Somatic mutations and proteins*

BY O. SMITHIES

Department of Medical Genetics, University of Wisconsin, Madison, Wisconsin, U.S.A.

A large body of data has now been accumulated concerning the ways in which heritable mutations can affect protein structure. Such mutations are usually presumed to arise in the germ line, and are studied by progeny analyses of various types. Today, I would like to consider how somatic mutations might be genetically predisposed and thereby made available to an organism during its adaptation to environmental hazards or during its development. The highly specialized immune system of higher animals can profitably be discussed from this viewpoint, and I shall describe a possible molecular mechanism (Smithies 1965) for implementing the immune response, and then consider very briefly its relevance to other adaptive and developmental processes.

THE IMMUNE RESPONSE

The essential questions which are posed by the immune system can be stated simply:

(i) How can many different antibodies be specified by an initially fixed genetic endowment?

(ii) How can selection be made against anti-self antibodies and for anti-foreign antibodies?

(iii) How can an antigen instruct the organism to respond more effectively to repeated encounters with the same foreign substance?

(i) Antibody variability could be the consequence of the organism having a relatively large number of related but not identical genes in the initial zygote from which to select combinations forming useful antibodies. This possibility is unattractive for reasons of economy, and such a system would probably be genetically unstable. I have consequently preferred to explore one of many alternative possibilities, namely that there are a limited number of genetic loci determining antibody structure which have evolved so that they can provide many different proteins as a result of genetically predisposed somatic mutations.

Point mutations, due most probably to changes in single base pairs in the corresponding structural genes, are well illustrated by the human haemoglobins (Ingram 1963), bacterial tryptophan synthetase (Yanofsky 1963), and tobacco mosaic virus coat protein (Tsugita & Fraenkel-Conrat 1963). Although there appears to be no known mechanism which would permit base pair replacements to be used for genetically predisposed somatic mutations in complex organisms, such

---

* This work, paper number 1019 of the Laboratory of Genetics, University of Wisconsin, was supported in part by grants from the National Institutes of Health (GM 08217-03) and the National Science Foundation (G 14240).
Somatic mutations and proteins

Joint mutations might nevertheless be observed in different antibodies. If we assume, as an order of magnitude, that there are $10^8$ cells in the immune system of a new-born animal and that the probability of a point mutation occurring at a particular locus during a cell division is $10^{-5}$, then at birth there could be a thousand antibody genes containing point mutations. These mutations would also be expected to arise after birth, so that their *de novo* appearance at any stage in the life of an animal must be compatible with tolerance. If strong somatic selection occurs for favourable (anti-foreign) point mutations and against unfavourable (anti-self) mutations, such mutations might be to some extent useful in the immune response. I hope to show that strong selective forces are indeed possible, but this would still not make point mutations genetically predisposed.

Deletions and additions of base pairs as mutational events in the case of the acridine mutants of the $r_{11}$ locus of T$_1$ bacteriophage are strongly suggested by the elegant work of Crick et al. (1961). However, predisposition to such small deletions and duplications does not appear to be adapted to genetic control, even though in the presence of strong selective pressures they might be observed. Nor does it seem likely that the *de novo* occurrences of large duplications, such as Bar in *Drosophila* or the partial gene duplication $Hp^a$ at the haptoglobin locus in man, could be genetically predisposed. On the other hand, once a duplication is established in the germ line, as the result of a chance event during the evolution of an organism, subsequent mutational events are genetically predisposed to the extent that there is now a permanent and in principle predictable tendency to further changes. Presumptive evidence in favour of this view has been reviewed and the suggestion made that the predisposed events can be represented as the consequence of crossing-over between regions of homology (duplications) at non-equivalent places in the genome (Smithies 1964). Although the examples quoted involve inter-chromosomal crossing-over, similar events have been found to occur *intra-*chromosomally (Laughnan 1961). These phenomena, then, suggest that a genetic predisposition to somatic inter- or intra-chromosomal rearrangements could be obtained by the presence of suitable duplications in the genome.

Chromosomal rearrangements might occur outside or inside the boundaries of genes determining protein structure. The former would probably lead to the synthesis of no new types of polypeptide, but the latter could lead to quite new polypeptides. These would formally be described as due to somatic mutations of the corresponding genes, although in fact only due to their rearrangement. The formation of hybrid genes related to two ‘ancestral’ genes ($ABC$ and $PBQ$) would be possible, e.g.

$$\begin{align*} ABC \times PBQ & \rightarrow ABQ + PBC. \end{align*}$$

Since the two nucleic acid duplications ($B$ in this case) need not have the same relation to the reading frame of the coding units (triplets) in the two ancestral genes, the hybrid genes would not always be recognized from the amino acid sequences of their products as being directly related to their ancestors. Also difficult to recognize, until we know the amino acid code completely, would be
two genes related to each other by an intragenic inversion. Perhaps the two mutant forms of the genetically unstable allele \textit{ad-3} at the \textit{ad-3} locus in \textit{Neurospora} (Barnett & de Serres 1963) are related in this way.

These considerations led me to suggest (Smithies 1963) that variability in antibody structure could be the consequence of somatically occurring rearrangements in the corresponding genes, made possible by suitable intragenic duplications (or inverted duplications). Of course, this does not exclude the possible involvement of other types of genetically predisposed mutations arising by some as yet unknown mechanism.

(ii) \textit{Selection of antibodies}. There must be some means available to the organism for removing genes specifying undesirable antibodies and for activating genes specifying desirable antibodies, and the choice between these alternatives should be explicable without invoking any unspecified ability to distinguish ‘self’ from ‘non-self’. Previous attempts at explaining this have suggested that selection between non-desirable and desirable antibodies occurs at the cellular level, and that time is the factor which determines rejection or retention (Burnet & Fenner 1949). Thus during the period of the establishment of tolerance, cells are presumed to be destroyed when they meet an antigen with which they can react; subsequently, during the period of immune competence, the same types of encounter lead to antibody production. Increasing experimental evidence is accumulating which indicates that time alone is \textit{not} a sufficient factor to decide between tolerance or response. For example, Steinberg & Wilson (1963) have found individuals with antibodies against the Gm factor of their maternal \g-globulin despite the fact that they had been exposed to this \g-globulin \textit{in utero}. Even more striking are the observations of Stone \textit{et al.} (1965) that dizygotic cattle twins with persisting erythrocyte chimerism are not always tolerant to each other’s skin antigens; skin grafts between partners of 21 pairs of chimeric twins were rejected in more than half of the tests. Since time does not appear to be a sufficient factor in determining tolerance versus response, I have attempted to find another explanation, using the following reasoning:

No molecular basis is known for presuming a direct selective interaction between an antigen and the nucleic acid corresponding to an antibody. We must therefore assume that the antibody itself is also involved in the interaction. We can then ask, at what stage in protein synthesis can the antibody and its nucleic acid co-exist? One possibility is at the ribosomal level, and with sufficient \textit{ad hoc} assumptions a mechanism can be devised which would permit the nucleic acid (messenger \textit{RNA}) and the antibody to form a stable intracellular complex until exposed to a reacting antigen. However, further development of this idea was abandoned since the mechanism lacked plausibility, was without parallels in other systems, and failed completely to account for tolerance. This led me to look for other biological systems in which informational nucleic acid and a corresponding protein product are associated. Once the problem is stated in this way another obvious possibility appears: viruses show an association of informational nucleic acid and a corresponding protein product (the virus coat protein). In addition, viruses can be stable both intra- and extracellularly. I therefore considered the possibility...
that 'antibody viruses' might exist in which nucleic acid specifying one particular antibody structure is associated with a coat protein containing the corresponding antibody.

The next step in the reasoning was to ask, how could interaction between the coat protein of an antibody virus lead to the destruction of one type of nucleic acid (specifying anti-self) and to the activation of another (specifying anti-foreign)? Time of interaction as the sole factor, as we have seen, seems to be excluded, but place is not. If we suppose that the interaction between antibody in the antibody virus coat protein and a reacting antigen releases the nucleic acid, then the fate of this informational nucleic acid will be determined by the place of release. If the nucleic acid is released in an environment in which protein synthesis cannot occur it will probably be eventually destroyed, but in a different environment it could function and direct protein synthesis, and perhaps its own replication. For parallels in virus studies, I would refer you to the works of Holland & Hoyer (1962), Hoyle (1962), Abel & Trautner (1964) and Bayreuther & Romig (1964).

For antibody viruses to provide a mechanism for establishing immune competence and ensuring tolerance one must assume that the cells in which antibody variability arises are not the same as the cells which execute antibody synthesis. The antibody viruses, containing the information for all antibody combining-site variability, must be released by one set of cells and exposed to self-antigens in an environment not permitting protein synthesis. Those antibody viruses which escape destruction must be taken up by a second set of effector cells. Interaction of an antigen with a corresponding antibody virus inside the effector cells must then initiate the primary antibody response, including synthesis of excess coat protein (the primary response antibody) and virus replication. I have suggested elsewhere (Smithies 1965) that the thymus may be the source of antibody viruses, although there are reservations to this suggestion which I cannot detail here; other cells, possibly circulating lymphocytes, may be the effector cells.

(iii) The anamnestic response to the second administration of an antigen is faster, more extensive, and frequently involves a different class of antibody from that of the primary response. Thus the organism has an immunological memory, but its 'remembered' (anamnestic) response may be different from the primary response. This type of anamnestic response could be explained if the primary response leads to a controlled somatic mutation in some cells which then divide and, after a second exposure to the antigen, mediate a genomically instructed response. However, if tolerance is to be preserved, no new specificity of the antibody vis-à-vis antigen can be introduced by this mutation. These necessary conditions suggest a strong parallel between the acquisition of immunological memory and virus-mediated transductions. Thus, if the replicating antibody virus can insert into the host cell genome a region of nucleic acid with information for the (pre-screened) antibody specificity of the primary response antibody then the phenomena can be understood. For such a specific transduction to be possible, one must assume that there are receptor genes for the transduced antibody specificity. Such receptor genes, if we are to adhere to the rules I have tried to set out, must have regions of homology with part of the non-variable region of the nucleic acid of the antibody.

(iii) The anamnestic response to the second administration of an antigen is faster, more extensive, and frequently involves a different class of antibody from that of the primary response. Thus the organism has an immunological memory, but its 'remembered' (anamnestic) response may be different from the primary response. This type of anamnestic response could be explained if the primary response leads to a controlled somatic mutation in some cells which then divide and, after a second exposure to the antigen, mediate a genomically instructed response. However, if tolerance is to be preserved, no new specificity of the antibody vis-à-vis antigen can be introduced by this mutation. These necessary conditions suggest a strong parallel between the acquisition of immunological memory and virus-mediated transductions. Thus, if the replicating antibody virus can insert into the host cell genome a region of nucleic acid with information for the (pre-screened) antibody specificity of the primary response antibody then the phenomena can be understood. For such a specific transduction to be possible, one must assume that there are receptor genes for the transduced antibody specificity. Such receptor genes, if we are to adhere to the rules I have tried to set out, must have regions of homology with part of the non-variable region of the nucleic acid of the antibody.

(iii) The anamnestic response to the second administration of an antigen is faster, more extensive, and frequently involves a different class of antibody from that of the primary response. Thus the organism has an immunological memory, but its ‘remembered’ (anamnestic) response may be different from the primary response. This type of anamnestic response could be explained if the primary response leads to a controlled somatic mutation in some cells which then divide and, after a second exposure to the antigen, mediate a genomically instructed response. However, if tolerance is to be preserved, no new specificity of the antibody vis-à-vis antigen can be introduced by this mutation. These necessary conditions suggest a strong parallel between the acquisition of immunological memory and virus-mediated transductions. Thus, if the replicating antibody virus can insert into the host cell genome a region of nucleic acid with information for the (pre-screened) antibody specificity of the primary response antibody then the phenomena can be understood. For such a specific transduction to be possible, one must assume that there are receptor genes for the transduced antibody specificity. Such receptor genes, if we are to adhere to the rules I have tried to set out, must have regions of homology with part of the non-variable region of the nucleic acid of the antibody.
viruses, in order that the transduction may occur specifically and be under genetic control. No \textit{a priori} limit can be set to the number of receptor loci nor need they all be identical. However, the receptor loci (of which the \textit{Gm} locus in man would be an example) must be incomplete before transduction; i.e. they must have no information for the antibody combining site, otherwise tolerance could not be ensured.

The existence of more than one type of receptor locus would follow from the diversity of class of anamnestic antibodies (\(\gamma G\)-globulin, \(\gamma A\)-globulin, tissue-bound antibodies, etc.)—but all loci would have some degree of homology with the invariable portion of the locus determining the primary response antibodies. The greater the length of homology at any one receptor locus, the greater would be the probability that transduction would occur at that particular locus. The relative proportion of cells transduced at the different loci could consequently be genetically controlled by the extent of the homology between the virus nucleic acid and each receptor locus.

The postulated processes can be summarized in a few sentences:

Antibody variability arises by rearrangements of primary response antibody genes in one set of cells. Nucleic acid corresponding to these genes is passed to other cells in the form of antibody viruses. Antibody viruses meeting an antigen before their uptake by effector cells are destroyed. Antibody viruses encountering a reacting antigen in the effector cells initiate virus proliferation and production of primary response antibody (excess coat protein). Some of the cells mediating the primary response are transduced by the viral nucleic acid. During the transduction, nucleic acid specifying the antibody combining site is inserted into receptor genes. The receptor genes are incomplete before the transduction and cannot instruct the synthesis of antibody of any type until completed. The transduced cells divide and can later mediate an anamnestic response with information from their own genomes. The secondary response antibodies have the same combining-site specificity as the corresponding primary antibodies. The non-variable portions of the anamnestic antibodies are controlled by their respective receptor loci.

The primary response is thus assumed to be a virus-mediated reaction, with the information for the structure of the antibody in the nucleic acid of the virus rather than in the effector cell genome. (Whether or not the effector cell divides during this response cannot be stated on \textit{a priori} grounds. The secondary response involving a different class of antibody is mediated by transduced cells with the information for the antibody structure in the genome of the effector cells. Considerable cell division may be associated with establishing and perhaps executing the secondary response. The observation that virus-transformed cells frequently divide rapidly (Temin 1962) and also incorporate into their cell surfaces protein coded by the nucleic acid of the virus (Habel 1962) suggests that a similar situation might exist in the case of antibody virus-transduced cells.

Some genetic linkage relations implied by the suggested mechanism are of interest. To discuss these I must remind you that both primary response and anamnestic antibodies have two types of polypeptide chains, light and heavy. The light chains of anamnestic antibodies may be the same as the light chains of the
primary response antibodies, but their heavy chains differ. Good evidence is accumulating that the antigen-combining specificities of antibodies depend upon both the light and the heavy chains. It is therefore important for the maintenance of tolerance that no combinations of the variable portions of the light and heavy chain genes occur in antibody-producing cells which have not been pre-screened during the initial transfer of antibody viruses to the effector cells. This suggests that the genes for the structures of the light and heavy chains of primary response antibody must be closely linked and that the virus nucleic acid is a single molecule corresponding to both genes (and possibly to others). This in turn suggests that the rearrangements in the light chain and heavy chain genes need not always be unrelated. If the rearrangements can occur across the nucleic acid region specifying the end of one polypeptide and the start of the second, then the same region of nucleic acid may sometimes specify an amino acid sequence on the light chain and at other times a sequence on the heavy chain. The postulated mechanism also implies that the somatic nucleic acid linkage relations may be different from those in the gametic nucleic acid. Thus the Inv light chain locus in man is gametically unlinked to the Gm (receptor) locus which specifies some characteristics of the heavy chain of γG-globulin antibodies. Yet I am suggesting that, in a functional γG-globulin-producing cell, the Inv and Gm loci are linked as a result of the transduction.

NUCLEIC ACID REARRANGEMENTS IN OTHER SYSTEMS

I want now to discuss briefly some of the more general implications of the mechanism postulated here.

There are many parallels between the set of relations described by Jacob & Monod (1961) in their regulator–operator concept and the series of molecular events we have been discussing. I would like to propose for further consideration the following molecular interpretation of the lactose system of Escherichia coli. With suitable changes it appears to be applicable to other systems:

(a) The operon o^+z^+y^+ is incomplete (like a receptor locus) and lacks an initiator for the synthesis of messenger RNA.

(b) The i^+ allele is complete and instructs for the synthesis of a messenger RNA which becomes encapsulated in its own protein (repressed).

(c) Inducers can interact with this protein and release the nucleic acid (de-repression).

(d) The released nucleic acid can ‘prime’ the o^+z^+y^+ operon, as a result of homology between part of the i^+ messenger RNA and the DNA of the o^+ region. The primer messenger RNA is then completed by z^+ and y^+. (This priming behaviour is the only important new assumption needed.)

(e) Operator constitutive mutants can arise by mechanisms which put a functioning messenger RNA initiator in front of the z^+y^+ region so that it requires no primer (small deletions are known to be effective in changing o^+ to o^-; Jacob, Ullman & Monod 1964).

(f) i^- mutants cannot functionally encapsulate their i primer nucleic acid, but do not prevent the protein product(s) of i^+ from doing so.
(g) \(i^e\) primer nucleic acid cannot be released in the presence of the inducer. The \(i^e\) protein can also prevent the induced release of primer corresponding to \(i^+\) (cf. phenotypic mixing).

The several stages of the immune response as considered here involve sequential gene action, with function of the primary response antibody locus a necessary prerequisite of function of the anamnestic antibody loci. Since the anamnestic antibody loci are presumed to be incomplete in the initial zygote, the system also shows differential activation of these loci, and can lead to divergence of cell lines. This can be symbolically expressed as follows: The initial zygote contained the genes \([A, -B, -C, -D]\), but two different transduced cells might contain the genes \([A-B, -C, -D]\) and \([-B, A-C, -D]\) respectively. Since the genes \(-B, -C\) and \(-D\) are incomplete and the products of the two genes \(A-B\) and \(A-C\) differ, this situation would be formally described by the statement that the gene \(A-B\) was activated and the gene \(A-C\) was suppressed in the first cell, and vice versa for the second cell. The implications of these relations with respect to differentiation need no elaboration, although I do not wish to imply that gene rearrangements are the only mechanism involved.

**Generalization**

My thesis has been that interactions between regions of homology in nucleic acids can form the basis of the genetic control of somatic mutations and of gene expression. Sequential gene action, antibody or enzyme induction, differences in the structures of genetically related polypeptides in divergent tissues, and the apparent expression of genes in some tissues and their apparent suppression in others may be explained in terms of such interactions. In suggesting these explanations I have had to postulate (a) the existence of a class of incomplete genes which can be functionally completed by messenger RNA from other genes or truly completed by gene rearrangements or transductions; (b) that nucleic acid taking part in gene completion may be encapsulated in protein corresponding to its own information; (c) that stereo-specific interaction between the encapsulating protein and inducers, antigens, hormones, etc., can control the release of this nucleic acid; (d) that in some cases this nucleic acid can initiate its own replication.

In preparing a paper attempting a synthesis of observations in several fields, I have necessarily considered and have to some extent used ideas previously discussed by others. Particularly I would like to draw attention to papers by Fox, Yoon & Mead (1962) and Jerne, Nordin & Henry (1963), both of which discuss the possibility of self-replicating RNA and of induced gene changes. I would like also to thank my colleagues and students for their helpful criticisms and suggestions during the development of the ideas I have been discussing.
Somatic mutations and proteins

References (Smithies)

Abel, P. & Trautner, T. A. 1964 Z. Vererbungslehre, 95, 66.


Smithies, O. 1965 Science, 149, 151.


