Molecular biology, nucleic acids, and the future of medicine

E. L. Tatum

Rockefeller University, New York, New York, USA

This paper was presented at the dedication symposium of the new research laboratories of Merck Sharp and Dohme in New York, May, 1966. Columbia University College of Physicians and Surgeons was a joint sponsor. All the papers presented will be published in book form by Merck & Co.

Cost of publication of this article has been covered by the National Foundation.

It seems particularly fitting, in this symposium, to try to assess the probable and possible impact of molecular biology and its important component, the structure and function of nucleic acids, on the future of medicine. Although I attempt specific predictions with considerable trepidation, not being qualified in either fortune-telling or medicine, I do so in the firm belief that the findings and concepts of molecular biology will play a leading role in the future of medicine. Medicine, after all, primarily involves the application of biological concepts and understanding to the health and welfare of man.

After considering various alternative presentations, I have decided first to point out some of the principal problem areas facing medical science in which molecular biology is most immediately concerned. I will then do my best to predict some developments in these areas over the next ten to twenty years.

Before doing this, however, I feel a real obligation to mention a problem that is peripheral in one sense, but that in another is related to and even supersedes most others in basic importance to man. This is the world problem of population and the means of arriving at and maintaining an effective balance between the population explosion and natural resources. Although this problem is outside the main thread of this symposium, it is actually of primary importance. Unless it can be solved, and a new “dark age” avoided, scientific progress can hardly be maintained, and the application to man of new findings will be defeated by sheer numbers.

However, let us assume that this problem will be solved, and proceed to problem areas in medical science. Some of the most important of these are being discussed by the other symposium speakers, auto-immune diseases by Sir Macfarlane Burnet, brain function by Dr. Schmitt, and degenerative diseases by Sir George Pickering.

Accordingly, I will concentrate primarily on the following areas in which progress seems to me to be particularly directly related to the concepts of molecular biology and dependent on nucleic acid research. These major areas are:

1. viruses and virus diseases;
2. hereditary metabolic defects, enzymatic, and regulatory;
3. developmental, congenital, and structural defects; and
4. cancer.

What developments can be predicted in these general areas during the next twenty years or so?

In the field of viruses and viral diseases, it can be anticipated fairly confidently that the study of viruses — bacterial, plant, and animal — will continue to hold as important a place in molecular biology and genetics as has been true during the past decade or so. Most, if not all, viral diseases will be conquered either through immunological means or by the design and synthesis of specific antiviral chemicals. With this, and with definitive understanding of the roles of viruses in human problems involving development and with regulation of cell growth, will come effective prevention and hence control of human problems attributable to viral disease. Finally, it can be anticipated that viruses will be effectively used for man’s benefit, in theoretical studies in somatic-cell genetics and possibly in genetic therapy.

In the area of metabolic disorders, the recognition of the genetic basis of many more disorders can be anticipated. The specific enzymatic defects will be identified in many more instances, as already has been done for PKU (phenylketonuria), galactosemia, certain amino acidurias, and abnormalities in hemoglobin and other serum proteins. Rapid, simple, and sensitive methods for the detection of carriers and for the early diagnosis of affected in-
individuals will be developed, thus facilitating both more effective
eugenic measures and more effective therapy by dietary and other
means. We can even be somewhat optimistic on the long-range
possibility of therapy by the isolation or design, synthesis, and
introduction of new genes into defective cells of particular organs.

In the field of developmental biology, as the consequence of a
better understanding of the molecular and spatial-temporal se-
quences involved in differentiation and development, we can foresee
the effective prevention and alleviation of developmental
errors, such as congenital malformations, whether these be due
to genetic defects or to faulty gene expression or regulation or
are indirectly related to gene activity via hormone production or
target-organ receptivity.

Perhaps in no area is the foreseeable rapid accumulation of basic
information more pertinent to molecular biological research and
concepts, or more promising for the future, than in the area of
neoplasia—cancer. I feel that we can reasonably anticipate that the
basic causes of many, if not all, forms of cancer will be established
within the next few decades. All suspected causes, viral, mutational,
or regulatory factors, center on cell genetics, and on nucleic acid
structure and function. Hence, we can be reasonably optimistic of
the development, first, of effective preventive measures and, later,
of curative therapy. These will come by epidemiological, immuno-
logical, and chemotherapeutic means, by modification and regula-
tion of gene activities, or by means of gene repair or replacement.

Let us now explore the basis for my optimism, which rests, first,
on the general validity of the concepts of molecular biology and
the significance of nucleic acid structure and function therein and,
second, on the „state of the art“ in nucleic acid research and on
the exponential rate at which knowledge in this and other areas of
biology is increasing.

It is now generally accepted that the basic unit of heredity in all
forms of life is DNA, deoxyribonucleic acid, consisting of the
now familiar complementary stranded double helix of Watson
and Crick. This structure uniquely possesses all the qualifications
essential for genetic function: specificity through purine and py-
rimidine base sequence; mutation through alterations in base se-
quence; replication by the enzymatic assembly of new strands
on the two parental strands as templates; and translation into
cell function by transcription of one strand into a complemen-
tary strand of messenger RNA (ribonucleic acid). The messenger
RNA is transferred from the nucleus to the cytoplasm and there
directs the assembly of amino acids into enzymes and other pro-
teins in the ribosomes, with the sequence of base triplets in the
gene thus specifying a particular sequence of amino acids in the
protein product.

The past several years have seen the fleshing out of these bare
conceptual bones in considerable detail and with some remark-
able observations and phenomena.

Genetic material is frequently, or even perhaps usually, present
in the form of circular molecules, such as in bacterial „chromo-
somes“ in bacteriophages, in polyoma virus, and perhaps even in
chromosomes of higher forms. It would appear that this circular-
ity has a control function in the replication process, which, at least
in bacteria, appears to start at a particular location and proceed
in one direction at a single growing point which moves along the
chromosome until all the genes have been copied and the process
is completed [1]. The circularity may also serve in part to protect
dNA from enzymatic attack on free ends of the molecule.

The evidence suggests that a double helical form of nucleic
acid is essential for its replication. For example, replication
of a single-stranded phage DNA (ΦX174), of single-stranded
RNA of the phage f2, and of several plant and animal virus-
eses has been shown to involve the enzymatic formation of a
double-stranded replicative form. Only one of these (-strand),
the complement to the parental strand, would then serve as
template for the synthesis of new viral RNA (+strand). An ap-
parent exception to the requirement for replication of a double-
stranded structure, the in vitro replication of a biologically ac-
tive bacterial virus RNA, was recently reported by Spiegelman
and co-workers [2]. However, later experiments by Weissman
and Feix [3] strongly suggest the presence of a double-stranded
replicating form of RNA in this system as well.

Strandedness would appear also to be important in repair of dam-
aged DNA, as with ultraviolet radiation. It should be recalled that
ultraviolet-produced thymine dimers are split in photoreactivation
but are excised and the strand repaired in a dark reaction. Bacte-
rial mutants deficient in the excision reaction are known, and the
ultraviolet-damaged double-stranded replicative form of ΦX174
is capable of dark repair, whereas the single-stranded form of the
bacteriophage DNA is not. Incidentally, it might also be pointed
out here that exonuclease III may well be involved in the excision
reaction and therefore might be absent in the „excision negative“
bacterial mutants just mentioned.

Strandedness and circularity seem also to be important in deter-
mining the template specificity in RNA synthesis on either DNA or
RNA templates. In vitro, cellular RNA polymerase can use either
dNA or RNA templates, either single or double stranded. With
double-stranded DNA, both strands usually serve as templates.
However, in in vivo m-RNA synthesis and in vivo synthesis of
viral RNA, only one of the double strands of the nucleic acid
appears to serve as template. This also seems to be true in vitro in
several other instances involving transcription to RNA of double-
stranded nucleic acids related to bacterial viruses. That circularity
may be involved in determining single-strand transcription is sug-
gested by the experiments of Hayashi, Hayashi, and Spiegelman
[4] with ΦX174 replicating form DNA. In vitro, only one strand
of the circular form was transcribed, but both strands of the open
linear form were so used.

In connection with the structure of DNA and its transcription,
an important field of investigation was initiated a few years ago
with the discovery by Reich and co-workers that actinomycin D
primarily acts by binding specifically to the amino group of gua-
nine in the minor groove of double-stranded DNA. In so doing,
prevents the DNA from functioning as a template for cellular
RNA synthesis but does not affect RNA-dependent synthesis of
viral RNA [5]. Other studies [6] have shown that chromomycin
and related antibiotics react similarly and that ethidium bromide
and daunomycin bind to DNA and inhibit its function non-spezif-
ically, perhaps, like proflavin, by intercalation between adjacent
base pairs. Bhuyan and Smith [7], in similar studies, have shown
that nogalo-mycin binds to DNA, probably to either adenine or
thymine residues. In contrast, another antibiotic, tubercidin, an analogue of adenosine, apparently is incorporated into both RNA and DNA [8], as are some other base analogues. Such studies in general not only provide an understanding of the molecular basis of activity of these antibiotics but lead to the recognition and use of valuable tools in investigating the structure and functioning of nucleic acids.

The second step of gene transcription involves not only m-RNA but two other classes of RNA, both of which are gene determined. Transfer or soluble RNA (s-RNA) attaches activated amino acids at its CCA terminal end and transfers them to the growing polypeptide chain in the ribosome, which contains at least two ribosomal structures containing RNA (18S and 23S). Each amino acid is carried by at least one specific s-RNA. Some evidence suggests that the DNA loci responsible for ribosomal and transfer RNA are bunched and may be transcribed more or less as a group, whereas the loci for m-RNA are more widely and randomly distributed [9].

Each s-RNA has two separate recognition sites, one specific for its amino acid and one specific for the corresponding base triplet or codon on the m-RNA. The first complete base-sequence analysis was reported just this last year by Holley and collaborators for yeast alanine s-RNA [10]. When sequences of other s-RNA molecules are established, as can be anticipated fairly soon, it should be possible to define structurally both recognition sites, the anticodon, and the amino acid recognition site.

The final step in protein synthesis, the assembly of polypeptide chains, takes place in the ribosomes. These consist of two different-sized subunits and function most effectively as aggregates or polysomes, with each ribosome involved in transcription of a different section of the m-RNA molecule that holds the ribosomes together in the aggregate. In some instances at least, a single polycistronic m-RNA molecule may code for several polypeptide chains.

Protein synthesis on the ribosome is subject to regulation and control in a number of ways. One of the important recent questions has been the punctuation in the genetic code. Are there start and stop signals? It seems that there are! The start signal for most Escherichia coli proteins appears to be a particular nucleotide sequence which specifies at least N-formyl-methionine, and perhaps even N-formyl-methionyl-alanyl-serine. After completion and release, one or more of the N-terminal groups is enzymatically removed. The evidence for this process comes from the elegant work of Webster, Engelhart, and Zinder [11] on the in vitro synthesis of f2 virus coat protein and from related findings of Adams and Capecci [12].

That there are also stop signals seems equally probable. These may function in a manner analogous to the termination of peptide synthesis by puromycin, which is added to the growing carboxyl end of the chain, stopping further additions and causing release of the incomplete chain from the ribosome. At least two possible nucleotide triplets (UAG and UAA), which seem not to code for any amino acid, have been suggested by Sarabhai, Stretton, Brenner, and Bolle [13] as giving a signal for peptide chain termination in some strains of E. coli.

Ribosomal protein synthesis in vitro can also be modified by conditions or substances that bind to ribosomes, disassociate their subunits, or otherwise affect their structure or binding of s-RNA. Examples are basic compounds, such as spermidine, and the antibiotics lincomycin and, probably, chloramphenicol. Interestingly, organic solvents such as alcohol and certain salts may change the specificity of recognition between m-RNA codon and the s-RNA anticodon. Related to this is the finding that suppressor mutations may involve alterations in the recognition site of a particular s-RNA and thus restore normal transcription. Another very important related finding is the discovery by Davies, Gilbert, and Gorini [14] that streptomycin alters the specificity of m-RNA codon recognition by s-RNA, and hence disarranges normal transcription, but can function as a „suppressor substance“ in correcting faulty transcription due to gene mutation.

In view of the widespread interest in and knowledge of the intricacies of the triplet genetic code, it should here suffice to point out that, as the result of the brilliant pioneering work of Nirenberg, Ochoa, and Khorana, and their collaborators, we now have an almost complete key to the triplet codons for all amino acids. This accomplishment has required particularly the techniques of nucleic acid chemistry and biochemistry, in producing the necessary oligonucleotides of defined sequence and length. I would only remind you, in addition, that the universality of the code has now been fairly convincingly established by work with viruses, bacteria, plants, and animals. This is particularly pertinent to our later consideration of genetic engineering.

Another area of nucleic acid research that should be mentioned here is that of mutation, which in its simplest form represents the substitution of one base for another in a DNA triplet. We already know that certain substitutions are more frequent than others and can selectively be made still more frequent by the incorporation into DNA of base analogues which alter base pairing specificities or, under more natural conditions, by the presence of a „mutator“ gene. These observations, together with knowledge of the code, make the prospects of directed mutation somewhat more hopeful for the future.

We perhaps should here also remind ourselves that genes and gene functions in living organisms are not isolated entities and phenomena in a test tube but are subject to regulation or control mechanisms which turn them on, or turn them off, either separately or in operon groups, in feedback response to repressor or activator molecules. Developmental geneticists believe that gene activation and repression are of primary significance in processes of differentiation and development. The more we learn about these repressors and activators and how they work, the more optimistic we can be about the prospects of controlling and correcting faulty developmental processes.

Finally, in this general discussion, both for aesthetic completeness and because of its possible role in infective processes and in development, I want to mention extrachromosomal inheritance. The last few years have seen a considerable clarification of the physical basis of extrachromosomal genetic phenomena. Cell organelles such as mitochondria and plastids, and entities such as infective bacterial episomes, replicate and divide independently of the nucleus, can mutate, and control certain typical characteristics of the cells in which they exist. These characteristics are inherited in a non-Mendelian pattern, often completely
maternally. It is intellectually satisfying that all such entities investigated have been found to contain double-stranded, helical, high-molecular-weight DNA. It is particularly gratifying to me that it is now known that Neurospora mitochondria grow and divide [15], that they contain DNA [16], and that a cytoplasmic character in Neurospora is transmitted from cell to cell by pure isolated mitochondria, as shown in our laboratories [17]. Drs. Luck and Reich have most recently [18] produced evidence consistent with the semiconservative replication of Neurospora mitochondrial DNA, as is true for bacterial plant, animal, and some viral DNA. They have also shown, as followed by a species-specific density difference, that this mitochondrial DNA is inherited maternally. A role of mitochondrial DNA in maternally inherited respiratory-deficient mutants of Neurospora has very recently been suggested by Woodward and Munkres [19] to involve the control of mitochondrial structural protein. It is of considerable interest that chemotherapy is of potential value in controlling the replication of extrachromosomal DNA, as suggested by the effects of agents such as proflavin, streptomycin, and nitrosourea in micro-organisms.

Virus infection introduces new genetic material and interferes with the functioning of host genes. Mutation changes genes, generally for the worse, as in inborn metabolic diseases characterized by single enzyme defects.

Developmental defects may be due to mutant genes, as for club-foot, harelip, etc.; to virus infection, as with Rubella or probably adenovirus; to drugs such as actinomycin or thalidomide; or to prenatal environmental factors such as a riboflavin deficiency. Some of these environmental factors probably act by changing the spatial-temporal sequence of gene activation and expression necessary for normal development. Certain developmental defects may be due to an extra chromosome as in mongolism or Turner’s and Klinefelter’s syndromes. These chromosomal effects appear to be due to defective gene expression as the consequence of chromosome imbalance. It is of considerable interest that some recent data on mongolism in Australia support the possibility that some unknown infective agent may affect oogenesis, causing the non-disjunction that leads to the extra chromosome 21. If this proves to be true, the significance for preventive therapy is obvious.

Certain heritable developmental abnormalities are accompanied by and even attributable to hormone deficiencies, as for thyroxin in cretinism or growth hormone in a type of dwarfism. The basis of hormone activity itself is still unsettled. However, evidence has been accumulating recently that many, particularly the sex hormones and the insect hormone ecdy-sone, act by regulating gene activity in the target cells, leading to the production of specific m-RNA, as shown for animals by the recent experiments of Kidson and Kirby [20]. It should be pointed out, however, that other modes of action have not been ruled out, for instance, that cell-membrane permeability is involved.

The last area I have emphasized, that of neoplasia or cancer, is at the same time one of the greatest problems facing medicine and one of the most complex. In an intellectual sense this complexity, centering on the variety of phenomena implicated in the causation and manifestations of cancer, is reason for optimism, since it provides the opportunity of understanding the process of careogenesis, „spontaneous,” viral, or chemically or physically induced. All of these would appear basically to involve genetic material. Similarly, there is cause for optimism in understanding and treating the manifestations, since they would appear to involve various steps in gene expression.

This general thesis is supported by the increasing evidence that many forms of cancer in animals are indeed due to particular viruses in a phenomenon basically analogous to lysogenesis in bacteria. Most simply it would appear that viral genes so introduced are integrated into the host-cell genome, transforming them into tumor cells, with characteristic new properties such as loss of contact inhibition, change of morphology and growth requirements, and new tumor or viral antigens. We are coming closer to the clarification of the role of viruses in human cancer through detailed studies of the “transformation” process in cultured cells, induced by human and annual viruses; through studies showing the need for a “helper” virus acting with a tumor virus in viral careogenesis in animals; and through epidemiological and other studies in human cancer such as malignant lymphoma and leukemia.

The possibilities of prevention and therapy would seem to be many and diverse, ranging from prevention of infection, replication, expression, and integration of viral genetic material, perhaps as with antiviral agents such as HBB, guanidine, and IUDR, to specific chemotherapeutic suppression of the transformed tumor cells, or even their reversion to normality. Effective attacks on these fronts will lean heavily on our knowledge of the molecular events involved, that is, on knowledge in molecular and nucleic acid biology.

Many, including myself, have talked and written about the application of the newer knowledge of molecular genetics and biology to the improvement of man’s life, heritage, and health in terms of engineering. In these terms, eugenic engineering operates at the level of existing genes and involves purposeful, conscious effort to decrease the prevalence and expression of undesirable genes. These efforts will be effective in proportion to the numbers of detrimental genes that can be identified; to the development of effective methods for their detection in the hidden, carrier state; and, most important, to the general acceptance by individuals of their social responsibility not to perpetuate these genes.

Eugenic engineering operates at the level of gene expression on the phenotype of the individual. It is already widely used in medicine, if not so recognized, as in the administration of vitamins, hormones such as insulin, and thyroxine. Better-recognized examples include the prevention of the harmful accumulation of toxic materials associated with genetic defects by dietary restrictions.
such as of phenylalanine in PKU, of galactose in galactosemia, or of certain branched amino acids in maple-sugar urine disease or the recently described isovaleric acidemia. Such diseases are of special interest and significance in that they characteristically involve brain function and mental retardation. This, and the known actions of drugs such as LSD, forecast the eventual understanding of the organic basis of other mental illnesses, such as schizophrenia, and their successful treatment.

Another much less developed type of eugenic engineering would make use of the concepts of gene regulation and control, by way of feedback regulation by the administration of compounds yet to be discovered. For example, the activities of harmful dominant genes in theory could be repressed as desired, or inactive genes could be turned back on or derepressed as needed, even in utero at critical periods of development. As already pointed out, hormone therapy may actually represent this type of gene regulation.

I would define genetic engineering as the alteration of existing genes in an individual. This could be accomplished by directed mutation or by the replacement of existing genes by others. In principle, and in respect to possible ways of accomplishing this replacement, there are only minor technical differences between genetic engineering as applied to genes in germinal and in somatic cells, although the net result would be considerably different. Precedents for the introduction or transfer of genes from one cell to another exist in microbial systems and are now being tried with mammalian cells in culture. Isolated DNA as such is physically taken up and integrated into the recipient bacterial genome in the transformation process, or in transduction is transferred from one bacterium to another by a virus, followed by integration. If this can be done successfully with animal cells, it will facilitate the development of a mammalian somatic-cell genetics. It will also bring us considerably closer to successful genetic engineering. It is pertinent to point out that, for the phenotypic correction of most genetic errors, only one of the two inactive genes in a diploid cell need be replaced by an active gene and that this may need to be done only in a certain critical number of cells in the particular organ in which the gene function is needed. Hence, it can be suggested that the first successful genetic engineering will be done with the patient’s own cells, for example, liver cells, grown in culture. The desired new gene will be introduced, by directed mutation, from normal cells of another donor by transduction or by direct DNA transfer. The rare cell with the desired change will then be selected, grown into a mass culture, and reimplanted in the patient’s liver. The efficiency of this process and its poten-tialities may be considerably improved by the synthesis of the desired gene according to the specifications of the genetic code and of the enzyme it determines, by in vitro enzymatic replication of this DNA, and by increasing the effectiveness of DNA uptake and integration by the recipient cells, as we learn more about the factors and conditions affecting these processes.

An even more speculative biological possibility for genetic engineering may be suggested, stemming from the finding by Harris [21] that in culture, mammalian cells, even from different species, can be caused to fuse by exposure to an as yet unidentified component of certain animal viruses. These “hybrid” cells can survive and grow, retaining all or part of the chromosome complements of the two “parental” cells. The possible applicability of this approach to the introduction of new genetic material for purposes of genetic engineering seems obvious.

Let me now try to summarize. I have attempted to point out and discuss some of the high spots in molecular biology and nucleic acid research that are particularly promising for the future of medicine. It is apparent that multidisciplinary concepts, approaches, and techniques are essential to the fulfillment of the potentialities which can now be only speculated on. This is particularly apparent if we try to list some of the areas and techniques involved and indicate some needs for further information and development.

One major area involves the design and synthesis of compounds for chemotherapeutic use. In relation to viruses, these will include agents affecting their adsorption, penetration, and replication. In relation to gene function, transcription to m-RNA, and protein, agents already exist and more surely can be designed that will even more specifically affect DNA and RNA structure and particular functions. Other “suppressor substances” can be imagined, which, like streptomycin, will change nucleic acid transcription at the level of protein synthesis and thereby correct genetic errors. Still others, patterned after repressor or regulator substances yet to be isolated and identified, will be able to modulate the activities of specific genes.

With the structures of key enzymes established, the DNA code completely established, and with suitable synthetic methods available, genes can be synthesized to order. The time may soon come when a few molecules of such synthetic genes, or of genes isolated in pure state from nature, will be replicated enzymatically in vitro. With more complete knowledge of the biological processes and techniques involved in DNA uptake and integration, these DNA’s can be incorporated into chromosomes. Genetic engineering will then be just around the corner.

Such speculations as these may be considered by some as too idle daydreaming for a serious symposium. Yet the phenomena of molecular biology which are now almost taken for granted were not even dreamed of a very few years ago! So, to paraphrase,

The time has come, it may be said,
To dream of many things;
Of genes—and life—and human cells—
Of Medicine—and kings—

References


