Primary Structures of Human Pituitary Growth Hormone and Sheep Pituitary Lactogenic Hormone Compared

Abstract. *The primary structures of human pituitary growth hormone and sheep pituitary lactogenic hormone have been compared with regard to similarities in their sequence. Three homologous segments in each molecule, comprising approximately 45 percent of either peptide chain, have been found. These two proteins have apparently evolved from a common ancestor molecule.*

A comparison of the primary structures (*1*) of ovine pituitary lactogenic hormone (LTH) and human pituitary growth hormone (HGH) is of special interest in view of the overlapping biological activities of these two molecules (*2*) and the possibility that an evolutionary relationship might exist between them.

The validity of establishing such relationships by comparison of amino acid sequences in closely related proteins is now well documented (*3*). There is considerable evidence that the individual members of each species contain within their genome a complex and detailed record of their own genetic origins and histories. At the present time, the most convenient way of "reading" this history is through a comparison of primary structures of the proteins manufactured in strict accordance with this genetic code. An evolutionary relation between proteins is then established by demonstration of a high degree of homology in their primary structures. Such a comparison may be made between proteins that carry out equivalent biological functions in various species or within a single species.

In the human, hormonal control of lactation and growth seems to be effected through a single pituitary protein, namely HGH. In other mammalian species, including sheep, control of these two functions is mediated by two separate molecular entities. We might therefore expect that there would be a demonstrable genetic link between these molecules.

To examine the structures for areas of homology, we aligned the two sequences (*1*) in the LTH and HGH molecules according to the best possible fit of certain reference residues. Cysteine, tryptophan, tyrosine, histidine, and proline were used as references because of their limited content in the proteins and their low relative mutability (*3*). Occasionally, the introduction of a gap into one or the other sequence was required to obtain the best alignment. The homology between any two sequences, or portions thereof, is determined from the number of identical residue pairs plus the number of acceptable mutations. By statistical analysis of a large number of sequences from several types of related proteins, Dayhoff (*3*) has classified residue mutations according to their relative rates of acceptance into proteins. We have designated as "highly acceptable" those mutations Dayhoff lists as having acceptance rates equal to or greater than 40 times that predicted by chance, and those listed as "acceptable" are mutations with acceptance rates from 21 to 39. In our analysis residue pairs representing mutations with acceptance rates from 0 to 20 were not considered homologous.

Thus, where replacements in the sequence can be tolerated, Asp would appear to be a "highly acceptable" replacement for Glu, as would Lys for Arg or Phe for Tyr (*4*). On the other hand, Gly is only an "acceptable" replacement for Glu since Glu is more often replaced by Asp than by Gly. Finally, Glu is almost never replaced by residues such as Tyr, Arg, or Cys,

Table 1. Homology in amino acid sequences of human growth hormone (HGH) and ovine lactogenic hormone (LTH).

Residues encompassed	Identical pairs (No.)	Highly acceptable mutations (pairs) (No.)	Acceptable mutations (pairs) (No.)	Potential residue pairs (No.)
LTH: 156 → 198 HGH: 146 → 188	14	14	7	45
LTH: 58 → 94 HGH: 68 → 90 17 → 29	14	14	5	37
LTH: 22 → 44 HGH: 28 → 50	7	8	1	23
LTH: 101 → 156 HGH: 90 → 146	8	13	8	55
LTH: 1 → 17 HGH: 1 → 17	4	2	3	17

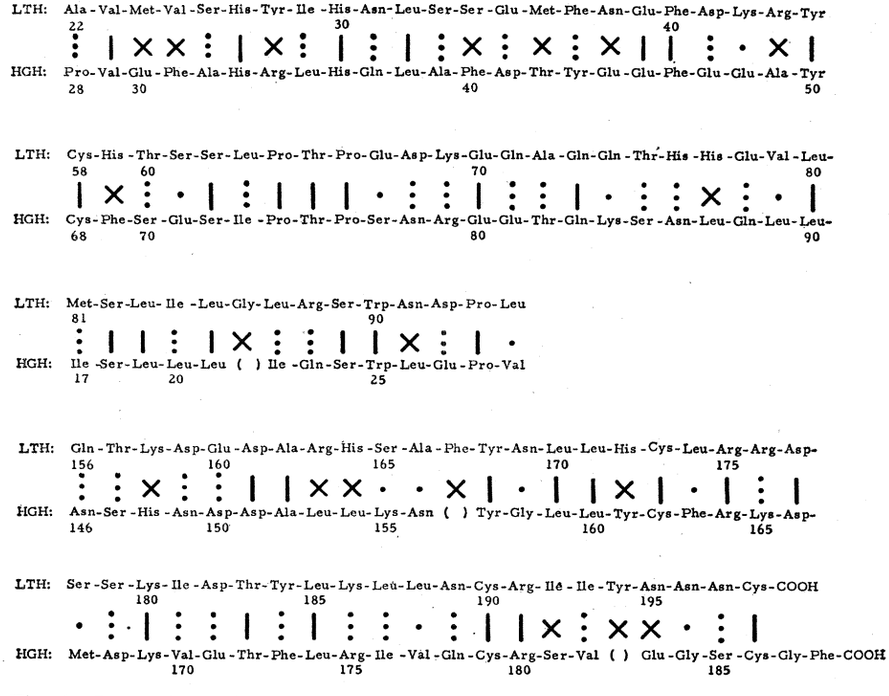

Fig. 1. Comparison of the structures of HGH and LTH. Homology is indicated by identical pairs, vertical bar; highly acceptable pairs, three dots; and acceptable pairs, single dot. Nonhomologous pairs are indicated by a cross.

these being classified as nonhomologous or unacceptable replacements. The acceptability or unacceptability of mutation pairs is quite understandable on the basis of the chemical similarities or dissimilarities existing between the members of each pair.

Although the inclusion of these acceptable or conservative replacements is certainly necessary in any evaluation of homology, it must be kept in mind that the highest emphasis must always be given to the number of identical residues.

In Fig. 1 are shown the segments of the LTH and HGH sequences having clearly demonstrable homology. The longest of these include the segments containing the 43 COOH-terminal residues from these two molecules. In order to align these two segments, gaps must be introduced into the HGH structure. Addition of these gaps results in a total of 45 residue positions. The homology between the segments consists of 14 pairs of identical residues, 14 pairs of highly acceptable mutations, and 7 pairs of acceptable mutations. Thus, 35 of the 45 residue positions show a high degree of homology. It may be significant that in the segments of HGH and LTH shown in Fig. 1 the homology is poorest adjacent to the positions where a gap has been introduced.

Additional regions of good homology are found near the middles of the two sequences. Figure 1 also shows the alignment from Cys-58 of LTH and Cys-68 of HGH. The comparison is quite good up to Leu-80 of LTH and Leu-90 of HGH. At this point a repositioning of the HGH chain is required to continue fitting to the LTH structure. If Ile-17 of HGH is aligned with Met-81 of LTH the homology may be extended for an additional 14 residue pairs. This is an important alignment since it identifies which of the two tryptophan residues in LTH (Trp-90) is homologous with the single tryptophan (Trp-25) in HGH. Final areas of homology are shown in Fig. 1, beginning with Ala-22 of LTH and Pro-28 of HGH and extending for 23 residue pairs.

The remaining two portions of each molecule do not show such a high degree of homology. In particular, the content of identical residues is considerably lower. These last two alignments are tabulated in Table 1.

Thus, three areas of these two hormone molecules appear to be homologous in terms of their primary structure (Table 1). These areas contain a total of 105 residues of which 35 are identical, 36 are highly acceptable replacements, and 13 are acceptable replacements giving a total of 84 homologous positions. This amounts to about 45 percent of either peptide chain. There appears to be sufficient similarity in the sequences of LTH and HGH to confidently postulate their separate evolution from a common ancestor molecule.

Since HGH is active both as a lactogenic and a growth hormone it is highly probable that some portion or portions of the three areas showing the highest degrees of homology contain important structural features intrinsically required for lactogenic activity. Of course, some or all of these same features may also be required for growth hormone activity, and we are not proposing that the two activities reside in mutually exclusive portions of the HGH molecule. It may be that these two biological functions are controlled by a single basic mechanism and that the real difference in the two molecules is in their specificity for receptors.

At the present time the significance of an evolutionary link between these two hormones in terms of their mechanisms of action is unknown. Nevertheless, the acceptance of such a link may provide directions along which useful theories may be made and ultimately tested in the laboratory.

THOMAS A. BEWLEY
CHOH HAO LI
Hormone Research Laboratory, University of California, San Francisco

References and Notes

1. C. H. Li, J. S. Dixon, T-B. Lo, Y. A. Pankov, K. D. Schmidt, *Nature* **224**, 695 (1969); C. H. Li, W-K. Liu, J. S. Dixon, *J. Amer. Chem. Soc.* **88**, 2050 (1966); C. H. Li, J. S. Dixon, W-K. Liu, *Arch. Biochem. Biophys.* **133**, 70 (1969).
2. C. H. Li, *Excerpta Med. Int. Congr. Ser.* **158**, 3 (1968); I. A. Forsyth, *ibid.*, p. 364.
3. M. O. Dayhoff, *Atlas of Protein Sequence and Structure* (National Biomedical Research Foundation, Silver Spring, Maryland 1969), vol. 4.
4. Abbreviations for amino acid residues are as follows. Alanine, Ala; arginine, Arg; aspartic acid, Asp; asparagine, Asn; cysteine, Cys; glutamic acid, Glu; glutamine, Gln; glycine, Gly; histidine, His; isoleucine, Ile; leucine, Leu; lysine, Lys; methionine, Met; phenylalanine, Phe; proline, Pro; serine, Ser; threonine, Thr; tryptophan, Trp; Tyrosine, Tyr; and valine, Val.
5. Supported by grants from the American Cancer Society and by NIH.

23 January 1970; revised 16 April 1970

Histocompatibility-2 (H-2) Polymorphism in Wild Mice

Abstract. *Red blood cells of 40 wild mice captured at four different locations in the Ann Arbor area were typed for the presence of 15 different H-2 antigens by direct hemagglutination test. Fifteen different phenotypes were found and all were different from those determined by known* H-2 *alleles of the laboratory mice. Great similarity between phenotypes of mice from the same location was observed. This is interpreted as further evidence for the deme structure of natural mouse populations.*

The major histocompatibility locus of the house mouse, *Mus musculus*, is *H-2*. Its products are the most potent antigens of tissue incompatibility and, at the same time, antigens of the most complex blood group system of this species. The complexity of the *H-2* locus is best illustrated by the so-called H-2 chart (*1*) which lists about 25 well-defined H-2 antigens in about 20 well-defined combinations (*H-2* alleles), according to their distribution in various inbred strains and stocks. Since most of the strains and stocks of the laboratory mouse were derived from a very small number of sources, the H-2 chart does not give a true picture of the H-2 polymorphism. To get information about the extent of polymorphism in natural populations, one has to study wild mice. In this report preliminary data on such analyses are presented.

The wild mice were trapped in four farms in the Ann Arbor area during the summer of 1969. The farms were 10 km or more apart and 5 km or more beyond the city limits. The mice were captured either in granaries (farms KP, KE, and SA) or in corncribs (farms KE and GA). H-2 antigens were identified with alloantiserums, most of them prepared by immunization of F_1 hybrids between two inbred strains with tissue from genetically different inbred strains. The donor-recipient combinations were chosen in such a way as to produce monospecific or nearly monospecific antiserums (*2*). Prior to their employment in the wild mice study, all the antiserums were analyzed serologically with a panel of red blood cells representing different *H-2* alleles. A total of