting such simplified accounts. "*
>
> F. Sanger (1988)

INTRODUCTION

I had the good fortune to arrive in the laboratory of Gobind Khorana, in September 1956, just one month after he had made the accidental discovery of the phosphodiester method for the chemical synthesis of deoxyribo-oligonucleotides (Khorana *et al* ., 1956), a synthetic approach whose full exploitation led to elucidation of the genetic code and the first total synthesis of a gene (Khorana, 1961, 1969, 1979). An identifying characteristic of the Khorana approach to the production of synthetic polynucleotides for the solution of biological problems has been a willingness to use both chemical and enzymatic tools as appropriate (Chambers *et al* ., 1957; Khorana, 1979) and this catholic approach is the foundation of all the molecular technology employed in modern molecular genetics (Ausubel *et al.,* 1987; Sambrook *et al. ,* 1989; Watson *et al. ,* 1993). It had a major impact on my thinking.

During my post-doctoral time in the Khorana group, I was mainly involved in small molecule synthesis; nucleoside-5' triphosphates (Smith and Khorana, 1958) and nucleoside-3',5' cyclic phosphates (Smith *et al.,* 1961). However, I was able to contribute to the technology of polynucleotide synthesis by the development of the methoxytrityl series of 5'-hydoxyl protecting groups (Smith *et al. ,* 1961) and with a first approach to the chemical synthesis of ribo-oligonucleotides (Smith and Khorana, 1959).

In 1961, I left the Khorana group to join the Fisheries Research Board of Canada Vancouver Laboratory. Whilst much of my research effort was directed at studies on salmonid physiology and endocrinology, it was possible to continue studies on the chemistry of phosphodiester synthesis and

this resulted in a new method for nucleoside-3'5' cyclic phosphate synthesis by alkaline transesterification under anhydrous conditions, a reaction which obviated the need for chemical protecting groups on the heterocyclic bases (Smith, 1964; Borden and Smith, 1966).

In 1966, I became a faculty member in the Department of Biochemistry of the University of British Columbia, which has been my academic home ever since. Studies on oligonucleotide synthesis continued, being directed at the reaction of deoxyribonucleoside phosphorofluoridates in anhydrous alkaline conditions (von Tigerstrom and Smith, 1970). Although this method, in obviating the need for base protecting groups, represented a significant advance, it did not have the generality required of a universal synthetic method (von Tigerstrom *et al.,* 1975a, 1975b).

An approach to deoxyribo-oligonucleotide synthesis which did prove to be versatile, simple and useful, although inefficient, involved the extension of short primers using *E. coli* polynucleotide phosphorylase, in the presence of Mn^{+2} and NaCl, with doxyribonucleosise-5' diphosphates as substrate (Gillam and Smith, 1972, 1974, 1980; Gillam *et al.,* 1978). This procedure, which made possible the synthesis of deoxyribo-oligonucleotides up to 12 to 13 nucleotides in length, proved to be a significant breakthrough for our small group because it allowed us to undertake a number of fairly ambitious molecular biological projects at a time, between 1970 and 1980, when oligonucleotides were not generally accessible. Of course, it has been displaced as a routine synthetic procedure with the advent of automated chemical synthesis (which still uses the methoxytrityl protecting group) with nucleoside-3' phosphoroamidites as the key intermediates (Adams *et al.,* 1983; Atkinson and Smith, 1984; McBride and Caruthers, 1983).

MODEL STUDIES ON OLIGONUCLEOTIDE DOUBLE HELIX STABILITY AND SPECIFICITY

In 1968, planned studies on the biosynthesis of salmonid protamine (Ingles *et al.,* 1966) and on the sequence of dA,dT-rich crab DNA (Astell *et al.,* 1969) became impossible for personal or for technological reasons. Casting around for a new project, I suggested to Caroline Astell, who was studying for her Ph.D., a series of model studies using chemically synthesized oligonucleotides directed at defining the stability and specificity of oligonucleotide duplexes of defined sequence. The objective was to see if a synthetically accessible oligonucleotide could be a tool for identifying and isolating a specific messenger RNA based on Watson-Crick hydrogen-bonding. From the statistical point of view it should be possible, since the length required of an oligonucleotide in order for it to be unique in a given genome size ranges from 9 nucleotides for phage lambda with a genome of 4.6×10^4 base pairs to 18 nucleotides for the higher plants, e.g. A. cepa with a genome of 1.5×10^{10} base pairs (Table 1). Also relevant was the work of Michelson and Monny (1967) and of Niyogi and Thomas (1968) on interactions of ribo-oligonucleotides with ribo-polynucleotides which indicated that stable dou-

Table 1. DNA content of haploid genomes and oligonucleotide length required for unique recognition.

Organism	Base pairs	Oligonucleotide length (N)*
Viruses		
Phage lambda	4.6×10^4	9
Bacteria		
D. pneumoniae	1.7×10^6	11
E. coli	4.1×10^6	12
Fungi		
S. cerevisiae	1.7×10^7	13
N. crassa	2.1×10^7	13
Plants		
C. reinhardtii	10^8	14
A. thaliana	10^8	14
Z. mays	6.6×10^9	17
A. cepa	1.5×10^{10}	18
Animals		
C. elegans	10^8	14
D. melanogaster	1.3×10^8	14
B. rerio	10^9	16
M. musculus	2.2×10^9	17
H. sapiens	3.3×10^9	17

* (N) is unique when $4^N \geqq 2 \times$ Base Pairs.

ble helices can be formed with as few as seven base pairs and, equally importantly, which demonstrated that there is a significant increase in duplex stability for each additional base pair up to about 16 base pairs (Figure 1). The work of Gilham with random length oligothymidylates covalently attached to cellulose demonstrated their potential power for affinity chromatography (Gilham, 1962; Gilham and Robinson, 1964) which led to an important procedure for isolation of eukaryote mRNAs (Aviv and Leder, 1972). However, although data was available on duplex stability involving *mixtures* of oligonucleotides of defined length with DNA (McConaughy and McCarthy, 1967; Niyogi, 1969), there was no systematic data obtained with pure oligonucleotides of defined sequence and length.

Our approach was to prepare deoxyribo-oligonucleotide-celluloses using simple oligonucleotides of defined length and sequence synthesized by the Khorana method and to use these oligonucleotide-celluloses, in the form of thermally-eluted columns, to establish the stability of duplexes with a variety of complementary, or partially complementary deoxyribo- and ribo-oligonucleotides (Astell *et al.,* 1973; Astell and Smith, 1971, 1972; Gillam *et al.,* 1975). The results and conclusions of this time-consuming but critical series
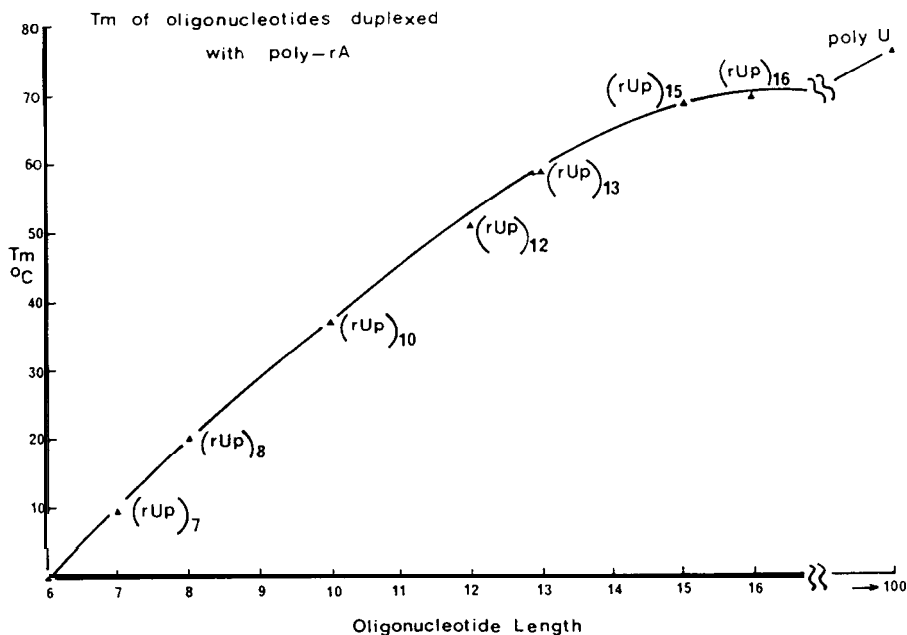
Tm of oligonucleotides duplexed
with poly—rA

poly U

$(rUp)_{16}$

$(rUp)_{15}$

$(rUp)_{13}$

$(rUp)_{12}$

$(rUp)_{10}$

$(rUp)_8$

$(rUp)_7$

Tm °C

Oligonucleotide Length

Fig. *1:* The relationship between oligouridylate length and the *Tm* of duplexes with polyadenylate in 0.02 M NaCl, 0.01 M sodium cacodylate, pH 7.0. Drawn from data of Michelson and Monny (1967).

of experiments has been summarized in more detail elsewhere (Smith, 1983) and only some highlights will be discussed here.

Firstly, it was clear that, under appropriate conditions of ionic strength, stable duplexes containing as few as six dA-dT base-pairs can be formed (Figure 2). Further, additional base-pairs increase duplex stability significantly, but in a non-linear manner with the incremental increase in stability diminishing with increased length. Also, a single base-pair mismatch of dT with dT or of dT with dG decreases duplex stability by the same amount (equivalent to removing about two base-pairs), but duplex formation is still possible at low temperature (Figure 2). In addition, it was obvious that oligonucleotide duplexes form rapidly at low temperatures, conditions under which denatured double strand DNA does not reanneal.

These experiments convinced me that it would be possible to use a synthetic oligonucleotide to identify an RNA or a DNA containing the exact Watson-Crick complement of its sequence and to differentiate it from similar but not identical sequences. An opportunity to demonstrate this principle was provided by mutants in the lysozyme-encoding locus of phage T4 (Doel and Smith, 1973). Also, at a later date, the data made me realize that it should be possible to use short synthetic deoxyribo-oligonucleotides as specific mutagens and to differentiate between a point mutant and wild-type DNA.

Experiments on the interaction of deoxyribo-oligonucleotide-celluloses and ribo-oligonucleotides confirmed earlier studies on the relative stabili-
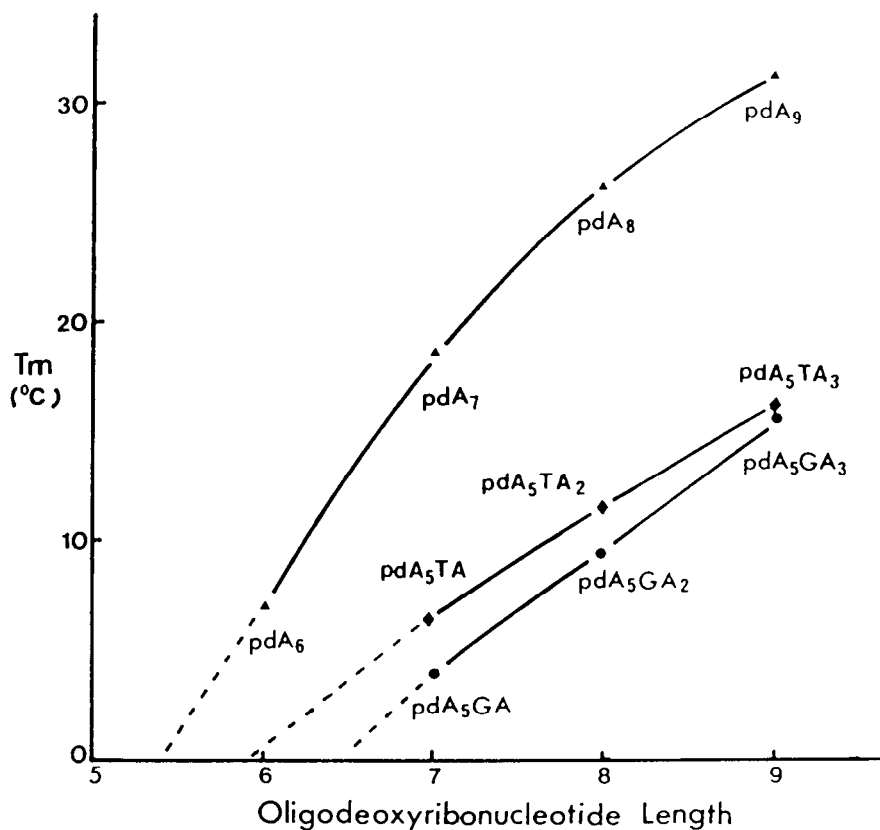
Fig. 2: The *Tm* values of the duplexes with cellulose-pdT$_9$ of complementary oligodeoxyribonucleotides containing a single base-pair mismatch in 1 M NaCl, 0.01 M sodium phosphate, pH 7.0. Drawn from data of Gillam, Waterman, and Smith (1975).

ties of DNA-DNA and DNA-RNA double strands (Chamberlin, 1965; Smith, 1983). They also demonstrated a further important principle: a dT, rG base pair is relatively stable (Figure 3). Differences in the stabilities of dT, dG and dT, rG base pairs (compare Figures 2 and 3), presumably, reflect differences in the base pairing allowed in the B-forms and the A-forms of double helix (Watson et al., 1987). The observation has two practical implications. First, the most specific and stringent interaction of an oligonucleotide probe will be the one between a *deoxyribo-oligonucleotide* and *DNA*. Second, one cannot assume that duplex structures that are formed by RNA will be formed by DNA or by deoxyribo-oligonucleotides. In connection with this, I have always been puzzled by the widespread acceptance of the idea that deoxyinosine is precocious in base pairing because inosine behaves in a precocious manner in t-RNA duplex interactions. Studies on deoxyinosine, in fact, indicate that it functions as a specific analog of deoxyguanosine, although it does not self-aggregate as does deoxyguanosine (Ausubel *et al.,* 1987; Kornberg, 1980; Sambrook *et al.,* 1989).
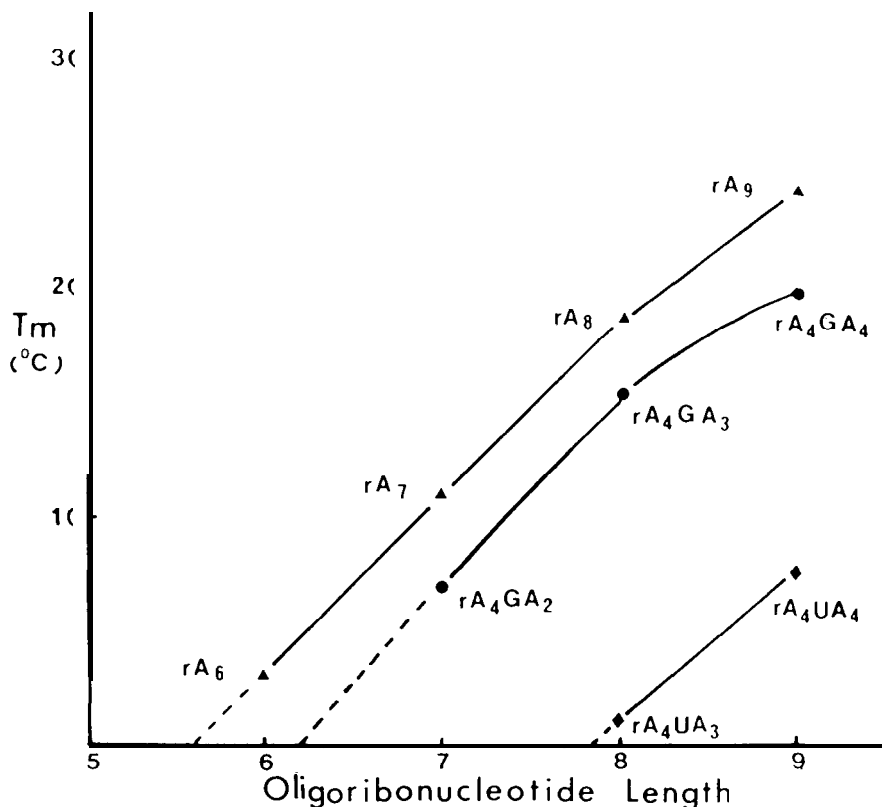
Fig. 3: The Tm values of the duplexes with cellulose-pdT, of complementary oligoribonucleo-
tides containing a single base-pair mismatch in 1 M NaCl, 0.01 M sodium phosphate, pH 7.0.
Drawn from data of Giliiam, Waterman, and Smith (1975).

The above empirical studies, and other related ones not discussed here
(Smith, 1983), provided the direct inspiration for and commitment to our
use of synthetic deoxyribo-oligonucleotides as probes for isolation of specif-
ic DNA fragments, as primers for the direct sequencing of double-stranded
DNA, as primers for precise definition of the ends of mRNAs, as specific
mutagens and as tools for the identification and isolation of point muta-
tions.

DEOXYRIBO-OLIGONUCLEOTIDES AS DNA PROBES

When the previously described model studies were commenced, in 1968,
the intent was to develop a method for isolation of specific m-RNAs by
affinity chromatography. However, the development of the full panoply of
DNA cloning technologies in the early 1970's (Watson *et al.,* 1993) made it
clear that the prime target for work with synthetic doxyribo-oligonucleo-
tides should be a method for monitoring gene isolation. Discussions with
Dr. Benjamin D. Hall made me aware of the important studies of F.
Sherman on double-frameshift mutants of yeast cytochrome *c* (Stewart and

Sherman, 1974). Knowledge of the amino-acid sequence of these mutants allowed the unambiguous prediction of the N-terminal coding sequence and, hence, provided a specific target for an oligonucleotide of defined sequence. A second attractive feature about a yeast gene as a first target was the fact that the length of oligonucleotide required for a specific probe, 13 nucleotides (Table 1), was close to the limit for our enzymatic method of synthesis (Gillam and Smith, 1972, 1980; Gillam *et al.,* 1978). The required probe was synthesized using the enzymatic method (Gillam *et al.,* 1977) and was used to isolate the gene, albeit with some difficulty (Montgomery *et al.,* 1978), thus clearly validating the principle that a synthetic oligonucleotide could function as a probe for gene isolation. The subsequent generalization of the method using mixtures of oligonucleotides (Wallace *et al.,* 1981; Suggs *et al.,* 1981) has made this approach to gene isolation into a powerful tool in molecular genetics (Ausubel *et al.,* 1987; Sambrook *et al.,* 1989).

SEQUENCING DOUBLE-STRANDED DNA USING OLIGONUCLEOTIDE PRIMERS

At the time when the iso-1-cytochrome *c* gene was isolated, the chemical method for DNA sequence determination (Maxam and Gilbert, 1977) and the chain terminating enzymatic method (Sanger et al., 1977) had newly been developed, the former being applicable to double-stranded DNA and the latter to single-stranded DNA. The sequences of the ends of the Eco RI-Hind III DNA fragment containing the cytochrome *c* coding sequence were obtained using the chemical method (Montgomery *et al.,* 1978). I decided to investigate the possibility of directly sequencing double-stranded DNA by the enzymatic method using synthetic oligonucleotide primers, based on my conviction that an oligonucleotide should hybridize *at low temperature* with a complementary DNA strand of a denatured DNA double helix in solution, even though the second DNA strand was present and capable of producing a thermodynamically much more stable structure. The critical question was whether the second DNA strand would displace a short oligonucleotide from a complex with the complementary DNA strand under the conditions of the Sanger enzymatic sequencing method (Sanger, 1981). The first experiment was completely successful, an enormously exciting event. As a consequence, the sequence of the gene was completed by "walking" along the sequence using a series of short oligonucleotide primers 9 or 10 nucleotides in length (Smith *et al.,* 1979). I believe that the full potential of this approach to gene sequencing has yet to be realized.

 Another important application of oligonucleotide-primed sequencing of double-stranded DNA is the precise identification of point mutations produced by classical genetic techniques at a given locus. A particularly good example was in the analysis of a large number of independently isolated point mutations in a yeast suppressor t-RNA gene (Koski *et al.,* 1980; Kurjan *et al.,* 1980).

DEFINING THE ENDS OF mRNAs PRECISELY

It is important to exactly define the sequences at the 5'- and 3'-ends of m-RNAs. For the 3'-end of an m-RNA, this can be done by oligonucleotide-primed sequencing using reverse transcriptase for the Sanger sequencing method with RNA enriched for the desired m-RNA as template. The family of 12 deoxyribo-oligonucleotides consisting of a tract of dT-residues terminated at the 3'-end by two other deoxyribonucleotides can be annealed to the 3'-end of polyadenylated eukaryote m-RNA. If the transcribed sequence has a unique 3'-end, only one of the twelve oligonucleotides should act as a specific primer. This strategy was used to define the sequence of the unique 3'-end of bovine growth hormone m-RNA (Sasavage *et al.,* 1980) and to demonstrate and precisely define the microheterogeneity at the 3'-end of the bovine prolactin m-RNA transcript (Sasavage *et al.,* 1982). The iso-1-cytochrome *c* proved to have a unique 3'-end (Boss *et al.,* 1981) while the 5'-end of the same m-RNA is very disperse indeed as was shown by oligonucleotide primed extension using the m-RNA as template (McNeil and Smith, 1985). There are no other methods for defining the 5'- and 3'-ends of m-RNAs with this precision, a degree of precision which is essential to the accurate definition of transcription initiation and termination signals (McNeil and Smith, 1985; Sasavage *et al.,* 1982).

OLIGONUCLEOTIDE-DIRECTED MUTAGENESIS

> *'I strongly believe that, through the products of organic synthesis, it will be possible to gain influence over the development of organisms and to produce changes that surpass all that can be achieved by conventional breeding. "*
>
> Emil Fischer (1917)

> *"The ignis futuus of Genetics has been the specific mutagen, the reagent that would penetrate to a given gene, recognize it, and modify it in a specific way. "*
>
> J. Lederberg (1959)

Initiation of the project to isolate the yeast cytochrome *c* gene made me realize that I needed to learn about DNA sequence determination. I was lucky to be able to spend a year, starting in the fall of 1975, in Fred Sanger's laboratory working with the "plus-minus" sequencing method as part of the team which was attacking the sequence of the *E. coli* phage $\varnothing$X1 74 (Sanger and Coulson, 1975; Sanger *et al.,* 1978). There are three pairs of overlapping genes in the 5386 nucleotide genome (Barrel1 *et al.,* 1976; Brown and Smith, 1977; Gillam *et al.,* 1985) all genes being encoded on the same DNA strand, with each overlapping pair using different codon reading frames. An integral essential component of defining the position and reading frame of these genes was the availability of nonsense mutants suppressible by

amber or ochre suppressors. This use of precisely located mutants of a particular phenotype highlighted the need for a very specific mutagenic method that would define the target to a specific base-pair in the genome and introduce a predetermined change with sufficiently high efficiency to allow genomic screening in order to identify phenotypically silent mutants. Given that the DNA of phage ∅X174 is single stranded and with the knowledge of the complete sequence of its genome, our earlier studies, which demonstrated that small oligonucleotides as short as 7 nucleotides in length could form stable duplexes at low temperature even with a mismatch (Figure 2), suggested that oligonucleotide-directed mutagenesis should be possible. There was additional useful information. It was known that point mutants could be reverted, albeit with low efficiency, by annealing mutant phage ∅X174 DNA with fragments from the complementary strand of wild-type DNA prior to transfection (Hutchison and Edgell, 197 1; Weisbeek and van de Pol, 1970). However, fragments very much longer than those that we could readily synthesize were needed and the low efficiency precluded genotypic screening (Hutchison and Edgell, 1971; Weisbeek and van de Pol, 1970). In discussing these issues, Clyde Hutchison (who, also, was spending one year in Fred Sanger's group and whose biological knowledge of ∅X174 was invaluable to the sequencing project) and I realized that the studies of Kornberg and Goulian (Goulian, 1968a; Goulian *et al.*, 1967, 1973; Kornberg, 1980) provided an obvious route to a mutagenic method since they had demonstrated that an oligonucleotide as short as nine nucleotides in length could act as a primer for *E. coli* DNA polymerase I on a circular single strand template and that the product could be converted to a closed circular duplex by enzymatic ligation. It was also known that most of the primer molecule is excised from the product (Goulian, 1968b), presumably due to the 5'-exonuclease activity of *E. coli* DNA polymerase I (Klett *et al.*, 1968). With this information and our knowledge of the stability of duplexes involving a mismatched oligonucleotide, we decided to use a 12 nucleotide oligomer, with a centrally positioned single nucleotide mismatch, as primer and with ∅X174 DNA as template and *E. coli* D N A polymerase I in which the 5'-exonuclease had been inactivated by subtilisin (Brutlag *et al.*, 1969; Klenow and Henningsen, 1970; Klenow et al., 1971) to construct a closed circular double-stranded DNA with the oligonucleotide in one strand. Transfection of *E. coli* with this DNA should produce wild-type and mutant phage. The specific mutations chosen for the first experiment were the production and reversion of a known nonsense mutation, *am3*, in the lytic function, gene E, of ∅X174, since convenient phenotypic screens were available. The mutations involved the interconversion of a Trp codon, TGG, and an amber codon, TAG, by G-T and A-C mismatches (Hutchison *et al.*, 1978). In the first experiments, mutation was achieved but at low efficiency. This was increased to a very encouraging level of about 15 % after removal of incompletely closed duplexes by adsorption to nitro cellulose or treatment with a single-strand specific nuclease under conditions where a single base-pair mismatch was not degraded (Hutchison *et al.,*

1978). Further studies directed at optimizing the mutagenic conditions resulted in efficiencies of up to 39 % with a 12 nucleotide mutagenic primer and significant mutation with a heptamer mutagenic primer (Gillam and Smith, 1979a). The wisdom of using subtilisin-treated DNA polymerase is evident from comparison of these results with those of another study which used intact *E. coli* DNA polymerase I (Razin *et al.,* 1978).

Further studies on the mutagenesis of phage ∅X174 demonstrated that, in addition to the two transition mutations of the first experiments, it was possible to produce transversion mutations and single nucleotide deletions, also using very short oligonucleotides (Smith and Gillam, 1981). Clearly, the method was highly specific, efficient and general within the context of the genome of ∅X174. It was time to further generalize the method so that it could be applied to any cloned fragment of DNA. The icosahedral bacteriophages, such as ∅X174, are not potential vectors for recombinant DNA because the amount of DNA that can be accommodated in the viral particle is limited. The filamentous bacteriophages, on the other hand, can accommodate additional DNA incorporated into the genome. At the time when Mark Zoller and I were planning the generalization of the method, there were those that argued in favour of a methodology applied to double-stranded recombinant plasmids. It seemed to us that the biology of a filamentous bacteriophage, which produces single-stranded circular DNA and then purifies it as phage particles, argued overpoweringly in favour of phage DNA as vector. Subsequent developments in recombinant phage biology and in the construction of phagemids (Ausubel *et al.,* 1987; Sambrook *et al.,* 1989), have vindicated this strategic decision, although viable procedures for plasmid mutagenesis are available (Smith, 1985; Zoller, 1991, 1992). Oligonucleotide mutagenesis with filamentous phage DNA as template proved to be more difficult than with ∅X174 DNA, presumably because of a greater degree of secondary structure which inhibited the production of full length double-stranded circles. Consequently, new procedures for enhancing the efficiency of the mutagenic procedure had to be developed (Zoller and Smith, 1982, 1983, 1984, 1987). The power of the methodology was graphically illustrated by our first involvement in protein engineering in collaboration with Greg Winter and Alan Fersht (Winter *et al.*, 1982). This, of course, presaged the enormous use of the methodology in protein structure-function analysis that has become a feature of modern biochemistry.

GENOTYPIC SELECTION OR SCREENING FOR MUTANTS

Because the immediate product of the enzymatic incorporation of a mutagenic oligonucleotide into a double-stranded recombinant DNA is a heteroduplex, biological replication of the DNA will produce mutated and unmutated progeny DNAs, providing there is no asymmetry in the mechanism of DNA replication or providing that there is not an asymmetric mismatch repair mechanism. In fact, it is possible to achieve mutagenic efficiencies of

close to 50 % using standard procedures (Gillam and Smith, 1979a; Zoller and Smith, 1984). With this level of efficiency, screening by DNA sequence determination is the most effective procedure for detecting the desired mutant DNA. However, often the yield of mutant is considerably less than 50%. It is my opinion that the predominant reason for this is the use of template DNA that is contaminated with random oligonucleotides resulting from bacterial DNA degradation; others ascribe it to biological mismatch repair favouring the sequence of the template DNA (Smith, 1985). Whatever the cause of the reduced yield of mutant, procedures other than sequence determination are required for mutant identification. The ones that are available fall into two categories, those using the mutagenic oligonucleotide as a tool for identifying mutant DNA and those which select for progeny derived from the newly synthesized, oligonucleotide-containing DNA strand.

The significantly enhanced stability of a perfectly matched oligonucleotide duplex relative to one with a mismatch (Figure 2) suggested that a mutagenic oligonucleotide should be able to identify mutant complementary DNA. This principle was used to develop a method for *selection* of mutant DNA. In this procedure the mixture of wild-type and mutant single-stranded phage DNA produced by oligonucleotide mutagenesis is used as a template for double-stranded DNA synthesis using the mutagenic oligonucleotide as primer under conditions when it should not form a mismatched duplex (Gillam and Smith, 1979b). Following transfection into *E. coli,* this procedure can select mutant DNA with close to 100 % efficiency.

An alternative procedure is to use the mutagenic oligonucleotide as a probe to screen for mutant DNA in plaques produced by *E. coli* with mutagenized phage (Zoller and Smith, 1982, 1983). Again, this procedure is very reliable in the identification of mutant DNA. The same principle is used in the identification of point mutations in human DNA (Conner *et al.,* 1983) where, of course, the complexity of the human genome requires a probe containing a minimum of 17 nucleotides (Table 1).

A number of procedures for selection of progeny derived from the DNA strand containing the mutagenic oligonucleotide have been developed (Smith, 1985). Two of these, one based on *in vivo* selection and the other on *in vitro* selection have proved most popular and are widely used. *In vivo* selection involves the use of a template DNA where about 1% of the deoxythymidine residues have been replaced by deoxyuridine residues. This template is incorporated into double-stranded DNA where the mutagenic oligonucleotide is part of the *in vitro* synthesized second strand. On transfection of an appropriate *E. coli* host, the deoxyuridine strand is selectively degraded by the bacterium's DNA damage repair system, resulting in progeny derived primarily from the mutant strand (Kunkel, 1985). *In vitro* selection involves the use of a normal recombinant phage template DNA with the mutagenic oligonucleotide incorporated in the second strand which contains enzymatically synthesized thiophosphate inter-nucleotide linkages. These are resistant to cleavage by restriction endonuclease which,

as a consequence, nicks the template. The nicked template then is partially
degraded with an exonuclease and repaired using a DNA polymerase and
DNA ligase. The result is a double-stranded DNA which, predominantly, is
mutant in both strands and which produces, predominantly, mutant recom-
binant phage on transfection of *E. coli* (Taylor *et al.,* 1985).

CONCLUSION

This concludes my account of our studies on the use of synthetic dioxyribo-
oligonucleotides as tools for characterizing naturally occurring nucleic
acids, which is as accurate as I can make it, bearing in mind the caveat of
Fred Sanger cited at the beginning of this article. Clearly, many other
applications of oligonucleotide in biology exist, notably the polymerase
chain reaction (Mullis, 1994) and its multitudinous variations, but also in
double strand DNA synthesis and in the use of oligonucleotide probes *in
vitro* and *in vivo*. In the latter area there is the potential for development of
a whole new chemistry directed at pharmaceuticals which block the expres-
sion of specific genes and at diagnostic agents (Buchardt *et al.,* 1993;
Crooke, 1993).

   With regard to our own work, it is very satisfying that the model studies
started in 1968 provided the inspiration for so many applications and
scientific collaborations (only a few of which are cited in this article). The
potential of oligonucleotide-directed site-specific mutagenesis was apparent
from the outset; to quote:

> *"This new method of mutagenesis has considerable potential . . . to define
> the role of. . . origins of DNA replication, promotes and the sequences for
> ribosome-building sites. . . . specific mutation within protein structural
> genes . . . will allow precise studies of protein structure-function relation-
> ships. "*
>
> Hutchison *et al.* (1978)

However, we could not have anticipated the explosion of gene isolations,
the improvements in DNA sequence determination methodology and the
advances in the chemistry of nucleic acids synthesis that have occurred since
1978. This has resulted in an amazing increase in the use of site-directed
mutagenesis as an analytical tool in biochemistry and biology. And it has
been accompanied by continual improvements in the basic methodologies
and versatility of site-directed mutagenesis (Smith, 1985; Zoller, 1991,
1992) and the initiation of new scientific publications such as Protein
Engineering and Protein Science. It would not be too much of an exaggera-
tion to say that the prediction of Emil Fischer in 1917 has been fulfilled and
the dilemma posed by Joshua Lederberg in 1959 has been resolved.

ACKNOWLEDGMENTS

REFERENCES

Adams, S.P., Kavka, KS., Wykes, E.J., Holder, S.B. and Galluppi, G.R. (1983). Hindered dialkylamino nucleoside phosphite reagents in the synthesis of two DNA 51-mers. J. Am. Chem. Soc. 105: 661-663.

Astell, C.R., Suzuki, D.T., Klett, R.P., Smith, M. and Goldberg, I.H. (1969). The intracellular location of the adenine- and thymine-rich component of deoxyribon-ucleate in testicular cells of the crab, *Cancer productus.* Exptl. Cell Res. 54: 3-10.

Astell, C.R. and Smith, M. (1971). Thermal elution of complementary sequences of nucleic acids from cellulose columns with covalently attached oligonucleotides of known length and sequence. J. Biol. Chem. 246: 1944-1946.

Astell, C.R. and Smith, M. (1972). Synthesis and properties of oligonucleotide-cellulose columns. Biochemistry *11:* 4114-4120.

Astell, CR., Doel, M.T., Jahnke, P.A. and Smith, M. (1973). Further studies on the properties of oligonucleotide cellulose columns. Biochemistry 12: 5068-5074.

Atkinson, T. and Smith, M. (1984). Solid-phase synthesis of oligodeoxyribo- nucleo-tides by the phosphite-triester method. In: Oligonucleotide Synthesis. A Practical Approach (M.J. Gait, ed.). IRL Press, Oxford. pp. 35-81.

Ausubel, F.M., Brent, R, Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (1987). Current Protocols in Molecular Biology, (2 volumes, supple-mented quarterly). Wiley Interscience, New York, N.Y.

Aviv, H. and Leder, P. (1972). Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. Proc. Natl. Acad. Sci. USA 69: 1408-1412.

Barrell, B.G., Air, G.M. and Hutchison, C.A. III (1976). Overlapping genes in bacteriophage ∅X1 74. Nature 264: 34-41.

Borden, R.K. and Smith, M. (1966). Preparation of nucleoside-3',5' cyclic phos-phates in strong base. J. Org. Chem. 31: 3247-3253.

Boss, J.M., Gillam, S., Zitomer, R.S. and Smith, M. (1981). Sequence of the yeast iso-I-cytochrome c mRNA. J. Biol. Chem. 256: 12958-12961.

Brown, N.L. and Smith, M. (1977). The sequence of a region of bacteriophage ∅X174 coding for parts of genes A and B. J. Mol. Biol. *116:* l-28.

Brutlag, D., Atkinson, M.R., Setlow, P. and Kornberg, A. (1969). An active fragment of DNA polymerase produced by proteolytic cleavage. Biochem. Biophys. Res. Commun. 37: 982-989.

Buchardt, O., Egholm, M., Berg, R.H. and Nielsen, P.E. (1993). Peptide nucleic acids and their potential applications in biotechnology. Tibtech *11: 384-386.*

Chamberlin, M.J. (1965). Cooperative properties of DNA, RNA and hybrid homopo-lymer pairs. Fed. Proc. 24: 1446-1457.

Chambers, R.W., Moffatt, J.G. and Khorana, H.G. (1957). A new synthesis of guanosine-5'-phosphate. J. Am. Chem. Soc. 79: 3747-3752.

Conner, B.J., Reyes, A.A., Morin, C., Itakura, K., Teplitz, R.L. and Wallace, R.B. (1983). Detection of sickle-cell β^s-globin allele by hybridization with synthetic oligonucleotides. Proc. Natl. Acad. Sci. USA 80: 278-282.

Crooke, ST. (1993). Oligonucleotide therapy. Curr. Opin. Biotechnol. 3: 656-661.

Doel, M.T. and Smith, M. (1973). The chemical synthesis of deoxyribo-oligonucleo-tides complementary to a portion of the lysozyme gene of phage T4 and their hybridization to phage-specific RNA and phage DNA. FEBS Letts. 34: 99-102.

Fischer, E. (1917). Source unknown.

Gilham, P.T. (1962). Complex formation in polynucleotides and its application to the separation of polynucleotides. J. Am. Chem. Soc. 84: 1311-1312.

Gilham, P.T. and Robinson, W.E. (1964). The use of polynucleotide-celluloses in sequence studies of nucleic acids. J. Am. Chem. Soc. 86: 4885-4989.

Gillam, S. and Smith, M. (1972). Enzymatic synthesis of deoxyribo-oligonucleotides of defined sequence. Nature New Biol. 238: 233-234.

Gillam, S. and Smith, M. (1974). Enzymatic synthesis of deoxyribo-oligonucleotides of defined sequence. Properties of the enzyme. Nucleic Acids Res. *1:* 1631-1648.

Gillam, S., Waterman, K. and Smith, M. (1975). The base-pairing specificity of cellulose-pdT $_g$. Nucleic Acids Res. 2: 625-634.

Gillam, S., Rottman, F., Jahnke, P. and Smith, M. (1977). Enzymatic synthesis of oligonucleotides of defined sequence: synthesis of a segment of yeast iso-l-cytochrome *c* gene. Proc. Natl. Acad. Sci. USA 74: 96-100.

Gillam, S., Jahnke, P. and Smith, P. (1978). Enzymatic synthesis of oligo-deoxyribonucleotides of defined sequence. J. Biol. Chem. 253: 2532-2539.

Gillam, S. and Smith, M. (1979a). Site-specific mutagenesis using synthetic oligo-deoxyribonucleotide primers: optimum conditions and minimum oligodeoxyribonucleotide length. Gene 8: 81-97.

Gillam, S. and Smith, M. (19798). Site-specific mutagenesis using synthetic oligo-deoxyribonucleotide primers: in vitro selection of mutant DNA. Gene 8: 99-106.

Gillam, S. and Smith, M. (1980). Use of *E. coli* polynucleotide phosphorylase for the synthesis of oligodeoxyribonucleotides of defined sequence. Methods Enzymol. 65: 687-701.

Gillam, S., Atkinson, T., Markham, A. and Smith, M. (1985). Gene K of bacteriophage $\varnothing$X174 codes for a protein which affects the burst size of phage production. J. Virol. 53: 708-709.

Goulian, M., Kornberg, A. and Sinsheimer, R.L. (1967). Synthesis of infectious phage $\varnothing$X174 DNA. Proc. Natl. Acad. Sci. USA 58: 2321-2328.

Goulian, M. (1968a). Incorporation of oligodeoxynucleotides into DNA. Proc. Natl. Acade. Sci. USA *61: 284*-291.

Goulian, M. (1968b). Initiation of the replication of single-stranded DNA by *Escherichia coli* DNA polymerase. Cold Spring Harbor Symp. Quant. Biol. 33: 11-20.

Goulian, M., Goulian, S.H., Codd, E.E. and Blumenfield, A.Z. (1973). Properties of oligodeoxynucleotides that determine priming activity with *Escherichia coli* deoxyribonucleic acid polymerase I. Biochemistry *12:* 2893-2901.

Hutchison, C.A. III and Edgell, M.H. (1971). Genetic assay for small fragments of bacteriophage $\varnothing$X174 deoxyribonucleic acid. J. Virol. 8: 181-189.

Hutchison, C.A. III, 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.

Ingles, C.J., Trevithick, J.R., Smith, M. and Dixon, G.H. (1966). Biosynthesis of protamine during spermatogenesis in salmonid fish. Biochem. Biophys. Res. Comm. 22: 627-634.

Khorana, H.G., Tener, G.M. Moffatt, J.G. and Pol, E.H. (1956) A new approach to the synthesis of polynucleotides. Chem. and Ind. London, 1523.

Khorana, H.G. (1961). Some Recent Developments in the Chemistry of Phosphate Esters of Biological Interest. J. Wiley and Sons. N.Y.

Khorana, H.G. (1969). Nucleic acid synthethis in the study of the genetic code. In: Les Prix Nobel en 1968. P.A. Norstedt and Sons, Stockholm, pp. 196-220.

Khorana, H.G. (1979). Total synthesis of a gene. Science 203: 614-625.

Klenow, H. and Henningsen, 1. (1970). Selected elimination of the exonuclease activity of the deoxyribonucleic acid polymerase from *Escherichia coli* B by limited proteolysis. Proc. Natl. Acad. Sci. USA 65: 168-175.

Klenow, H., Overgaard-Hansen, K. and Patkar, S.A. (1971). Proteolytic cleavage of native DNA polymerase into two different catalytic fragments. Influence of assay conditions on the change of exonuclease activity and polymerase activity. Eur. J. Biochem. 22: 371-381.

Klett, R.P., Cerami, A. and Reich, E. (1968). Exonuclease VI, a new nuclease activity associated with *E. coli* DNA polymerase. Proc. Natl. Acad. Sci. USA 60: 943-950.

Kornberg, A. (1980). DNA Replication. W.H. Freeman and Company. San Francisco, CA.

Koski, R.A., Clarkson, S.G., Kurjan, J., Hall, B.D. and Smith, M. (1980). Mutations at the yeast *SUP4* tRNATyr locus; transcription of the mutant genes *in vitro* . Cell 22: 415-425.

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

Kurjan, J., Hall, B.D., Gillam, S. and Smith, M. (1980). Mutations at the yeast *SUP4* tRNA^Tyr locus: DNA sequence changes in mutants lacking suppresor activity. Cell 20: 701-709.

Lederberg, J. (1959). A view of genetics. In: Les Prix Nobel en 1958. Norstedt and Sons, Stockholm. pp. 170-189.

Maxam, A.M. and Gilbert, W. (1977). A new method for sequencing DNA. Proc. Natl. Acad. Sci. USA 74: 560-564.

McBride, L.J. and Caruthers, M.H. (1983). An investigation of several deoxyribonucleoside phosphoramidates useful for synthesizing deoxyoligonucleotides. Tetrahedron Lett. 24: 245-248.

McConaughy, B.L. and McCarthy, B.J. (1967). The interaction of oligodeoxyribonucleotides with denatured DNA. Biochim. Biophys. Acta *149:* 180-189.

McNeil, J.B. and Smith, M. (1985). Saccharomyces *cerevisiue CYC1* mRNA 5'-end positioning: analysis by *in vitro* mutagenesis, using synthetic duplexes with random mismatch base pairs. Mol. Cell Biol. 5: 3545-3551.

Michelson, A.M. and Monny, C. (1967). Oligonucleotides and their association with polynucleotides. Biochim. Biophys. Acta *149: 107-126.*

Montgomery, D.L., Hall, B.D., Gillam, S. and Smith, M. (1978). Identification and isolation of the yeast cytochrome c gene. Cell 14: 673-680.

Mullis, K. (1994). Les Prix Nobel en 1993. Norstedt and Sons, Stockholm. pp. 107-117.

Niyogi, S.K. and Thomas, C.A., Jr. (1968). The stability of oligoadenylate-polyuridylate complexes as measured by thermal chromatography. J. Biol. Chem. 243: 1220-1223.

Niyogi, S.K. (1969). The influence of chain length and base composition on the specific association of oligoribonucleotides with denatured deoxyribonucleic acid. J. Biol. Chem. 244: 1576-1581.

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

Sambrook, J. Fritsch, E.F. and Maniatis, T. (1989). Molecular Cloning. A Laboratory Manual, 2nd Edition (3 volumes). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Sanger, F. and Coulson, A.R. (1975). A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. J. Mol. Biol. 94: 441-448.

Sanger, F., Coulson, A.R., Friedmann, T., Air, G.M., Barrell, B.G., Brown, N.L., Fiddes, J.C., Hutchison, C.A. III, Slocombe, P.M. and Smith, M. (1978). The nucleotide sequence of bacteriophage ⌀X174. J. Mol. Biol. *125: 225-246.*

Sanger, F. (1981). Determination of nucleotide sequences in DNA. In: Les Prix Nobel en 1980. Norstedt and Sons, Stockholm. pp. 143-159.

Sanger, F. (1988). Sequences, sequences and sequences. Ann Rev. Biochem. 57: l-28.

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

Sasavage, N.L., Smith, M., Gillam, S., Astell, C., Nilson, J.H. and Rottman, F. (1980). Use of oligodeoxynucleotide primers to determine poly(adenylic acid) adjacent sequences in messenger ribonucleic acid. Biochemistry 19: 1737-1743.

Sasavage, N.L., Smith, M., Gillam, S., Woychik, R.P. and Rottman, F.M. (1982). Variation in the polyadenylation site of prolactin messenger RNA. Proc. Natl. Acad. Sci. USA 79: 223-227.

Smith, M. and Khorana, H.G. (1958). An improved and general method for the synthesis of ribo- and deoxyribo-nucleoside-5′ triphosphates. J. Am. Chem. Soc. 80: 1141-1145.

Smith, M. and Khorana, H.G. (1959). Specific synthesis of the C5′-C3′ inter-ribonucleotide linkage: the synthesis of uridylyl-(3′N5′)-uridine. J. Am. Chem. Soc. *81:* 2911.

Smith, M. Drummond, G.I. and Khorana, H.G. (1961). The synthesis and properties of ribonucleoside-3′,5′ cyclic phosphates. J. Am. Chem. Soc. 83: 698-706.

Smith, M., Rammler, D.H., Goldberg, I.H. and Khorana, H.G.. (1961). Synthesis of **uridylyl-(3′« 5′)-Uridine** and **uridylyl-(3′«5′)-adenosine**. J. Am. Chem. Soc. 84: 430 - 440.

Smith, M. (1964). Synthesis of deoxyribonucleoside-3′5′ cyclic phosphates by base-catalysed transesterification. J. Am. Chem. Soc. 86: 3586.

Smith, M., Leung, D.W., Gillam, S., Astell, C.R., Montgomery, D.L. and Hall, B.D. (1979). Sequence of the gene for iso-l-cytochrome *c* in *Saccharomyces cerevisiae.* Cell *16:* 753-761.

Smith, M. and Gillam, S. (1981). Constructed mutants using synthetic oligodeoxyribonucleotides as site-specific mutagens. In: Genetic Engineering. Principles and Methods (J.K. Setlow and A. Hollaender, eds.). Volume 3. Plenum Press, New York, N.Y. pp. l-52.

Smith, M. (1983). Synthetic oligodeoxyribo nucleotides as probes for nucleic acids and as primers in sequence determination. In: Methods of DNA and RNA Sequencing (S.M: Weissman, ed.). Praeger Publishers, New York, N.Y. pp. 23-68.

Smith, M. (1985). In vitro mutagenesis. Ann. Rev. Genet. 19: 423-462.

Stewart, J.W. and Sherman, F. (1974). Yeast frameshift mutations identified by sequence changes in iso-1-cytochrome *c*. In: Molecular and Environmental Aspects of Mutagenesis (L. Prakash, F. Sherman, M.W. Miller, C.W. Lawrence and Taber, H.W., eds.). Charles C. Thomas, Springfield, IL pp. 102-107.

Suggs, S.V., Wallace, R.B., Hirose, T., Kawashima, E.H. and Itakura, K. (1981). Use of synthetic oligonucleotides as hybridization probes: isolation of cloned cDNA sequences for human β_2-microglobulin. Proc. Natl. Acad. Sci. USA 78: 6613-6617.

Taylor, J.W., Ott, J. and Eckstein, F. (1985). The rapid generation of oligonucleotide-directed mutations at high frequency using phosphothioate-modified DNA. Nucleic Acids Res. 13: 1451- 1455.

Tigerstrom, R. von and Smith, M. (1970). Oligodeoxyribo-nucleotides: chemical synthesis in anhydrous base. Science *167:* 1266-1268.

Tigerstrom, R. von, Jahnke, P. and Smith, M. (1975a). The synthesis of the internucleotide bond by a base-catalysed reaction. Nucleic Acids Res. 2: 1727-1736.

Tigerstrom, R. von, Jahnke, P., Wylie, V. and Smith, M. (1975b). Application of base-catalysed reaction to the synthesis of dinucleotides containing the four common deoxyribonucleosides and of oligodeoxythymdylates. Nucleic Acids Res. 2: 1737-1743.

Wallace, R.B., Johnson, M.J., Hirose, T., Miyake, T., Kawashima, E.H. and Itakura, K. (1981). Hybridization of oligonucleotides of mixed sequence to rabbit b-globin DNA. Nucleic Acids Res. 9: 3647-3656.

Watson, J.D., Hopkins, N.H., Roberts, J.W., Steitz, J.A. and Weiner, A.M. (1987). Molecular Biology of the Gene. 4th Edn. The Benjamin/Cummings Publishing Company, Inc. Menlo Park, CA.

Watson, J.D., Gilman, M., Witkawski, J. and Zoller, M. (1993). Recombinant DNA. 2nd Edn. Scientific American Books, New York, N.Y.

Weisbeek, P.J. and van de Pol, J.H. (1970). Biological activity of ∅X174 replicative form DNA fragments. Biochim. Biophys. Acta 224: 328-338.

Winter, G., Fersht, A.R., Wilkinson, A.J., Zoller, M.J. and Smith, M. (1982). Redesigning enzyme structure by site-directed mutagenesis: tyrosyl t-RNA synthetase and ATP binding . Nature 299: 756-758.

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. Nucleic Acids Res. 10: 6487-6500.

Zoller, M.J. and Smith, M. (1983). Oligonucleotide-directed mutagenesis of fragments cloned in Ml3 vectors. Methods Enzymol. 100: 468-500.

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

Zoller, M.J. and Smith, M. (1987). Oligonucleotide-directed mutagenesis: a simple method using two oligonucleotide primers and a single-stranded DNA template. Methods Enzymol. *154: 329-350.*

Zoller, M.J. (1991). New molecular biology methods for protein engineering. Curr. Opin. Biotechnol. 2: 526-531.

Zoller, M.J. (1992). New recombinant DNA methodology for protein engineering. Curr. Opin. Biotechnol. 3: 348-354.