B. Wiberg and B. J. Nist, *The Interpretation of NMR Spectra* (Benjamin, New York, 1962).
2. J. W. Emsley, J. Feeney, L. H. Sutcliffe, *High Resolution Nuclear Magnetic Resonance Spectroscopy* (Pergamon, London, 1965–66).
3. J. A. Pople, W. G. Schneider, H. J. Bernstein, *High-Resolution Nuclear Magnetic Resonance* (McGraw-Hill, New York, 1959).
4. F. Bloch, W. W. Hansen, M. E. Packard, *Phys. Rev.* **69**, 127 (1946); F. Bloch, *Science* **118**, 425 (1953).
5. E. M. Purcell, H. C. Torrey, R. V. Pound, *Phys. Rev.* **69**, 37 (1946); E. M. Purcell, *Science* **118**, 431 (1953).
6. Varian Associates, Palo Alto, Calif., technical literature.
7. F. A. Nelson and H. E. Weaver, *Science* **146**, 223 (1964).
8. J. D. Roberts, personal communication.
9. A. Loewenstein and T. M. Connor, *Ber. Bunsenges. Physik. Chem.* **67**, 280 (1963).
10. J. A. Pople, *J. Chem. Phys.* **24**, 1111 (1956); C. E. Johnson, Jr., and F. A. Bovey, *ibid.* **29**, 1012 (1958).
11. E. D. Becker and R. B. Bradley, *ibid.* **31**, 1413 (1959); W. S. Caughey and W. S. Koski, *Biochemistry* **1**, 923 (1962).
12. J. S. Griffith and L. E. Orgel, *Trans. Faraday Soc.* **53**, 601 (1957).
13. R. Freeman, G. R. Murray, R. E. Richards, *Proc. Roy. Soc. London Ser. A* **242**, 455 (1957).
14. H. M. McConnell and D. B. Chesnut, *J. Chem. Phys.* **28**, 107 (1958).
15. D. R. Eaton and W. D Phillips, in *Advances in Magnetic Resonance*, J. S. Waugh, Ed. (Academic Press, New York, 1965), vol. 1, pp. 103–48.
16. D. W. Larsen and A. C. Wahl, *J. Chem. Phys.* **41**, 908 (1964).
17. G. N. La Mar, *ibid.*, p. 2992.
18. A. Kowalsky, *Biochemistry* **4**, 2382 (1965).
19. N. Muller and D. E. Pritchard, *J. Chem. Phys.* **31**, 768 (1959).
20. M. Karplus, *ibid.* **30**, 11 (1959); *J. Phys. Chem.* **64**, 1793 (1960).
21. H. J. C. Berendsen, *J. Chem. Phys.* **36**, 3297 (1962).
22. A. Saupe and G. Englert, *Phys. Rev. Letters* **11**, 462 (1963).
23. L. C. Snyder and E. W. Anderson, *J. Chem. Phys.* **42**, 3336 (1965).
24. R. C. Ferguson, *J. Polymer Sci. A* **2**, 4735 (1964).
25. F. A. Bovey and G. V. D. Tiers, *J. Polymer Sci.* **44**, 173 (1960).
26. C. W. Wilson, III, *J. Polymer Sci. A* **1**, 1305 (1963).
27. F. A. Bovey, F. P. Hood, III, E. W. Anderson, L. C. Snyder, *J. Chem. Phys.* **42**, 3900 (1965).
28. C. C. McDonald and W. D. Phillips, in *Proc. Conf. Magnetic Resonance in Biology 2nd* (Pergamon, New York, in press).
29. O. Jardetzky, N. G. Wade, J. J. Fischer, *Nature* **197**, 183 (1963).
30. E. W. Thomas, *Biochem. Biophys. Res. Comm.* **24**, 611 (1966).
31. C. C. McDonald, W. D. Phillips, J. Lazar, *J. Amer. Chem. Soc.*, in press.
32. A. Allerhand and H. S. Gutowsky, *J. Chem. Phys.* **41**, 2115 (1964); H. S. Gutowsky, R. L. Vold, E. J. Wells, *ibid.* **43**, 4107 (1965).

Antibody Variability

Somatic recombination between the elements of "antibody gene pairs" may explain antibody variability.

O. Smithies

The amino acid sequences of polypeptides are uniquely controlled by the DNA of corresponding structural genes. Mutational events in the germ line alter these genes and lead to changes in the structure of polypeptides in later generations. A frequently observed result of mutations is the replacement of single amino acid residues, probably caused by single base-pair changes in the DNA (*1*). Less frequently the mutational events are more complex: a gene may be duplicated and subsequent divergent evolution may cause the descendants of the ancestral gene to code eventually for related but not identical polypeptides (*2*). Once a gene has been duplicated, triplications and higher orders of repetition arise relatively easily by unequal but homologous crossing-over between the duplicated genes (*3*). An unequal crossing-over between tandem duplicated genes can be represented as $2 \times 2 \rightarrow 3 + 1$ which correctly implies that the loss of the duplication is as likely as the reciprocal event, the formation of a triplication. By extending these principles, one sees that multiple copies of related genes can arise but that they may be unstable in the germ line. Despite the potential instability of a multiple gene system, an obvious explanation of the variability of antibodies is that an animal carries in its genome a gene for every type of antibody polypeptide it may need. I will refer to this explanation of antibody variability as the *simple multiple-gene hypothesis*.

Lederberg (*4*) in 1959 proposed an alternative explanation that antibodies might be controlled by a small number of hypermutable genes. These hypermutable genes, transmitted regularly in the germ line, were presumed to undergo somatic mutation during the development of the immune system. I will refer to this as the *somatic mutation hypothesis*. The simplicity of the idea is appealing, and genetic mechanisms might be expected to have evolved to ensure that the needed hypermutability would be maintained and controlled. Lederberg in 1959 had little experimental information on the types of variability actually existing in antibodies, and so did not propose any specific genetic mechanisms for obtaining hypermutability.

Experience with the genetic behavior of a partial gene duplication (Hp^2) in the haptoglobin system (*5*) led me, in 1963, to propose a mechanism that could permit antibody genes to be hypermutable in a genetically controlled way (*6*). I suggested that chromosomal rearrangements, caused by somatic crossing-over between duplicated regions of DNA in antibody genes, would lead to genetically predisposed variability. A specific model was considered requiring inversions within the antibody genes; it was compatible with the limited peptide data available at that time (see *7*). The model was sufficiently specific that it could very easily be tested.

The major source of experimental data available at present on the nature of antibody variability is indirect. Natural antibodies produced by single cells can be detected, but their isolation and characterization is still impossible. Individually uniform monoclonal immunoglobulins are available only from humans (*8*) and mice (*9*) with multiple myeloma—a malignancy frequently leading to the production of large quantities of antibody-like proteins. These myeloma immunoglobulins consist of light and heavy polypeptide chains (*10*) and show serological types (*8*) similar to those observed in natural antibodies (*11*). The myeloma proteins are, however, much more homogenous.

Dr. Smithies is a professor of genetics and medical genetics at the University of Wisconsin, Madison. Portions of this material were presented at the Cold Spring Harbor Symposium on Quantitative Biology, 3 June 1967.

Bence Jones proteins, consisting of unique immunoglobulin light chains, are often excreted into the urine in myelomatosis. Studies of the peptides and amino acid sequences of these proteins by Putnam, his collaborators, and later workers have been vital to the development of present understanding of antibody variability.

The combined data of Titani *et al.* and Hilschmann and Craig (*12*), since confirmed and extended by others, indicated that Bence Jones proteins of a given serological type (there are kappa and lambda types) are almost identical in the COOH-terminal halves of their molecules. The "invariant" COOH-terminal half of kappa light chains is not, however, the same as the "invariant" COOH-terminal half of lambda light chains. Each example of a Bence Jones protein is itself constant, and remains so in mice after serial transplantation of the tumor (*13*), but the proteins differ so much from one another that no two have been found identical in humans. [This may not be true in inbred strains of mice (*14*).] The individual variations in Bence Jones proteins occur in the NH_2-terminal half of the molecule and are not random, as we shall see later. First, the kappa and lambda chains show specific patterns of variability characteristic of their types. Secondly, within a given light-chain type there is also a pattern; some positions in the NH_2-terminal half of the chains are essentially invariant, others show regularities in the observed amino acid variations.

Although the COOH-terminal halves of all human kappa chains from approximately position 107 to 214 are almost identical, there is an inherited difference in the amino acid residue at position 191. This difference correlates (*12, 15, 16*) with the Inv serological types of the proteins; the proteins reacting positively with Inv antiserum "a" have leucine at position 191; proteins reacting positively with Inv antiserum "b" have valine at position 191. The Inv types are inherited in a simple Mendelian fashion; that is to say the genes segregate as single factors in a heterozygous person (Inv^a/Inv^b) so that progeny receive one or the other of the two genes, but not both. This effectively precludes the correctness of the simple multiple-gene hypothesis with a single separate gene for each of the many varieties of light chain. This is because any extensive number of multiple genes would not be expected to continue to segregate as a single Mendelian factor; crossing-over would soon put Inv^a and Inv^b genes on the same chromosome.

A modified multiple-gene hypothesis was accordingly proposed by Dreyer and Bennett (*17*) to accept these facts. They made the *ad hoc* assumption that the constant part of each type of antibody polypeptide is controlled by a single gene that segregates normally, but that the variable regions are coded by about 1000 genes which can later be inserted individually into the DNA in front of the constant gene. The chief difficulty of this hypothesis, apart from its *ad hoc* nature, is the evolutionary unwieldiness of the proposed system. The fact that the kappa and lambda light chains, and presumably the heavy chains also, each show type-specific amino acid sequences implies, on this hypothesis, the parallel evolution of several large groups of discretely maintained multiple genes. This is difficult to imagine. Harder still to explain on any multiple-gene hypothesis is the observation of Doolittle (*18*) that all rabbit kappa chains have a different NH_2-terminal amino acid from those found in human and mouse kappa chains. Complete replacement of even a few identically situated codons in such a large group of genes seems unlikely in the evolutionary time available for this speciation.

Brenner and Milstein (*19*) attempted to account for the variable and constant regions of light chains within the framework of a somatic mutation hypothesis by making the *ad hoc* assumption that a DNA-degrading enzyme attacks the gene somewhere near its midpoint and continues degradation in the direction of the NH_2-terminus. Subsequent errors during repair were invoked to explain the mutations. No attempt was made in their discussion to account for any regularity in the patterns of variability within the NH_2-terminal half of the molecule. The authors recognized the difficulty of proposing any definitive tests of their theory.

Two very recent hypotheses require a limited number of multiple genes with somatic recombination between them (*20, 21*). The numbers of genes are not clearly stated, but Edelman and Gally (*20*) suggest that "perhaps 50 would suffice"; Whitehouse (*21*) implies that a low estimate of the number would be 16. Both these multiple genes plus somatic recombination hypotheses are so permissive that they will be difficult to disprove.

By the end of 1965, sufficient data on the amino acid sequences of Bence Jones proteins had been published for me to test my specific model of a somatic hypermutable gene containing intragenic inversions. There was no clear evidence in its favor (*22*). On the contrary, the available data showed that most but not all of the differences between Bence Jones proteins could be understood as being due to single base-pair changes in the corresponding genes (*23*).

The disproof of the specific intragenic inversion model led me to consider other simple models within the general framework of genetically predisposed chromosomal rearrangements. One alternative was obvious, namely crossing-over between two related but not identical genes. I found it was consistent with the data of Titani, Whitley, and Putnam (*24*) but was apparently not compatible with Milstein's data (*15*). Indeed, Milstein considered the same general model in his paper and rejected it on the grounds that more than two DNA sequences would be needed to explain all of the different amino acids found at some positions in different Bence Jones proteins. I temporarily rejected the model for essentially the same reasons. However, the simplicity of a model with recombination between a pair of antibody genes, and the ease of imagining its evolution by duplication, has caused me to reexamine the idea in the light of newly available data. In analyzing the data, I sought evidence for two fundamental sequences of amino acids from which the observed sequences could be derived. These two sequences would correspond to a pair of germline genes, and the derived "mixtures" would correspond to somatic recombinations between these genes. The two sequences were sought by looking for (i) the occurrence of one of two alternative amino acids at some positions in the proteins, and (ii) the more frequent occurrence of the alternative amino acids in certain linear combinations (linkage groups) than in others.

The analysis of the sequence data, which follows, shows the existence of a general pattern of alternative amino acids, linkage groups, and recombinants and leads me to conclude that a major part, if not all, of antibody variability can be explained by chromosomal rearrangements resulting from somatic recombination between similar but not identical genes in "antibody gene pairs." The model is simple, can

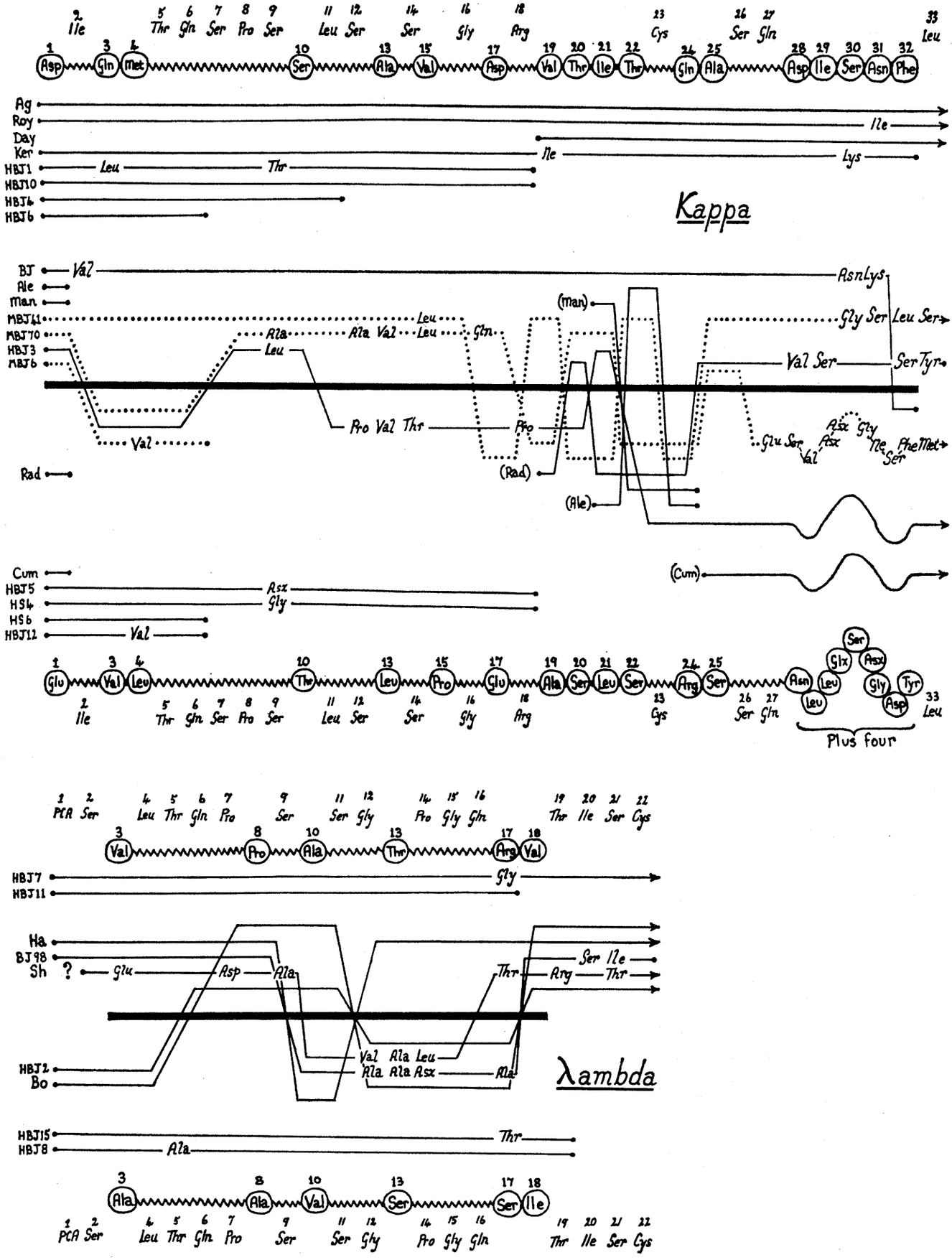

Fig. 1. A summary of available amino acid sequences for 27 human and 3 mouse Bence Jones proteins and myeloma globulin light chains. The sources of the data are listed in reference (*25*). See text for details of the notation used in the diagram to illustrate the patterns of variation. Solid line: human; dotted lines: mouse.

readily be tested, yet has considerable heuristic value as indicated by several new sorts of investigation which it suggests.

Analysis of Amino Acid Sequence Data

Amino acid sequences from a total of 27 human (and three mouse) Bence Jones proteins and myeloma globulin light chains have been published or kindly made available to me prior to publication (*25*). These data are summarized in Fig. 1 which illustrates the patterns of variation in the first 33 positions in kappa chains and the first 22 of lambda, regions for which at least six human proteins have been studied at every position. In some cases, where peptide compositions are known but not amino acid sequences, I have assumed sequences which minimize differences, provided they are not in conflict with any of the data. Glutamic and aspartic acid residues of unknown amide status have also been assigned the status minimizing differences, unless a choice would be prejudicial to subsequent steps in the analysis. The mouse data are included in the figure (the dotted lines) for later comparison, but were not used in its construction.

The figure was constructed by noting the following generalities.

1) Many positions in each of the kappa and lambda polypeptides are essentially invariant. They are numbered, and the relevant amino acids are shown in italics above and below the respective parts of the figure. For example, Thr at position 5 in kappas, Ser at position 2 in lambdas (*26*).

2) At 18 of the first 32 positions in the kappa chains and at 6 of the first 18 lambda positions one of two alternative residues is usually found. The appropriate positions are numbered, and the pairs of alternative amino acids are printed in upright letters in circles above and below the bulk of each part of the figure. Different alternative residues are found in the kappas and lambdas.

3) The alternative residues for each type of light chain can be arranged into two linkage groups, shown by the wavy lines connecting the circles, which link those alternative residues occurring together in the largest number of proteins. For example, Asp, Gln, and Met occur at positions 1, 3, and 4 in six kappa proteins, and Glu, Val, and Leu occur at the same three positions in three other kappa proteins. The heavy black lines across the figure show the regions covered by the linkage groups which can be constructed with the present data. The amino acid at any position is read from the amino acids shown at the top or bottom of the figure, depending on whether the thin line used to indicate a given protein is above or below the heavy black line at that position. For example, the indicating line for protein HBJ3 is below the black line at positions 19 and 20 so that the relevant amino acids, read below, are Ala and Ser; the indicating line for protein Ag at these positions is above the black line, and the amino acids are Val and Thr.

4) *Recombinants* occur; for example, Asp is found with Val and Leu at positions 1, 3, and 4 in the human Bence Jones protein HBJ3. Many of the recombinant proteins show multiple exchanges, that is, the indicating line crosses the heavy black line several times.

5) Once the invariant and alternative residues are taken out of the data, single differences can be recognized which are usually unique to a given protein and occur more or less randomly at any position. They are printed in the bulk of the figure in light italics. For example, Val is an apparently unique single difference occurring at position 2 in BJ; Leu occurs at position 3 in HBJ1 (replacing the expected alternative residue Gln).

6) Positions 9 in the kappas and 17 in the lambdas appear to be hypervariable.

7) Block differences, also printed in light italics, occur in some proteins. The block differences are too large and occur too frequently to be random associations of single differences. The block differences are not randomly distributed; three blocks of three occur at positions 12, 13, and 14 in HBJ3, Sh, and BJ98; BJ and Rad have block differences in the region 28 to 33. All seven block differences occur in the ten recombinant proteins, and none in the 17 nonrecombinants.

General Comments

Before discussing the relevance to antibody variability of the patterns of variation described here, particularly the linkage groups and the recombinants, possible trivial reasons for their observation must be considered. An obvious possibility is that the proteins classed as recombinant and nonrecombinant (proteins for which the respective indicating lines does or does not cross the heavy black line) are the result of some accidental associations of amino acids at the relevant positions which mimic recombination. I cannot directly exclude this, but three arguments make it improbable. First, the general pattern was written down before the data from all proteins were available to me; the sequences of proteins Ale, Man, HBJ3 from position 27 to 33, and of Ha, BJ98, and Bo were obtained later and fitted into the general scheme with no significant changes. Second, many of the recombinant proteins show what looks like multiple exchanges; this would be expected. Although most classical studies with moderately distant genetic markers are most easily interpreted by assuming that recombination is due to a single breakage and reunion event, recent data strongly suggest that many recombinational events lead to what appear to be multiple exchanges when at least three very closely linked markers are examined (*27*). Third, the detection of separate pairs of linkage groups and their respective recombinants in both kappa and lambda type proteins is so striking that any accidental correlation is exceedingly unlikely.

If the described recombinant and nonrecombinant classes are not due to accidental associations, the question arises whether their origin is germinal or somatic. The following arguments make a solely germinal origin unlikely. Polymorphism in the germ line can probably be excluded, since the same general pattern is observed (see the dotted lines in Fig. 1) within a single strain of inbred mice (BALB/c). The classes could not be due to unsuspected differences associated with the Inv system, since no associations of the recombinant and nonrecombinant classes with Inv types are apparent. The nonrecombinants Ag, Roy, Day, Ker, and Cum are respectively b^+, a^+, b^+, a^+, and b^+; whereas the recombinants Man and Rad are b^+ (the other proteins have not been tested). Another explanation of the pairs of nonrecombinants and their recombinants might be that they reflect subclasses of light chains controlled by still more sets of multiple genes. The already considerable evolutionary problems of multiple-gene hypotheses would be further compounded by such a suggestion.

A final question is whether the presence or absence in some proteins of any of the differences described might

be due to events associated with myelomatosis rather than with normal antibody variability. This question is partly answered by the demonstration (*28*) of the amino acids Asp and Glu at position 1 and Gln and Val at 3 in pooled light chains from normal human γ-globulins; these same amino acids are observed as alternative residues in discrete myeloma proteins. This demonstration does not, however, completely exclude myelomatosis as causative of some of the other sorts of variation, or lack of them. For example, might the apparent absence of recombination in some of the human proteins be a feature of myeloma proteins but not of normal antibodies? At least two possibly identical lambda type Bence Jones proteins and two minimally different kappas have been observed in separately induced myelomas in BALB/c mice (*14*), which also suggests that variability in myeloma proteins may not always occur.

Genetic Mechanisms

I now describe an extremely simple genetic model which, with its corollaries, will account for most of the differences revealed by the present analysis. Suppose antibody polypeptide chains of a given type are controlled by an antibody gene pair consisting of a master gene (*29*) and a second scrambler gene, similar but not identical to the master gene. Chromosomal rearrangement by somatic recombination between the two genes could lead to the observed patterns of alternative residues, linkage groups, and recombinants.

I will illustrate the proposed mechanism with kappa type light chains using one of the several possible chromosomal configurations which could provide the physical basis for the chromosomal rearrangement. In this illustrative configuration the master and scrambler genes are depicted as being partial inverted duplications on the same chromosome (Fig. 2). The inverted configuration on one chromosome was selected as requiring the fewest subsidiary conditions for the recombinational event; it does not require pairing of homologous chromosomes or even of sister chromatids; odd or even numbers of complete crossing-over can occur during any number of recombinational events without losing any genetic material and without generating nonsense (gibberish). I stress, however, that these arguments are not compelling and other configurations are admissible. The scrambler gene is depicted as a partial gene duplication, this again being the simplest assumption; the minimum amount of information to account for the structure of a given polypeptide and the observed patterns of variability in one half of it is one and a half genes, but the duplication could be more extensive than "half" a gene.

Figure 2 shows an antibody gene pair with the master and scrambler genes differing in six alternative residues, but identical in the remaining 101 residues of the "variable" half. (A more realistic antibody gene pair differing at 20 positions could generate over a million, 2^{20}, possible, recombinants.) The original chromosome is shown at the top of the figure. The synaptic configuration during the recombination of the elements of the antibody gene pair is in the middle. The rearranged chromosome is shown at the bottom of the figure, with the recombinant master and scrambler genes. Four reciprocal exchanges are depicted, although in reality exchanges are not always found to be reciprocal during recombination (*27*). Note that the recombinant master

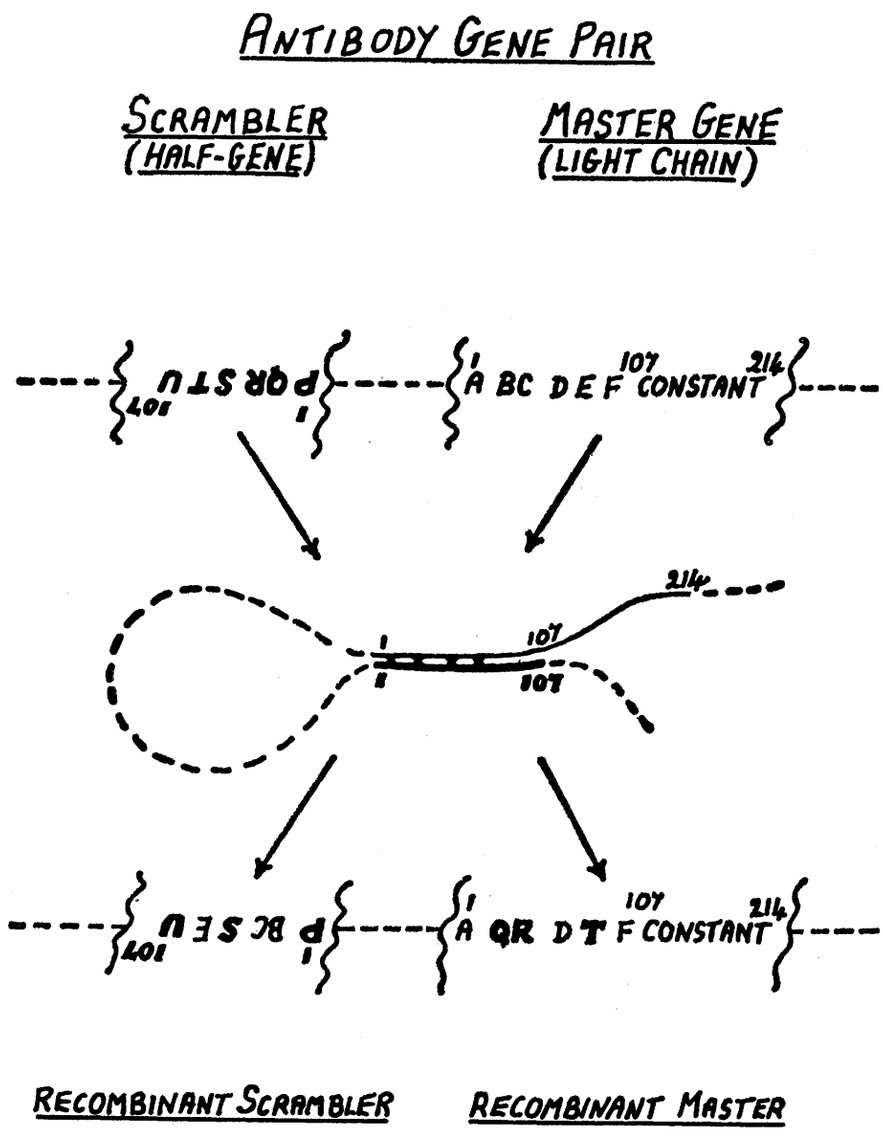

Fig. 2. One of several possible configurations which would permit somatic chromosomal rearrangement to produce a recombinant antibody gene from the elements of an *antibody gene pair*. This particular illustrative example depicts a *master* gene for a light chain and its *scrambler* gene, an inverted duplication of the NH_2-terminal half of the master gene identical to it in 101 places, but differing from it in six places: ABCDEF versus **PQRSTU**. The original chromosome is shown at the top of the figure, the synaptic configuration in the middle, and the rearranged chromosome below. Note the recombinant antibody gene, **AQRDTF** CONSTANT, differing from the nonrecombinant only in the "variable" half of the molecule.

gene differs from the nonrecombinant gene only in the "variable" half. Which of the two linkage groups for the kappa and lambda proteins (Fig. 1) is the master and which is the scrambler cannot be determined at this time.

Corollaries and Extensions

Several corollaries follow from the present general thesis that variability in antibodies arises from recombination between the elements of antibody gene pairs.

First, rapid production of variability in antibody genes should be readily obtained in a developing animal because even a single recombinational event can lead to multiple exchanges (*27*).

Second, the amount of useless or deleterious (directed against self) variability should be minimal, since selection during evolution could have optimized the number and nature of the differences between masters and scramblers.

Third, the evolution of different antibody gene pairs for different purposes is readily imagined; presumably the kappa and lambda antibody gene pairs represent duplicate divergent descendants of some primitive gene pair. Gene pairs can be expected for heavy chains. Whether one scrambler is ever used with several masters, or vice versa, is presently an open question.

Fourth, the particular antibody variability needs of a given species, or even a given racial group, might be expected to be readily met by selection of suitable germ-line mutations in either or both elements of the gene pair. This fourth corollary leads to a possible interpretation of the single and block differences I described above but did not attempt to explain. These sorts of differences are, of course, the exceptions to the recombinational pattern which caused Milstein (*15*) to reject a simple recombinational hypothesis and which made me put it aside temporarily.

If antibody variability is largely due to the mechanism here proposed, then there would be strong evolutionary selection in favor of heterozygosity at loci of antibody gene pairs. Animals heterozygous for any particular scrambler gene, or the relevant part of the master gene, would have a greater potential range of variability than of a homozygote. The advantages to any individual of a greater range of available antibodies in combating environmental microorganisms are obvious. These advantages would lead to a stable polymorphism at loci of antibody gene pairs for the following reasons (*30*). Each new mutant antibody gene pair capable of producing a different range of antibody specificities to which the prevailing microorganisms were not adapted would enjoy a selective advantage, although this advantage would diminish as the mutant became more frequent.

A balance between the rate of occurrence of novel favorable mutants, their selective advantage while still relatively infrequent, and loss of alleles by random drift, would lead to a stable polymorphism with a large number of alleles. (A similar situation has been known for many years in a number of plant species, where polymorphisms with many alleles are found at the loci controlling self-fertilization incompatability.) For these reasons, I suggest that a likely explanation for at least some of the deviations from the general pattern of alternative amino acids is diversity in the germ line. Conceivably polymorphism for block differences might be confined to the scrambler gene and could account for the apparent association of block differences with recombinant proteins.

Future Experimental Work

The suggestion that some of the single and block differences which do not fall into the recombination pattern may be due to germ-line polymorphism is eminently testable within the framework of my hypothesis. For example, the lambda protein Sh is very unusual in lacking the NH_2-terminal PCA (pyrrolidonecarboxylic acid) residue, and in having five single amino acid differences and one block difference which deviate from those expected in a recombinant between the two lambda linkage groups. (The Thr at position 17 could be an intracodon recombinant.) The individual was Japanese! Some of these particular deviations from expectation consequently could well be due to germ-line differences, in which case any children of Sh have a chance of showing the same differences. It is not necessary, within my hypothesis, to wait for two individuals with myelomatosis in successive generations to look for inherited differences in the variable regions of the light chains of single donor normal immunoglobulins. Inherited differences should stand out above the recombination pattern. (The need for keeping adequate records of family and racial origins of proteins under intensive study is apparent.) An immediately obvious way of eliminating many complications caused by diversity in the germ line is to work with myelomas and Bence Jones proteins in inbred BALB/c mice (*9, 13*). Studies in normal identical twin humans would also be valuable and even more reliable.

At this time I have no basis for excluding the possibility that genetic mechanisms other than recombination within an antibody gene pair have become superimposed during evolution on top of the basically simple mechanism described here. Is this the explanation of the hypermutable positions? Quite extensive data may well be required to elaborate all the fine details. Well-founded linkage maps for antibody gene pairs, based on a statistically adequate sample of proteins, will be needed together with a study of their germ-line transmission. Different linkage groups may well be found in different races and even in different families. The consequences of the difference in length of four amino acid residues in the two kappa linkage groups will be extremely interesting: will any recombinants be found indicating mispairing of the two genes in this region? How much configurational difference will be found in recombinant proteins with and without the four residues? Finally, will the basic pattern of my hypothesis (alternative residues, linkage groups, and recombinants) continue to be observed as more examples of kappa and lambda type light chains are studied, and as new classes of immunoglobulins are investigated? This will constitute the major test of the hypothesis.

Summary

I have analyzed the available amino acid sequence data from 30 myelomatosis-derived proteins. Several types of variation are apparent. I conclude that a major and genetically predetermined contribution to the variability of these proteins and of antibodies could be provided by chromosomal rearrangements resulting from somatic recombination between similar but not identical genes in antibody gene pairs. My hypothesis suggests many new types of experiment and can be tested (*31*).

References and Notes

1. V. M. Ingram, *Nature* **180**, 326 (1957); —— in *Genetics*, H. E. Sutton, Ed. (Macy Foundation, New York, 1960), p. 112.
2. ——, *Nature* **189**, 704 (1961).
3. O. Smithies, *Cold Spring Harbor Symp. Quant. Biol.* **29**, 309 (1964).
4. J. Lederberg, *Science* **129**, 1649 (1959).
5. O. Smithies, G. E. Connell, G. H. Dixon, *Nature* **196**, 232 (1962); W. E. Nance and O. Smithies, *ibid.* **198**, 869 (1963).
6. O. Smithies, *Nature* **199**, 1231 (1963).
7. F. W. Putnam, *Biochim. Biophys. Acta* **63**, 539 (1962); ——, S. Migita, C. W. Easley, *Protides Biol. Fluids, Proc. Colloq.* **10**, 93 (1963).
8. F. W. Putnam, *Physiol. Rev.* **37**, 512 (1957).
9. M. Potter, *J. Exp. Med.* **115**, 339 (1962).
10. M. D. Poulik and G. M. Edelman, *Nature* **191**, 1274 (1961); G. M. Edelman and M. D. Poulik, *J. Exp. Med.* **113**, 861 (1961).
11. J. B. Fleischmann, R. H. Pain, R. R. Porter, *Arch. Biochem. Biophys. Suppl.* **1**, 174 (1962).
12. K. Titani, E. Whitley, L. Avogardo, F. W. Putnam, *Science* **149**, 1090 (1965); N. Hilschmann and L. C. Craig, *Proc. Nat. Acad. Sci. U.S.* **53**, 1403 (1965).
13. M. Potter, W. J. Dreyer, E. L. Kuff, K. R. McIntire, *J. Mol. Biol.* **8**, 814 (1964).
14. E. Appella, K. R. McIntire, R. N. Perham, *ibid.*, in press.
15. C. Milstein, *Nature* **209**, 370 (1966).
16. C. Baglioni, A. Carbonara, D. Cioli, L. Alescio-Zonta, *Science* **152**, 1517 (1966).
17. W. J. Dreyer and J. C. Bennett, *Proc. Nat. Acad. Sci. U.S.* **54**, 864 (1965).
18. R. F. Doolittle, *Proc. Nat. Acad. Sci. U.S.* **55**, 1195 (1966); —— and K. H. Astrin, *Science* **156**, 1755 (1967).
19. S. Brenner and C. Milstein, *Nature* **211**, 242 (1966).
20. G. M. Edelman and J. A. Gally, *Proc. Nat. Acad. Sci. U.S.* **57**, 353 (1967).
21. H. L. K. Whitehouse, *Nature*, in press.
22. O. Smithies, in *Regulation of the Antibody Response*, B. Cinader, Ed. (Thomas, Springfield, Ill., in press).
23. D. W. Talmage, *Science* **150**, 1484 (1965); F. W. Putnam and K. Titani, *ibid.*, p. 1485.
24. K. Titani, E. Whitley, F. W. Putnam, *ibid.* **152**, 1513 (1966).
25. The most recent available sources were used for the sequence data. In most instances, earlier references will be found in the literature cited. I am particularly grateful to Drs. L. Hood, W. R. Gray, W. J. Dreyer, and F. W. Putnam for allowing me to use extensive amounts of sequence data from manuscripts in preparation. The designations of the proteins are those of the investigators as follows.
   Ag: K. Titani, M. Wikler, F. W. Putnam, (*24*). At positions 7 and 8 the sequence is Ser•Pro, K. Titani and F. W. Putnam, personal communication.
   Roy, Cum: N. Hilschmann and L. C. Craig, *Proc. Nat. Acad. Sci. U.S.* **53**, 1403 (1965).
   Day: C. Milstein, (*15*).
   Ker, BJ, Rad: C. Milstein, *Proc. Roy. Soc. London Ser. B* **166**, 138 (1966).
   Ale, Man: C. Milstein, personal communication; S. Cohen and C. Milstein, *Nature* **214**, 449 (1967).
   HBJ1, HBJ10, HBJ4, HBJ6, HBJ5, HS4, HS6, HBJ12, MBJ6: L. E. Hood, W. R. Gray, W. J. Dreyer, (*27*); manuscript in preparation (1967).
   HBJ3: W. R. Gray, W. J. Dreyer, L. Hood, manuscript in preparation. (*HBJ3* will be referred to as *Mil.*)
   MBJ4l, MBJ70: W. R. Gray, W. J. Dreyer, L. Hood, *Science* **155**, 465 (1967).
   HBJ7, HBJ11, HBJ2, HBJ15, HBJ8: L. Hood, W. R. Gray, W. J. Drever, *J. Mol. Biol.* **22**, 179 (1966); manuscript in preparation.
   Ha, Bo: F. W. Putnam, T. Shinoda, K. Titani, M. Wikler, *Science*, in press.
   BJ98: C. Baglioni, *Biochem. Biophys. Res. Commun.* **26**, 82 (1967).
   Sh: M. Wikler, K. Titani, T. Shinoda, F. W. Putnam, *J. Biol. Chem.* **242**, 1668 (1967).
26. The abbreviations for the amino acid residues are as follows: Ala, Alanine; Arg, arginine; Asn, asparagine; Asp, aspartic acid; Cys, cysteine; Glu, glutamic acid; Gln, glutamine; Gly, glycine; His, histidine; Ile, isoleucine; Leu, leucine; Lys, lysine; Met, methionine, Phe, phenylalanine; Pro, proline; Ser, serine; Thr, threonine; Try, tryptophan; Tyr, tyrosine; Val, valine; Asx and Glx, aspartic and glutamic acids of unknown amide status.
27. M. Meselson in *Heritage from Mendel*, R. A. Brink, Ed. (Univ. of Wisconsin Press, Madison, 1967), p. 81.
28. L. E. Hood, W. R. Gray, W. J. Dreyer, *Proc. Nat. Acad. Sci. U.S.* **55**, 826 (1966).
29. I have adopted the descriptive word "master" from a related but not identical usage by H. G. Callan and L. Lloyd, *Phil. Trans. Roy. Soc. London Ser. B* **243**, 135 (1960) and H. L. K. Whitehouse (*21*).
30. This argument was suggested to me by Dr. J. F. Crow.

31. This work (paper No. 1134 from the Laboratory of Genetics, University of Wisconsin) was supported by grants NIH (GM08217) and NSF (GB4362). I think my students and colleagues for patient discussions of this work.

Mathematical Theory of Automobile Traffic

Improved understanding and control of traffic flow has become a fast-growing area of scientific research.

Denos C. Gazis

Much has been written about the influence, good and bad, which the automobile has had on our lives. In particular, the subject of traffic congestion has appealed to popular writers, cartoonists, crusaders for saner living, and other serious workers. Some of them have recognized that traffic congestion is not a characteristic of our time only, but may be traced back to Roman times. It certainly existed in major cities near the turn of the century (see Fig. 1), without much contribution from automobiles. However, the emergence and popularity of the automobile has multiplied the problems of congestion and exported them to the suburbs. The centers of the major cities, of course, continue to extract the heaviest toll of delays and frayed nerves. This is illustrated in Fig. 2, which I like particularly because of the unintended caption in the top left-hand corner. The "ecstasy" is, of course, the emotion planned for the users of automobiles by the automobile manufacturers. The "agony" is the true emotion of drivers in New York City, and every other major city.

The investment of money and effort in automobile transportation is nothing less than staggering. In this country alone, the yearly expenditure is of the order of $80 billion, or one tenth of the gross national product. Perhaps it is not surprising that the rapid growth in the use of automobiles has caught us somewhat unprepared to handle the traffic. Compounding the difficulty of understanding and managing so vast an operation has been the fact that it involves not only inanimate objects but also the often unpredictable human being. Under the circumstances, it is not surprising that the management of automobile traffic developed largely as an art, with tools ranging from ingenious to hit-or-miss. Empiricism was of necessity the first approach in management of automobile traffic, and still is the mainstay of traffic engineering. However, in the last 10 years or so, a growing effort has been made to develop a science of vehicular traffic flow and control. The contributions to traffic science have been made by scientists with very diverse backgrounds, each of whom has left the imprint of his discipline on the literature of traffic theory. The growing club of traffic scientists has already convened three times in international symposia (*1–3*) and has contributed to a rapidly growing literature (*4*). In this article I attempt to present some highlights of this work, with emphasis on the more mathematical areas, some

The author is affiliated with the IBM Watson Research Center, Yorktown Heights, New York.