Indian J. Phys. **65B** (1), 12-24 (1991)

Glimpses of the emerging technology, genetic engineering, and its applications in medical sciences

S N Chatterjee
Biophysics Division, Saha Institute of Nuclear Physics,
37 Belgachia Road, Calcutta-700 037, India

Abstract : The history of development of the Genetic Engineering Technology has been brought out with a brief description of the current status of some of the techniques used in gene cloning. The achievements and potential applications of the technology in medical sciences and its possible impact on the Society in future have been discussed, as far as possible, in a semitechnical language.

Keywords : Genetic Engineering, gene cloning, medical applications, social impact.

PACS Nos : 87 Biophysics, medical physics and biomedical engineering; 87.90.+y Other topics in biophysics, medical physics and biomedical engineering.

1. Introduction

This article is addressed basically to intelligent laymen belonging to physical, biological and medical sciences. All attempts have been made to present facts and figures in a semitechnical language. The article presents the New Technology through four different facets : (i) introduction of the technology by bringing out findings and discoveries which led to its emergence, (ii) brief description of the essential methods that are used in the practice of the technology, (iii) potential applications of this technology and achievements made in the field of medical sciences with particular reference to the emergence of new pharmaceutical industries and (iv) brief analysis of the wider applications and implications of this technology to the human race of the days to come.

In the living world, Man has the distinct advantage of being endowed with the noble virtue of the power of thinking and accordingly he is being constantly goaded by the eternal thirst for knowing the unknown and seeing the unseen. And to his amazement and wonder he finds that he himself, his origin and action, is a mystery of mysteries. He has thus been puzzled through ages, through generations, in understanding himself. Great philosophers have exposed their learned views on clarifying what life is and how it originated. It has been expressed that life is something supernatural, something which goes beyond explanation, something

which is created by someone who is omnipotent, omniscient and omnipresent. But all along some other philosophers have been constantly refusing to accept this obscurantism as a solution and have been trying to formulate the basic concepts of life in terms of the principles and theories of the natural sciences. They argue that all forms of life have so far been found to consist of the same atoms or elements which in some way or other constitute the various inanimate objects. But the molecules which form by the combination of atoms might have some differentiating features. The simplest manifestation of life has been found in the viruses, the very small organisms whose dimensions might be as small as 100 Angstrom (10^{-6} cm). The viruses are unique in the sense that they share the properties of both the living and non-living entities. They are living in the sense that they can reproduce themselves in the exact sense. They are non-living in the sense that they can be crystallised and beautiful crystals of different viruses have been extensively used for understanding their structural organisations. Surely we have no experience of crystalline dogs so far and I shudder even to think of crystalline human beings ! Many of the simpler viruses have been found to consist of only two different types of chemical molecules, viz., the proteins and the nucleic acids and these are present in all forms of life known so far. It thus appears that these two chemicals are the minimum requirements for the manifestation of life. One may now ask as to whether the secrets of life lie in -

(i) the peculiar association of both the chemical molecules, protein and nucleic acid ;

or, (ii) one of them is of primary character and dictates the process of life ;

or, (iii) none of them is at all necessary and essential for the manifestation of life. The key to the secret of life probably lies in some still unknown physico-chemical agent ;

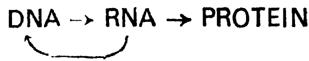
or, (iv) the mysteries of life are beyond the scope of the natural sciences.

The unequivocal answer to these questions has yet to be obtained. However, the achievements and progress of the science of molecular biology (which originated only some 30 years ago) have already been so spectacular as to rouse the natural enthusiasm and hope that life may be exactly known in terms of the physico-chemical forces in the days ahead. It has by now been apparent that protein is rather the food stuff of life, the message of which is actually coded in the structure of the nucleic acid molecules.

2. DNA, genes and diseases

The nucleic acid is mainly of two types, the deoxyribonucleic acid or DNA and ribonucleic acid or RNA. DNA is mostly of primary character. A nucleic acid molecule contains a linear array of four different simpler molecules known as nucleotides and the sequence in the arrangement of these nucleotides in the nucleic acid molecule of any species is supposed to decide its overall genetic and functional potential (Figure 1). This sequence of nucleotides determines the type or types

of protein molecules the particular species will or can produce and so on. Information in living systems is thus supposed to flow in general from DNA to RNA and then to protein molecules.



Information normally flows unidirectionally, from DNA to RNA and then to Protein. Only in cases of some specific RNA viruses, the information initially flows

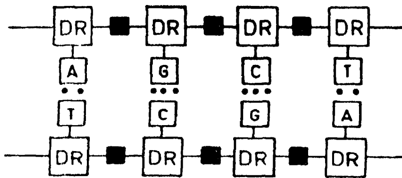


Figure I. Chemical structure of a double stranded DNA chain. The two strands are each made of a deoxyribose sugar molecule (DR) and a phosphate group (■) alternately joined to each other by covalent bonds. The DNA chain here shows an arbitrary sequence of the four bases, adenine (A), guanine (G), cytosine (C) and thymine (T) along its length. The complementary nature of base pairing, i.e., adenine with thymine (by two hydrogen bonds) and guanine with cytosine (by three hydrogen bonds) and vice-versa should be noted. A nucleotide molecule comprising one sugar molecule, one phosphate group and one of the four bases (A, T, G or C) forms an elementary or monomeric unit of the polymeric DNA molecule.

in the reverse direction, from RNA to DNA, and then as usual to RNA and Protein. Elucidation of the structure of different DNA molecules and in particular of the sequence of arrangement of the four different nucleotides along their chains thus form a very important area of investigation in molecular biology. A simple mathematical analysis can give us an idea of the information content in a DNA molecule, assuming that the information content is coded in the sequence of bases along the chain. If a DNA chain length is such as to contain only five nucleotides, then the number of ways in which these five nucleotides can be arranged along the DNA chain length is simply $4^5 = 1024$ assuming that there is endless supply of each of the four nucleotides (A*, T*, G* or C*) and that there is no restriction in the repetition of any nucleotide along the DNA length. If the DNA molecule is so long as to contain 10^8 nucleotides along its length (as in human DNA), then the number of ways these four nucleotides can be arranged along the DNA length is 4^{10^8}, i.e., an extremely large number. For such a DNA molecule, the information content may thus be varied by almost an infinitely large number of possibilities. This is a simple but crude explanation of the existence of infinite number of variations in the living world. A gene is a small segment of a DNA molecule having a specific base sequence and a specific information

A*—adenylic acid ; T*—thymidylic acid ; G*—guanylic acid ; C*—cytidylic acid.

content. The size of the entire DNA molecule (genome DNA) of an organism varies in accordance with the nature and size of the organism. Similarly, the size of a gene also varies in accordance with the nature of the gene (Table 1). Also

Table I. A comparative picture of the sizes of some genomic DNAs and genes.

Genome DNA	Size* (no. of bases)	Genes	Size* (no. of bases)
Virus			
$\phi\chi$ 174	5×10^3	transfer RNA	75-90
λ wild type	46.5×10^3	Xenopus laevis 5S-RNA	120
Simian Virus SV40	5.3×10^3	Sea Urchin histone gene	6.54×10^3
Bacterium			
Escherichia coli	4.0×10^6	Ovalbumin natural gene	7×10^3
Commonly used plasmid vehicle	$3\text{-}30 \times 10^3$	Coding sequence	1859
Human	10^9		

*For double stranded DNA, size is in no. of base pairs.

genes in the cells of more complex organisms are different in structure from the genes in bacteria. In such cells, the coding region of a gene is broken up along the DNA strand. Determination of base sequence in a DNA molecule is thus an important proposition and the men who devised methods for sequence determination were aptly rewarded with Nobel Prize. In fact, the development of rapid DNA sequencing techniques by the mid '70s contributed the final element required to make an entire recombinant DNA technology. In patients suffering from sickle-cell anaemia, the amino acid valine replaces glutamic acid in the chain of the vital blood protein, haemoglobin. The root of this defect could be traced to the alteration of a single nucleotide in corresponding DNA chain, i.e. thymidylic acid replaced by adenylic acid.

Segment of the β-Chain of haemoglobin and its gene

DNA ... TGT GGG TAC TAC TTT ...
RNA ... ACA CCC AUG AUG AAA ...
HbA ... Thr – Pro – Glu – Glu – Lys ... Normal adult

DNA ... TGT GGG AAC TAC TTT ...
⇑
RNA ... ACA CCC UUG AUG AAA ...
⇑
HbS ... Thr – Pro – Val – Glu – Lys ... Patient with sickle-cell anaemia

Haemoglobin contains two 'alpha' and two 'beta' chains. Each 'alpha' chain contains 141 amino acids and each 'beta' chain 146 amino acids. In sickle-cell

anaemia patients, the beta chain is defective, the glutamic acid in position six of the peptide chain being replaced by valine. This defect could be traced to a defect in DNA nucleotide sequence in the haemoglobin gene for β-chain. The importance of the sequence of nucleotide can thus be appreciated. We are naturally not satisfied with this knowledge of the basis of the disease sickle-cell anaemia, we would obviously like to devise methods so that we can alter this defective sequence and set it right to cure the disease. We are getting involved in devising techniques for altering the sequence of nucleotides in DNAs of living species either by deletion or addition of an extra piece of a short DNA chain. This ushers in the new discipline of science, commonly known as genetic engineering, which basically involves DNA technology. Major advances achieved recently in the field of DNA technology have not only endowed credibility to it but also established the science of genetic engineering as one of the most prospective technologies born in the present days.

3. Genetic engineering : basic aspects

Much of the early works in the field of genetic engineering involve the manipulation of the genetic structure of a bacterium, *Escherichia coli* which is commonly found in our digestive tract and which can be easily grown in the laboratory (Cohen, 1975). These bacteria usually contain two types of DNA molecules within their cell, of which one is very large in size and is usually known as its chromosome or genome. The other type consists of one or more much smaller DNA molecules which the bacteria can shed off or acquire under various environmental conditions, can be propagated in alternative states, either autonomously in the cytoplasm or as an integral part of the bacterial chromosome and are generally known as plasmids. Sometimes a plasmid can pick up a short segment of DNA from the chromosome of its own cell and transfer it to the cell of a related bacterial species, and sometimes the plasmid and the segment of chromosomal DNA can become integrated into the chromosome of the recipient cell. By suitable microbiological and chemical processes, these plasmids can be removed from the bacteria, cut open at a specific site in the laboratory (since normally they do not provide any free end, i.e. they form closed circles), joined at one end to a new piece of gene bearing the required information (say, the information to produce any particular protein), resealed to form a bigger circle and then reinserted in the same bacterial cell. This bacterium can then be grown again in the laboratory and induced at will to synthesize the particular protein the information for which was thus inserted into its plasmid. Stanley N. Cohen and associates were thus able to insert into *E. coli* some genes of the toad Xenopus laevis. The composite plasmid DNA thus formed was termed DNA chimera in analogy with mythological chimera (e.g., a creature with the head of a lion, the body of a goat and the tail of a serpent).

4. Early breakthroughs

These achievements were made possible by a series of independent discoveries made in rapid succession in the late 1960's and early 1970's and included at least

several distinct elements : (i) the discovery of means for the cleavage of DNA at highly specific sites, (ii) the development of simple and generally applicable methods for the joining of DNA molecules, (iii) the discovery of a suitable gene

Table 2. Some selected restriction enzymes generating butt and cohesive ends with the specific nucleotide sequences they recognize and the organisms from which they are purified (origin).

Enzyme	Origin	Nature of Ends generated	Recognition sequence and cleavage site
Alu I	Arthrobacter luteus	Butt	5′-AGCT-3′ 3′-TCGA-5′
Bal I	Brevibacterium albidum	Butt	5′-TGGCCA-3′ 3′-ACCGGT-5′
Hinc II	Haemophilus influenzae Rc	Butt	5′-GT py pu AC-3′ 3′-CA pu py TG-5′
Bam HI	Bacillus amyloliquifaciens	Cohesive	5′-GGATCC 3′ 3′-CCTAGG-5′
Eco RI	*Escherichia coli* RY 13	Cohesive	5′-GAATTC-3′ 3′-CTTAAG-5′
Hind III	Haemophilus influenzae Rd	Cohesive	5′-AAGCTT-3′ 3′-TTCGAA-5′
Pst I	Providencia stuartii 164	Cohesive	5′ CTGCAG-3′ 3′-GACGTC-5′

carrier that can replicate both itself and a foreign DNA segment linked to it, (iv) the devising of means of introducing the composite DNA molecule into a

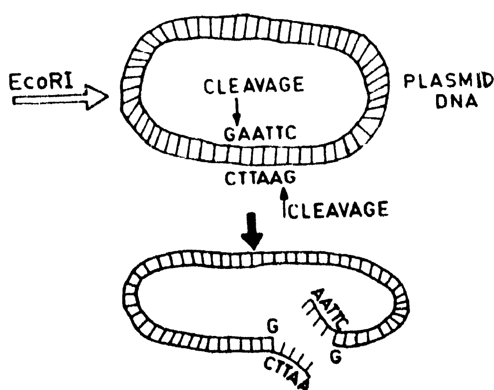


Figure 2. The restriction endonuclease Eco RI has the property of cleaving complementary strands of DNA at sites (shown by arrows) four nucleotides apart and recognises the base sequence GAATTC. Such cleavage yields DNA fragments with complementary overlapping single-stranded ends. As a result, the end of any DNA fragment produced by Eco RI cleavage can anneal with any other fragment produced by the same enzyme.

functional bacterial cell etc. These ingenuous biochemical and biological manipulations were made easier by the discovery of a group of enzymes called restriction endonucleases (Roberts 1980) which could cleave DNA at specific sites producing either complementary ended fragments or fragments with flush or butt ends (Table 2). The specific endonuclease, Eco RI has the property of cleaving complementary strands of DNA at sites four nucleotides apart (Figure 2). Such cleavage yields DNA fragments with complementary overlapping single-stranded ends. As a result the end of any DNA fragment produced by Eco RI cleavage can anneal with any other fragment produced by the enzyme. The 1978 Nobel Prize in Medicine was aptly awarded to the discoverers of these enzymes.

5. Gene cloning : essential steps

Some of the essential steps involved in the DNA cloning (Sinsheimer 1977, Glover 1980, Perbal 1984) procedure are (i) choice of cloning organism, (ii) choice of cloning vehicle (vehicle DNA), (iii) isolation of DNA to be cloned (Foreign DNA), (iv) integration of foreign DNA into vehicle DNA (production of DNA chimera), (v) insertion of DNA chimera into cloning organism and (vi) selection of cloned DNA (Figure 3). For each of these steps, various alternative procedures are available and the selection of any depends on the nature of the problem to be tackled. The ideal cloning organism is one which can be easily grown, which lacks the restriction enzymes (so that the cloned DNA is not chopped), which is preferably recombination deficient (*rec* A^-) etc. The *Escherichia coli* bacteria have been widely used for such purposes for many such obvious reasons. As for the cloning vehicle is concerned, it should be self replicating, should be amenable to easy manipulation (isolation, transfer etc.), should have the property that helps in the selection of cloned DNA (antibiotic resistance factors etc.), should help in the production of multiple copies of cloned DNA (gene amplification) and should be easily amenable to cleavage by restriction enzymes and splicing with foreign DNA. Some of the commonly used vectors are plasmids, DNA of the bacteriophage λ or the DNA of the virus SV-40. As for the isolation of foreign DNA (or gene) or the DNA (or gene) to be cloned, methods vary with the nature of the problem. In some cases it so happens that the isolation of a specific gene or DNA is a very difficult task but the isolation of the corresponding messenger RNA (m-RNA) in a pure form is relatively much easier. In such cases, the foreign DNA is synthesized from the m-RNA by enzymatic means (Figure 4). Such DNA is termed complimentary DNA or C-DNA. In some other cases it may so happen that a particular gene or DNA can be synthesized by chemical methods and such DNA is termed synthetic DNA. An example of such a synthetic DNA is the gene for tyrosine specific transfer RNA synthesized by Khorana. The integration of foreign DNA into vehicle DNA involves a number of steps : (i) cleavage of both the foreign and vehicle DNA by appropriate restriction enzymes and (ii) mixing and annealing of the cleaved vehicle and foreign DNA fragments. In some cases, the cleaved DNA fragments of both kinds (vehicle and foreign) are treated with enzymes such as

terminal deoxynucleotidyl transferase (TDNT), to add homopolymer tails of deoxynucleotides to each DNA fragment in such a way that annealing is much facilitated (Figure 5). In such cases, single stranded regions appear in the annealed DNA chimera and such single stranded gaps are to be repaired by treatment with the

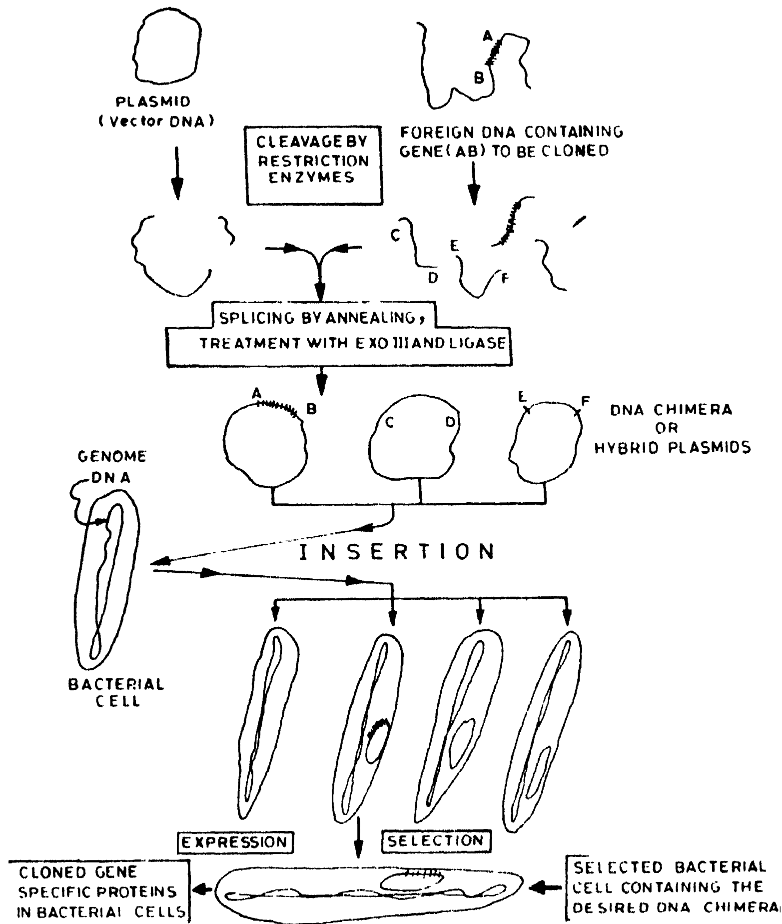


Figure 3. A schematic presentation of the essential steps involved in the cloning of a piece of DNA or gene using a bacterial cell. The human insulin gene was similarly inserted into the bacteria, *Escherichia coli*.

enzyme exonuclease III and finally (iii) the annealed DNA is sealed by treatment with DNA ligase. The reconstructed plasmid or plasmid chimera thus produced needed to be reintroduced in a bacterial cell by a process called transformation and allowed to be expressed. As stated earlier, *Escherichia coli* is often the organism of choice for cloning. For reinsertion of the plasmid chimera, the bacterial cell had to be treated with calcium chloride solution of appropriate molarity at 0°C and then

subjected to a heat shock by suddenly raising the temperature of the mixture. This treatment usually makes the bacterial cell permeable to foreign DNA and the plasmid chimera could thus be reinserted. As illustrated in Figure 3, different types of plasmid chimera can be formed during the annealing of the mixture of foreign and vehicle DNA components and as such different plasmid chimeras could

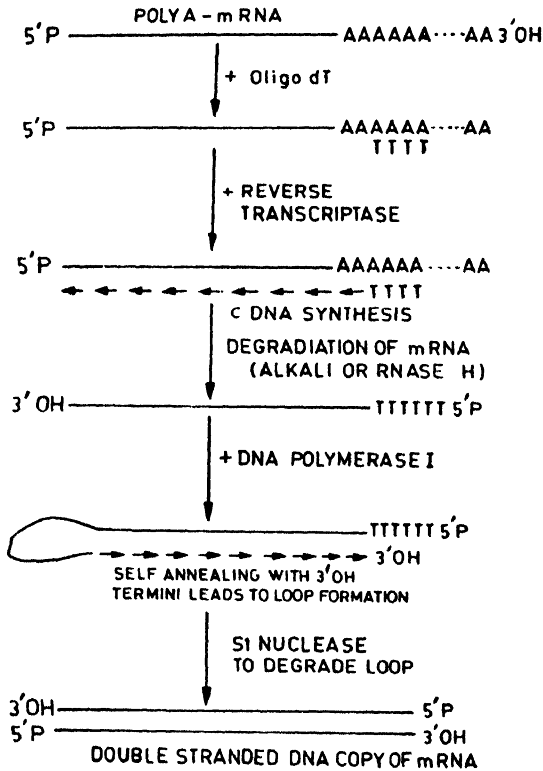


Figure 4. Schematic illustration of the steps involved in the synthesis of complementary DNA (C-DNA) from the messenger RNA (m-RNA) coupled to poly A. The enzyme reverse trancriptase has been used to synthesize the C-DNA by using the m-RNA as template and poly *dT* as the primer. The polymerase I has been used to catalyze a 5' → 3' elongation of DNA strand in the presence of primer and deoxynucleoside triphosphate and finally SI endonuclease has been used to degrade the single stranded DNA (portion in the form of a loop) part.

be introduced into the bacterial cell. It is thus necessary to select the right type of bacterial cells, i.e., those cells which had got within them the plasmid chimera containing the particular gene to be cloned, by some means or other. Various methods of selection are available. Some of these methods take recourse to purification of genes, by physical and/or microbiological methods, prior to cloning while others do the selection after cloning. In one such method, the vector DNA, plasmid rWL7, is cleaved by the restriction enzyme Pst I thereby inactivating the

ampicillin resistance factor normally present in the plasmid. The foreign DNA is then inserted at the cleavage site of the plasmid and the recombinant DNA plasmid is allowed to transform the bacterium *E. coli*. These bacteria are then cultured in media containing ampicillin and D-cycloserine besides the nutrients. Because of the presence of ampicillin in the growth medium, all growing bacteria except the

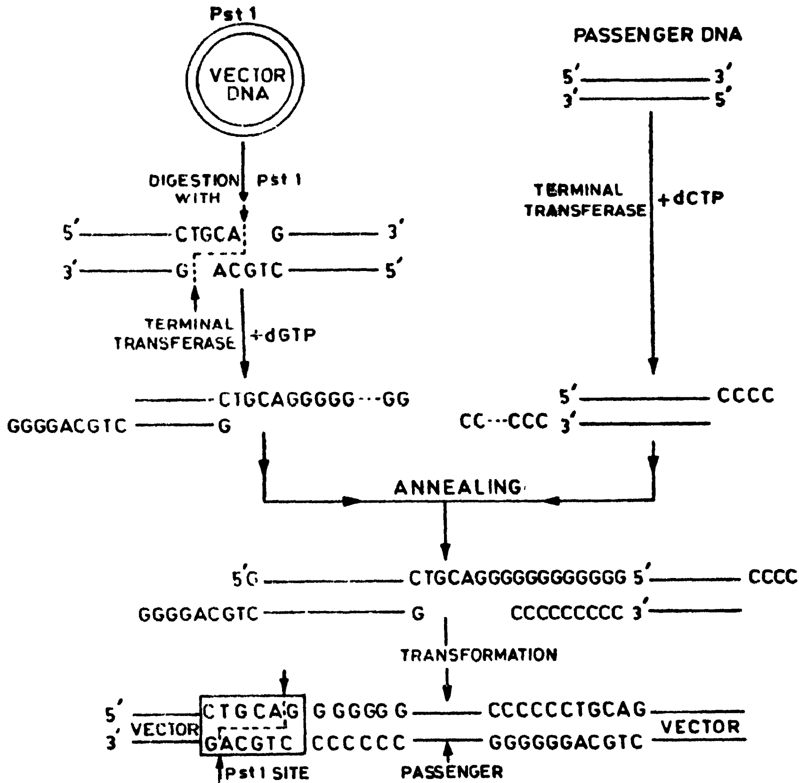


Figure 5. Illustration of the use of homopolymer tails of deoxynucleotides for cloning purpose. Here the foreign DNA with *dC* tails has been cloned into a vector DNA whose cloning site has been elongated with *dG* tails. In this method the cloned foreign DNA can be recovered because the Pst I recognition sites are regenerated after ligation.

actual transformants (those containing the foreign DNA inserted in the plasmid chimera by inactivating the ampicillin resistance factor) will be killed. Thus the actual transformants can be enriched and selected. The transformed cell, when it divides, will copy the hybrid DNA and pass it on to its daughter cells. If the new gene in the hybrid DNA is expressed, each of the daughter cells will make the new protein encoded for in the DNA.

6. Medical applications

Following similar procedures, a spectacular break-through in genetic engineering has been achieved very recently by the joint collaboration of the City of Hope

Medical Centre in California and the San Francisco firm, Genentech (Gilbert and Taunton-Rigby 1984). The achievement was the production of human insulin by the bacterium, *Escherichia coli*. Insulin is a protein based hormone normally produced in our pancreas. Diabetic patients need supply of insulin from outside sources, mostly from the pancreas of slaughtered cattle and pigs. It takes the insulin content of two cows to satisfy the needs of a diabetic for one year. If one takes account of the gradually increasing number of the diabetics in the world and also of the fact that their lives are now being prolonged by virtue of their growing awareness and better medical attention they receive, one can immediately appreciate the need for developing alternative source of insulin production. In this respect, genetically engineered bacteria can provide an unending source of insulin. The achievements of the City of Hope research team include the chemical synthesis of the human insulin gene in the laboratory test tube. Cattle insulin or pig insulin differs from human in several amino acids, and therefore may produce and does in fact produce allergic and untoward immunological reactions in many diabetic patients. The human insulin manufactured in bacteria will eliminate this drawback. In 1977, the same research team succeeded in synthesizing the human gene which produces the brain hormone, somatostatin. The next important step was to insert the insulin gene next to a cluster of bacterial genes called the "lac operon". The "lac operon" can be imagined as a "kind of switch system" which can regulate the production of protein in the bacterium. The production of insulin can then be turned on by adding certain chemicals in the bacterial growth medium.

The new technology promises to have unlimited applications in medical science (Gilbert and Taunton-Rigby 1984). The recombinant DNA technology can be used to produce commerical quantities of human enzymes and hormones. Production of insulin in bacteria, as elaborated earlier, is one of the early spectacular achievements of this technology. More than one hundred new biotechnology companies, e.g., (i) Cetus Corporation, Emeryville, California, (ii) Genentech Inc., San Francisco, (iii) Biogen, Geneva etc., have started functioning in this area. The way is now open for the manufacture of virtually any hormone, vaccine, antibodies etc. which has a protein base. Some of the first products that are emerging are human insulin, alpha interferon, gamma interferon, human and animal growth hormones, vaccines for foot-and-mouth disease virus and for hepatitis B etc. Plasminogen activators are forthcoming. These proteins can dissolve blood clots and therefore may be used to treat heart disease. Manufacture of these human products by this technology have additional advantage. Compounds extracted from animal or plant tissues exhibit a diverse range of adverse effects. Many plant products are effective but addictive, such as codeine. Animal proteins are frequently inactive in man or not well tolerated and cause immunological responses. Eventually we can expect to have available every active human protein in amounts sufficient for full clinical trials and at prices low enough for widespread medical use.

In the not too distant future, gene splicing technology is likely to engender a thorough understanding of how cells work on the molecular level. Precise knowledge will evolve on the differentiation of cells, on the signals that specify, dictate or modify growth, on how a normal cell is transformed into a malignant one etc. The knowledge could certainly be used to renew a kidney or a cirrhosed liver, to revert back a malignant cell to a normal one, to change fat deposits etc. With these achievements, the practice of medicine in the days to come will undergo a dramatic change. Now we treat illness after they occur, tomorrow we will focus on detection before symptoms appear, on testing for disease susceptibility, on prevention of diseases etc. Further, the ability to recombine DNA will lead to new and better diagnostic methods for detecting infectious agents and inherited diseases. It is no longer a fiction that someday the entire genetic code of each human being will be deciphered and our lives will undergo a total transformation. Besides application in medical science, the new technology has tremendous potential in bringing beneficial changes in diverse aspects of human civilization including environmental pollution, energy and food production.

7. Possible misuse of this technology

Any new scientific achievement may have both the potentially beneficial and also harmful possibilities and genetic engineering is no exception to this. The potential dangers were imagined to arise mostly from uncontrolled, rather ill-controlled, or motivated use of the techniques. Of particular concern is the fact that this new technology, which permits combination of genetic information from very different organisms, places us in an area of biology with many unknowns. These researches may thus lead to worldwide epidemics caused by newly created pathogenic bacteria let loose in the environment, the triggering off of catastrophic ecological imbalances, new tools for militarists and terrorists etc. Such hypothetical risks however could not be brushed aside lightly and a Committee on Recombinant DNA Molecules was established in 1973 in USA to consider imposition of restrictions on researches in genetic engineering. The Asilomar Conference held in 1975 at the Pacific Grove in California put forward a new concept of biological containment and certain guidelines in genetic engineering experiment. The matter grew into such a degree of seriousness that the City Council of Cambridge, Massachusetts imposed a six month ban on genetic engineering research at two leading institutes in USA, the Harvard University and the Massachusetts Institute of Technology. The situation however improved considerably in the subsequent days, thanks to the acquirement of increasing scientific confidence that the research is not as dangerous as some of its critics had suggested. Also, the actual laboratory use of the techniques of Recombinant DNA have largely dissipated the visions of new forms of life created by the insertion of fresh foreign DNA. The Recombinant DNA research thus received a new boost on January 2, 1979 when the guidelines and restrictions were relaxed significantly in the United States.

8. Possible impact on society

If we do not lose faith in mankind and can forget or take care of the possible ill-use or motivated use of the results of genetic engineering and related technologies, we can reasonably imagine of the days ahead when we shall have at our disposal methods of modifying our physical and mental constitution, of correcting errors of nature, of eliminating hereditary faults etc. leading to the evolution of a so-called superhuman race. But man does not live with science alone, he lives in societies and is guided by social and ethical values. Will there by any serious conflict between the scientific achievements and the social and ethical values of the human race ? Our social and ethical values will probably undergo a sea-change by then. Will science devour the human civilization ? Only Future can answer this. But for the time, let us recall the humorous forecast made by Jean Rostand while depicting the anguish of a 'Future Man', modified and sophisticated by varied and repeated treatments, most of which are still- and this must be emphasized imaginary : "I was born of stock chosen and irradiated by neutrons ; they chose my sex and I was born of a mother who was not mine ; during my development, I have been given injections of hormones and DNA ; I have benefited from treatment which superactivated the cortex of my brain ; after my birth, tissue grafts were made to improve my intellectual development and even now I submit myself every year to a course of maintenance treatment in order to keep my mind in good shape and my instincts at an optimum level. I have no reason to be discontented with my body, my sex, or my life, but who am I ?"

References

Cohen S N 1975 *Sci. American* **233** 24
Gilbert W and Taunton-Rigby A 1984 *Research and Development* **26** 176
Glover D M 1980 *Genetic Engineering* (London : Chapman and Hall)
Perbal B 1984 *A practical guide to molecular cloning* (New York : John Wiley)
Roberts R J 1980 *Nucleic Acid Res.* **8** 63
Sinsheimer R L 1977 *Ann. Rev. Biochem.* **46** 415