THE ORIGIN OF LIGHT AND HEAVY CHAINS OF IMMUNOGLOBULINS

M.S. KANUNGO

Biochemistry Laboratory, Department of Zoology,
Banaras Hindu University, Varanasi 5, India

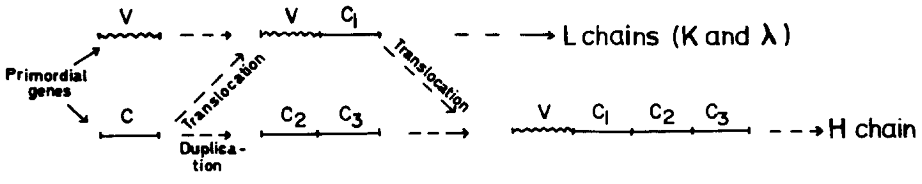
Received February 2, 1971

A model is proposed for the origin of H chains of Ig by the translocation of vc_1 gene for light chains to a c_2c_3 gene giving rise to a $vc_1c_2c_3$ gene. This explains why the common part of H chain is three times longer than the variable part.

Two characteristic features of the amino acid sequences of the light (L) and heavy (H) chains of immunoglobulins (Ig) are that (1) the L chain has a variable (V) region of about 107 residues beginning from the N-terminal,,and a constant (C) region of about 107 residues following the V region up to the C-terminal; (2) the H chain has a V region of 107 residues from the N-terminal, and a C region of 328 residues following it which is three times longer than the C region of the L chain.

Lennox and Cohn (1967) proposed that the germ line carries three genes, κ, λ and v_H, portions of which can be somatically varied. The genes, v_κ and v_λ code for complete L chains, whereas v_H is translocated to c_H to yield a vc cistron for H chains. According to the translocation hypothesis of Edelman and Gally (1967) translocation of two primordial genes, variable (v) and common (c), occurs to give rise to a vc gene which accounts for L and H chains. Smithies (1967) suggested the presence of a 'master' gene for the entire L chain and an inverted 'scrambler' half-gene for the variable part. The germ line (multigenic) hypothesis of Dreyer, Gray and Hood (1967) states that the primordial genes, v and c, are linked to form a common vc gene during somatic differentiation of the Ab-producing cell. This gene is responsible for the synthesis of L chain. A similar mechanism gives rise to the H chain. Different v genes are believed to arise by the normal process of chemical evolution involving gene duplication, mutation and selection. In a recent analysis, Hood and Talmage (1970) have shown that the v and c genes have evolved independently, and that 10,000 genes each for L and H chains may account for 10^7 different Ab molecules. This multigene system could evolve by nonhomologous pairing and recombination after gene duplication.

None of the above hypotheses, however, have any satisfactory explanation for the origin of the H chain which has a common region that is three times longer than the common region of L chains. The following hypothesis attempts to explain the origin of the H chain. (1) Originally there were two primordial genes, v and c, the former having a high frequency of mutation; (2) translocation of v and c genes gave rise to vc_1 genes which code for L chains having 214 amino acid residues; (3) also, duplication of the c gene occurred giving rise to c_2c_3 genes; (4) then translocation of vc_1 to c_2c_3 gene gave rise to $vc_1c_2c_3$ genes which code for H chains having 428 amino acid residues (Fig.).



Independent variations in vc_1 genes gave rise to $vc_{1\kappa}$ and $vc_{1\lambda}$ genes which code for $L\kappa$ and $L\lambda$ chains respectively. Thus the three basic genes present in all cells are $vc_{1\kappa}$, $vc_{1\lambda}$ and $vc_1c_2c_3$. This is corroborated by the finding that C_H1, C_H2 and C_H3 are similar in size, and are homologous to each other and to C_L also (Edelman, 1970). This is also supported by the sequence data for papain from various animals (Weinstein, 1970). The evolution of this protein is believed to have occurred by reduplication of an ancestral gene that coded for a peptide chain having 55-65 residues.

The Ig are unique proteins in that they exist in large varieties within a very limited range of molecular weights. According to the germ line theory (Dreyer, Gray and Hood, 1967) the origin of this diversity is due to the existence of multiple copies of genes. The different ratios of $L\kappa/L\lambda$ light chains in serum Ig (cow, 0.1; man, 2) is a reflection of the ratios of the genes for $L\kappa$ and $L\lambda$ (Weigert, Cesari, Yonkovich and Cohn, 1970). Thus it appears that the multiplication of the genes for $L\kappa$ and $L\lambda$ occurred independently during the evolution of various species, and after these genes arose from the primordial v and c genes.

A point of interest in this connection is the presence of multiple copies of genes for alcohol dehydrogenase (Hart, 1970) and alkaline phosphatase (Brewer, Singh and Sears, 1969) in wheat. The genes for the former are in triplicate and are located in chromosomes 4A, 4B and 4D chromosomes, and the genes for the latter are in duplicate and are in chromosomes 4B and 4D. We postulate that this may be the case for L and H chains of Ig. The location of copies of genes with small variations in different chromosomes may have contributed to independent variation of these genes which has resulted in the variability of Ab molecules, and the differences in the L_κ/L_λ ratios in different species.

References

Brewer, G.T., Singh, C.F. and Sears, E.R., (1969) *Proc. Natl. Acad. Sci.* U.S., *64*, 1224.

Dreyer, W.J., Gray, W.R. and Hood, L. (1967) *Cold Spring Harbor Symp. Quart. Biol.*, *32*, 353,

Edelman, G.M. (1970) *Biochemistry*, *9*, 3197.

Edelman, G.M., and Gally, J.A. (1967) *Proc. Natl. Acad. Sci.* U.S. *57*, 353.

Hart, H.E. (1970) *Proc. Natl. Acad. Sci.* U.S. *66*, 1136.

Hood, L. and Talmage, D.W. (1970) *Science*, *168*, 325.

Lennox, E.S. and Cohn, M. (1967) *Ann. Rev. Biochem.* *36*, 365.

Smithies, O. (1967) *Science*, 157, 267.

Weigert, M.G., Cesari, I.M., Yonkovich, S.J., and Cohn, M. (1970) *Nature*, *228*, 1045.

Weinstein, B. (1970) *Biochem. Biophys. Res. Commun.*, *41*, 441.