RODENT HEMOGLOBIN STRUCTURE:
A COMPARISON OF SEVERAL SPECIES OF MICE *

J. G. Gilman

Department of Human Genetics
University of Michigan Medical School
Ann Arbor, Michigan 48104

Introduction

Recent studies of the structure and genetics of hemoglobin have shown several instances of clusters of genes producing sets of closely related globin polypeptide chains. The first such cluster to be demonstrated was the pair of genes coding for the β and δ chains of humans.[1] Closely linked pairs of α-chain genes have been found in some individuals in Hungary,[2] although in Melanesia there are individuals with only one α-chain locus.[3] There may be as many as four γ-chain loci clustered together in humans.[4] I will call a closely linked pair, such as the genes coding for human β and δ chains, a "doublet," and a single gene a "singlet;" the terminology is based on that of Spofford.[5]
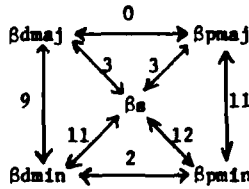
Mice provided one of the first examples of a globin chain doublet in non-humans. In 1962, Hutton and colleagues [6, 7] showed that some inbred strains of mice have an adult β-chain doublet producing two β chains in unequal amounts, which subsequently was called Hbb^d; other strains have an adult β-chain singlet, Hbb^s. Morton [8] discovered a third allele at the β-chain locus of mice, Hbb^p, with the p hemolysate bearing some resemblance to the d hemolysate. The conclusions of Hutton and colleagues remained in some dispute until recently, because there were no actual sequence data confirming that the d hemolysate contains two structurally different β chains. In 1972, I reported that the hypothesis of Hutton and colleagues was correct; I isolated two hemo-globins from the d hemolysate and showed that they possess structurally different β chains, β^{dmin} (produced as about 20% of the total) and β^{dmaj} (accounting for the remaining β-chain production).[9] I also confirmed that Hbb^p is a variant of the Hbb^d allele, since it produces a variant minor β chain, β^{pmin}. Rifkin [10, 11] presented tryptic peptide composition data on β^s and β^{dmaj}, and Popp has se-quenced most of β^s [12] and β^{dmaj}.[13] A summary of my data and those of Rifkin and Popp, showing the differences that are known among the β chains of laboratory strains of mice, is shown in FIGURE 1.

Mice have also been shown to have α-chain doublets. Rifkin and col-leagues [10] showed in 1966 that one inbred strain of mice has two α chains, which differ from each other in at least one position. Another inbred strain that they examined has only one α chain. Von Ehrenstein [14] had earlier found that the rabbits he examined have two α chains; he suggested that perhaps only a single gene was responsible for producing both chains, by means of "ambig-uous translation." Rifkin and colleagues suggested that translational ambiguity

might be operative in the case of mice as well, although they did not rule out the possibility of gene duplication.

In 1968 Hilse and Popp [15] showed that the two α chains found within a given inbred strain of mice differ at three positions. Two differences occur

$$
\begin{array}{ccc}
 & 0 & \\
\beta\text{dmaj} \longleftrightarrow & & \beta\text{pmaj} \\
& \searrow 3 \quad 3 \nearrow & \\
9 \downarrow & \beta\text{s} & \downarrow 11 \\
& 11 \nearrow \quad \searrow 12 & \\
\beta\text{dmin} \longleftrightarrow & 2 & \beta\text{pmin}
\end{array}
$$

	9	13	16	20	22	23	58	73	76	77	80	109	139
βs	Ala	Gly	Gly	Ala	Glu	Val	Ala	Asp	Asn	His	*	Met	Ala
βdmaj	Ala	Cys	Gly	Ser	Glx	Val	Ala	Asp	Asn	His	*	Met	(Thr)
βdmin	Ser	(Cys)	Ala	Pro	Glx	Val	Pro	Glu	Lys	Asn	Asn	Alx	(Thr)
βpmin	Syr	Cys	Ala	Pro	Ala	Ile	Pro	Glx	Lys	Asn	Asn	?	(Thr)

FIGURE 1. A comparison of *Mus musculus musculus* β chains β^s, β^{dmaj}, β^{dmin} and β^{pmin} at the positions where differences were found. No differences were found between β^{pmaj} and β^{dmaj}. Data of Rifkin [10, 11] and of Popp [12, 13] are used for β^s and β^{dmaj}, whereas my data [24] give partial sequence information on β^{dmin}, β^{pmin} and β^{pmaj}, as well as on β^{dmaj} and β^s. My data on β^s do not agree with those of Popp at positions 72 and 80; in this figure I use my data for position 80. The alanine-threonine difference between β^s and β^{dmaj}, in the C-terminal region covered by tryptic peptide βT–14, was demonstrated by Rifkin; [11] Popp has suggested that this difference occurs at position 139. [13] Bonaventura and Riggs [45] first demonstrated that cysteine is at position 13 of β^{dmaj}.

Uncommon symbols are $, Syr, serine or cysteine; Alx, probably alanine, possibly valine; ?, an unknown residue; *, probably serine; **, not arginine; ***, not methionine. Standard abbreviations are: A, Ala, alanine; R, Arg, arginine; N, Asn, asparagine; Γ Asp, aspartic acid; C, Cys, cysteine; E, Glu, glutamic acid; Z, Glx, glutamic acid or glutamine; G, Gly, glycine; H, His, histidine; I, Ile, isoleucine; L, Leu, leucine; K, Lys, lysine; M, Met, methionine; F, Phe, phenylalanine; P, Pro, proline; S, Ser, serine; T, Thr, threonine; W, Trp, tryptophan; Y, Tyr, tyrosine; V, Val, valine. Parentheses around a residue mean that tryptic peptide composition data are all that are available, and that the position in the sequence has been inferred by homology.

within a single tryptic peptide, yet only two of the four possible tryptic peptide variants were found, which is not compatible with any simple scheme of translational ambiguity. They also showed that, in some cases, an inbred strain with only one α chain had one of the α-chain variants possessed by a strain with

two α chains. Hilse and Popp interpreted their data to mean that mice with two α chains have an α-chain doublet, rather than an ambiguously translated singlet. In rabbits, as well, the translational ambiguity hypothesis has not been supported by recent evidence, which suggests that the two α chains found in rabbits are allelic variants.[16] Thus, globin-chain doublets have not been demonstrated in rabbits, although there is a good deal of evidence that doublets occur in mice.

The situation in rat hemoglobin bears a strong resemblance to that in the mouse. Early electrophoretic studies of hemoglobin from inbred and random bred rats showed several hemoglobin components in unequal proportions.[17] Three electrophoretic patterns were found among the various animals. The patterns could be interpreted as two pure patterns and a mixture of the two. In each pure pattern, four components were found. The two pure patterns differed in that two components of one pattern were displaced relative to the homologous components of the other pattern.

Using isoelectric focusing, Stein and colleagues[18] isolated and studied six hemoglobins of the inbred rat hemolysate, and they failed to show a polymorphism. However, results of Drysdale and colleagues[19] suggest the need for caution in interpreting isoelectric focusing data. They studied human and animal hemoglobins. The human study shows that the method is extremely sensitive, resolving hemoglobin components that differ from each other by chemical modification of the globin (N-terminal blockage, for example). They found that the rat hemolysate has eight components, though several of those components are present in very small amounts. Recent data of Garrick[20] suggest that isoelectric focusing data are correctly interpreted as suggesting six hemoglobin components; she found that the inbred rat hemolysate contains two α chains and three β chains.

Brdiča and colleagues[21] have presented partial sequence data on one of the α chains of rat hemoglobin, and Garrick has sequenced the major α chain and most of the minor α.[20] Comparing α-chain sequences of rat, mouse and rabbit, it is clear that substantial divergence has occurred. Rabbit[22] and mouse[23] α chains differ in at least 22 positions, and the data of Brdiča and colleagues and of Garrick, suggest that about the same number of differences would be found upon comparing rabbit and rat or mouse and rat. This contrasts with the relatively smaller number of differences found among variant α chains of mice. I obtained structural data on 79% of the α chain of *M. caroli*.[24] Compared with Popp's sequence of *M. musculus* α chain,[23] *caroli* α was seen to differ in only three positions (see FIGURE 6 below), despite the fact that DNA hybridization data[25] show that the two *Mus* species, *caroli* and *musculus*, diverged ⅓ as long ago as mouse and rat.

The major question with which my present paper tries to deal is whether some function is served by the existence of doublets. There are two limiting possibilities, complete redundancy of the two loci and their absolute importance to the survival of the animal. For example, Boyer and colleagues[26] suggested that the δ chain of humans probably serves no function not also fulfilled by the β chain, because δ is present at a level only a few per cent that of the β chain. This implies that the doublet could be replaced by a singlet producing only the β chain without any disadvantage. An example of the other limiting possibility is provided by the linked embryonic and adult β-chain loci in mice.[27] The embryonic locus serves a function quite important and distinct from that of the adult locus. Both loci are presumably essential to the mouse's survival.

In the evolution of doublets, there are likely to be intermediate stages between the limiting possibilities of redundancy and absolute importance of the two loci. A doublet may confer some selective advantage to an animal without both loci being absolutely essential for its survival. There is some evidence that this is true in the base of the mouse Hbb^d allele. The electrophoretic study by Selander and colleagues[28] of hemoglobins of wild North American mice, showed that the Hbb^d-Hbb^s polymorphism occurs in every mouse population that they had adequately tested. In some populations there was an excess of heterozygotes. These data are consistent with the hypothesis that the Hbb^d-Hbb^s polymorphism is being maintained by balancing selection, which means that one homozygote may be inferior to the other, but that the polymorphism is maintained because the heterozygote is superior to either homozygote. The major difference between the Hbb^d and Hbb^s alleles lies in the production by Hbb^d of a minor chain, β^{dmin}, since there are only three known differences between major chains β^{dmaj} and β^s but nine differences between β^{dmin} and β^{dmaj} and eleven between β^{dmin} and β^s (see FIGURE 1). One is led to ask whether the minor β chains may not have a more marked effect than the major chains in the maintenance of the doublet-singlet polymorphism.

I decided to investigate this question by studying the structures of hemoglobins from mouse species other than the laboratory species, *M. musculus musculus,* which was used in the previous studies. I reasoned that, if β^{dmin} served no function that contributed to the selective advantage of a heterozygous mouse possessing it, then there would be few selective constraints on it and its structure would be rather variable in different mouse species, since even deleterious mutations might be incorporated as substitutions. If I were to find a large degree of variability among minor β chains of the various mouse species being examined, or if the minor β chain were frequently not present in mice, I would tentatively conclude that there is no selective advantage in possessing the minor β chain produced by the Hbb^d locus. If very little variability were found among minor β chains, I would interpret this to mean that it is likely that the minor β chain is selectively important to the mouse.

Materials and Methods

Wild Asian mice of species *M. caroli, M. cervicolor* and *M. musculus castaneus* were originally trapped and characterized in Thailand by Dr. Joe T. Marshall. *M. musculus molossinus* were originally from Dr. Fusanori Hamajima (Kyushu University, Fukuoka, Japan). I obtained *M. musculus molossinus, M. musculus castaneus* and *M. caroli* from Dr. M. Potter (National Institutes of Health, Bethesda, Maryland), and *M. cervicolor* and additional *M. caroli* from Dr. T. C. Hsu (University of Texas, M. D. Anderson Hospital and Tumor Institute, Houston, Texas).

Starch gel electrophoresis of hemoglobins was performed using gels made with "tris-borate-versene" (TBV) buffer[29] but with the ethylenediamine-tetra-acetic acid (versene) concentration increased to .005 M. (Tris concentration was .045 M and boric acid concentration .025 M). Hemoglobin samples were run both with and without prior treatment with cystamine, which couples S-ethylamine groups to the accessible sulfhydryl groups of the protein. The procedure for cystamine treatment was reported by Wegmann and Gilman.[30] Untreated samples, when run on a gel containing treated samples, were subject

to the treatment procedure except that an equivalent amount of sodium chloride was substituted for cystamine. Alkaline urea gels were made with the same concentration of TBV buffer as gels without urea. Urea concentration was 6 M, mercaptoethanol concentration was .05 M, and gels were made using the procedure of Smithies and colleagues[31]

Sequencing was carried out using an Edman-Begg sequenator,[32] with procedures of sequenator sample processing as described by Smithies and colleagues[33] with one modification. Milder sodium hydroxide (NaOH) hydrolysis than they described (for conversion of thiazolinones to amino acids) permitted recovery of asparagine in about 50% yield, whereas the harsher procedures convert it all to aspartic acid. The mild procedure uses .15 N NaOH (instead of .2 N) and one hour hydrolysis (instead of 3.5 hour) in an autoclave of 21 pounds pressure (127° C). Following NaOH or hydriodic acid (HI) hydrolysis, samples were analyzed on a modified Technicon TSM-1 amino acid analyzer, as described by Smithies and colleagues. I also use their computer processing technique to correct for out of step cleavages and background, with one modification: For each amino acid, the same background was subtracted at every position so as to minimize background noise but not lose information.

Unfractionated hemolysate was sometimes run on the sequenator; to obtain sequence data on major or minor hemoglobins separately, hemolysates were fractionated by means of diethylaminoethyl-Sephadex® chromatography using a concentration gradient of ammonium acetate (from .03 M to .12 M) at about pH 9. Heme was removed from the globin before sequencing by the method of Rossi-Fanelli and Antonini.[34]

In order to gain information on the interior of β^{caroli}, unfractionated *caroli* globin was cleaved with cyanogen bromide (CNBr).[35] The solution contained 10 mg./ml. of protein, 40 mg./ml. of CNBr, 7 M guanidium hydrochloride and .33 N HCl, and the reaction was for 18 to 24 hours at room temperature. Salts were removed using a column of Sephadex G-50, which also achieved some fractionation of peptide material.

Results

### Asian *Mus musculus: molossinus* and *castaneus*

Alkaline starch gel electrophoresis suggests that the hemolysates of the Japanese mouse *M. musculus molossinus* have one major and one minor hemoglobin, which are similar to those of the d hemolysate, as shown in FIGURE 2a. The four mice that I tested of the Thai species *M. musculus castaneus* gave hemoglobin patterns, like that shown in FIGURE 2b, similar to that of a 50:50 mixture of p and d hemolysates. Hemolysate from another mouse, analyzed on acrylamide gel by Whitney,[36] showed a pattern like that of the d hemolysate. These data suggest that *M. musculus molossinus* has the Hbb^d allele, whereas *M. musculus castaneus* has both the Hbb^d and Hbb^p alleles segregating in the population. Urea starch gel analysis of the globin chains of these animals (FIGURE 2c) suggests, however, that both *M. musculus molossinus* and *M. musculus castaneus* possess a minor chain that migrates slightly differently from β^{dmin}. β^{pmin} migrates like the α chain at the alkaline pH of this gel,[24] so that this gel result is consistent with the presence in the *castaneus* hemolysate of β^{pmin} even though only one minor band is seen. Therefore, *molossinus* and

FIGURE 2. Starch gel electrophoresis of hemoglobin from Asian mice *Mus musculus molossinus* and *Mus musculus castaneus,* as compared to hemoglobin of laboratory strains of *Mus musculus musculus.* In A, B and C labels refer to hemoglobin components of *Mus musculus musculus* samples.

A. A comparison of hemolysates of *M. musculus molossinus* (left) and *M. musculus musculus* (genotype Hbb^d/Hbb^d). Sample 1: *Molossinus* hemolysate, +.0065 M. dithioerythritol (DTE). Sample 2: d hemolysate, +.0065 M. DTE. Sample 3: d hemolysate, no DTE. The heavy streaking in front of the major band of *molossinus* hemolysate, seen in the d sample only in the absence of DTE, may be due to the fact that the *molossinus* sample was several days old when tested.

B. A comparison of a hemolysate of *M. musculus castaneus* (left) with a 50:50 mixture of d and p hemolysates from *M. musculus musculus.* Both samples are .0065 M. in DTE. Sample 4: *M. musculus castaneus* hemolysate. Sample 5: 50:50 mixture of d and p hemolysates.

C. Alkaline-urea starch gel electrophoresis of globin from *M. musculus castaneus,* d and *M. musculus molossinus.* Sample 6: *M. musculus castaneus* globin. Sample 7: d globin. Sample 8: *M. musculus molssinus* globin.
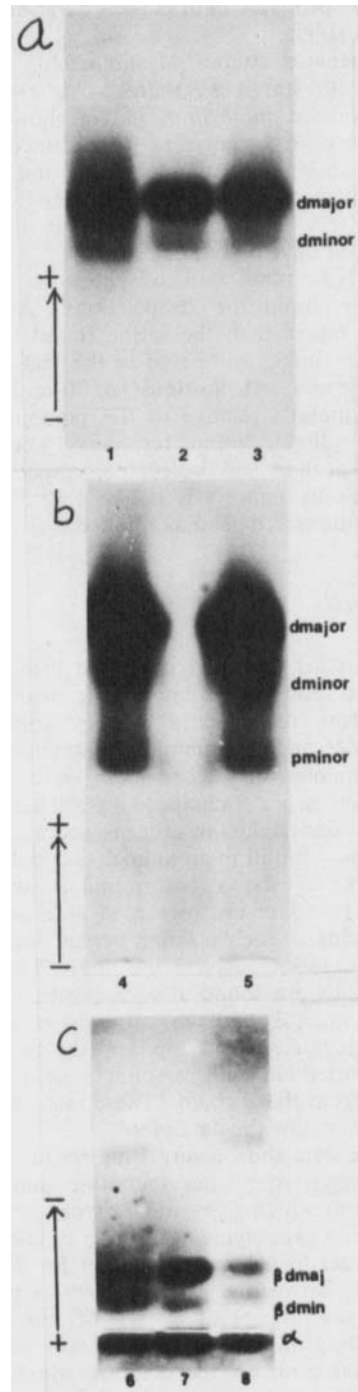
castaneus probably both possess a variant of *Hbb^d*, and *castaneus* also possesses the *Hbb^p* allele.

Sequenator studies of *molossinus* and *castaneus* hemoglobins were consistent with starch gel data.[24] An N-terminal sequence for 26 positions of unfractionated *molossinus* globin showed no differences compared to Popp's sequence of d hemolysate.[13] Sequence data were also obtained on 57% of purified *molossinus* minor β chain, and no differences were found compared to dminor hemoglobin. If, as the starch gel data suggest, there is a small difference between β^dmin and *molossinus* minor β, it must be in the 43% of the molecule for which I have no sequence data.

An N-terminal sequence analysis was also performed on unfractionated *castaneus* globin for 26 positions. At position 20, *castaneus* major β has alanine, rather than the serine found in β^dmaj and β^pmaj. That was the only difference found, compared to the sequence expected from a mixture of d and p hemolysates. At positions 16, 20 and 23 I could distinguish residues present in low amounts relative to the predominant amino acid at a given position, which I will call "minor residues." The minor residues at 16, 20 and 23 were alanine, proline and isoleucine, respectively, which are the ones expected if the *castaneus* hemolysate is like a 50:50 mixture of *molossinus* and p hemolysates, as the starch gel data implied.

Mus caroli

Starch gel data had suggested that the hemolysate of the Thai mouse *M. caroli* has one major component, with only traces of trailing components.[24] It therefore resembles s hemoglobin, the sole hemoglobin of homozygous *Hbb^s/Hbb^s* mice. Cystamine treatment studies similar to those described below for hemoglobin of *M. cervicolor* showed, however, that *caroli* hemoglobin has two cysteines per β chain, like β^dmaj and unlike β^s.[24] A sequence analysis was therefore undertaken in order to see why a β chain resembling β^dmaj in cysteine content was found in an animal that did not also produce β^dmin.

FIGURE 3 shows N-terminal sequenator data for unfractionated *caroli* globin. The data are presented as a semilogarithmic plot of relative yield of amino acids at each position versus position number, as described by Smithies and colleagues.[33] Above the "noise" cutoff line, no more than two different amino acids are found at each position, being from the sequences of α and β chains. Since β^s and the *musculus* α-chain sequences have been almost completely elucidated by Popp,[12, 23] the two residues at each position of the plot can be sorted out with reasonable assurance into the one from the α chain and the one from the β chain. These interpretive sequences of *caroli* α and β chain are shown below the data plot.

These data show a surprising result: In the region analyzed (the N-terminal 43 residues), β^caroli has the same amino acids as β^dmin at two of the three positions at which β^dmin differs from β^dmaj. The data also show that the *caroli* α chain has asparagine at position 8, compared to the serine of *musculus* α.

In order to determine whether β^caroli resembles β^dmin at the other points of difference between β^dmin and β^dmaj, a procedure was used to gain sequenator data on the interior of β^caroli. *M. caroli* globin was cleaved with CNBr, and the resulting mixture of peptides was partially fractionated on Sephadex G-50. The initial peak off the column, which contained one of the larger peptides
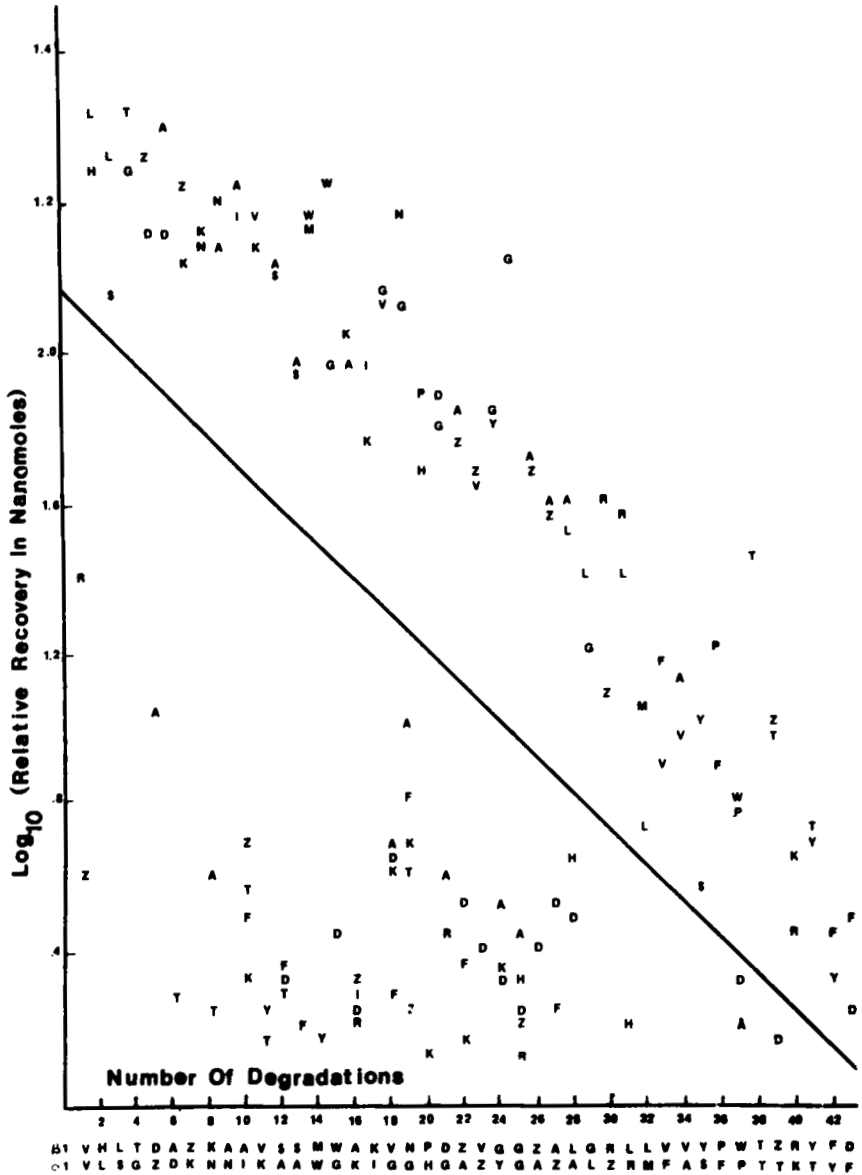
FIGURE 3. N-terminal sequence of unfractionated *Mus caroli* globin. Valine at Position 1 is not shown; it is off scale at a value of 1.8. The plot is that of HI hydrolysis data, as modified to include NaOH hydrolysis data where appropriate. (See legend to FIGURE 1 for symbols.)

and a considerable amount of heme, was sequenced and found to give predominantly a single sequence starting at position 56 of the β chain (by homology with Popp's sequence of β^s [12]). FIGURE 4 shows the semilogarithmic plot of the data for this peptide. Above the cutoff line, residues were recovered at every position but 4 and 14 (which correspond to $\beta 59$ and $\beta 69$). Other sequenator data on mixtures of CNBr fragments from *caroli* globin suggest that lysine and threonine are present at those positions, respectively, as is the case for *musculus* β chain. These residues are shown in parentheses in the interpretive sequence below the data plot. These data show that β^{caroli} is like β^{dmin} at positions 16, 20, 58 and 80, whereas it is like β^{dmaj} at 9, 73, 76 and 77.

Mus cervicolor

Hemoglobins of Thai species *M. cervicolor* show starch gel electrophoresis patterns quite different from the single band of *caroli* hemoglobin. FIGURE 5 shows an alkaline starch gel analysis of hemolysates from two mice of species
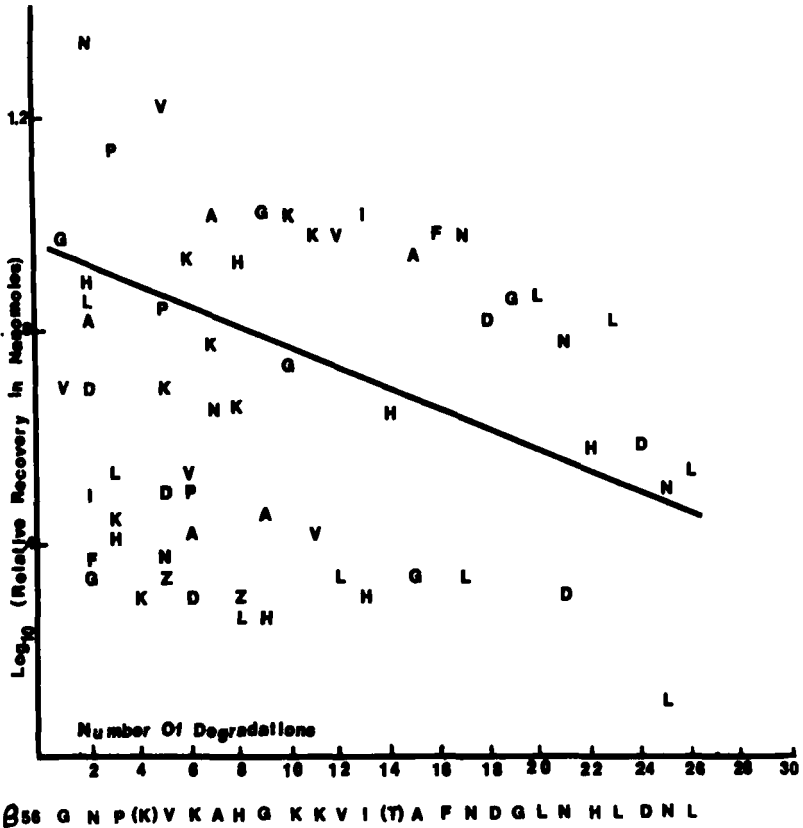

FIGURE 4. Sequence of a peptide from the interior of β^{caroli}, obtained by CNBr cleavage of unfractionated *caroli* globin and Sephadex G-50 fractionation of the mixture of peptides. The plot is that of NaOH hydrolysis data. (See legend to FIGURE 1 for symbols.)

FIGURE 5. Starch gel electrophoresis of hemoglobins of two mice (1 and 3) of species *Mus cervicolor*, as compared to hemoglobin of laboratory strains of *Mus musculus musculus*. All samples are .0065 M. in dithioerythritol (DTE). In both A and B, labels refer to hemoglobin components of the two samples on the right. A. Demonstration that the major hemoglobin of *cervicolor* mouse 3 has half as many free cysteines as pmajor (which has four per hemoglobin molecule, like dmajor). See text for the reasoning behind this conclusion. Sample 1: p hemolysate, +.25 M. NaCl. Sample 2: *Cervicolor* mouse 3 hemolysate, +.25 M. NaCl. Sample 3: p hemolysate, +.25 M. cystamine. Sample 4: *Cervicolor* mouse 3 hemolysate, +.25 M. cystamine. B. Demonstration that the major hemoglobin of *cervicolor* mouse 1 has the same number of free cysteines as dmajor. Sample 5: d hemolysate, +.25 M. NaCl. Sample 6: *Cervicolor* mouse 1 hemolysate, +.25 M. NaCl. Sample 7: d hemolysate, +.25 M. cystamine. Sample 8: *Cervicolor* mouse 1 hemolysate, +.25 M. cystamine.
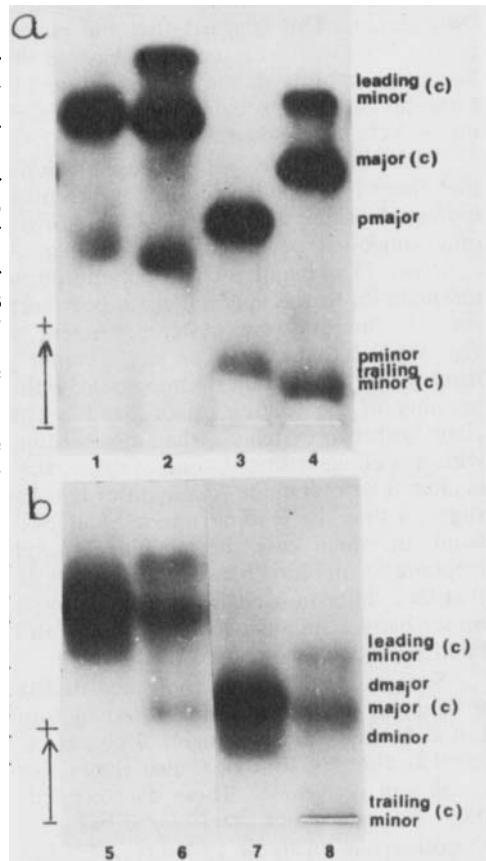
M. cervicolor, which shows that there is a polymorphism in that species for the major hemoglobin component. Hemolysates from the two mice were run both with and without cystamine treatment, which attaches a positively charged S-ethylamine group to each free cysteine of the hemoglobin molecule. In FIGURE 5a, cystamine treated and untreated samples of hemolysate from *cervicolor* mouse 3 are compared to treated and untreated samples of p hemolysate. It can be seen that the mobilities of the leading minor component and the major component of *cervicolor* hemolysate are altered by cystamine treatment to about the same extent, but that the mobility of the major component of the p hemolysate (pmajor) is altered by about twice that amount. Since pmajor has four free cysteines, two per β chain, this suggests that both the major and leading minor components of mouse 3 have only one cysteine per β chain. On the other hand, the mobility of the trailing minor hemoglobin of the p hemolysate (pminor) and that of the *cervicolor* hemolysate are altered to the same extent by cystamine treatment, which suggests that the *cervicolor* trailing minor component has a β chain with two cysteines as does the β chain (β^{pmin}) of the trailing minor component of the p hemolysate.

FIGURE 5b shows that the hemolysate of *cervicolor* mouse 1 does not yield the same result. The mobilities of the leading minor component and the major

component of this hemolysate are altered to the same extent as that of dmajor hemoglobin. This suggests that the major band and the leading minor band of the hemolysate of mouse 1 have β chains with two cysteines, like β^{dmaj}. Since the mobility of the trailing minor band of the mouse 1 is altered by the same amount, it probably has a minor β chain with two cysteines, like the minor β chain of mouse 3.

Thus, these starch gel electrophoresis data on *cervicolor* hemoglobin suggest that there is a polymorphism for the major β chain of *cervicolor*. The data indicate that one of the *cervicolor* major β chains has one cysteine, whereas the other major β produced by the other allele at the β-chain locus has two cysteines. The major β-chain polymorphism of *cervicolor* therefore bears some resemblance to the major β-chain polymorphism of *musculus* for β^s and β^{dmaj}, since β^s has only one cysteine whereas β^{dmaj} has two. The starch gel data also suggest that *cervicolor* has a minor α chain: When a given *cervicolor* hemolysate has a major hemoglobin with a β chain with two cysteines, the mobility of the leading minor band is altered by cystamine as if it had a β chain with two cysteines; when a *cervicolor* hemolysate has a major hemoglobin with a β chain with only one cysteine, the mobility of the leading minor band is altered by cystamine treatment as if it had a β chain with one cysteine. This suggests that the leading minor band has the major β chain, like the major band, in which case the most likely explanation for the difference in electrophoretic mobility between the major band and the leading minor band is that they differ in α chains, the major band having the major α and the leading minor band a minor α . By analogy with the p hemolysate, the trailing minor band is probably composed of the major α chain and a minor β chain.

Sequenator data have confirmed that there is a polymorphism in the major β chain of *M. cervicolor*. FIGURE 6 summarizes the differences that I have found between the two major β chains of *cervicolor* and compares them with β^{dmaj} and β^{caroll}; FIGURE 6 also shows a comparison of α chains of *musculus, caroli* and *cervicolor*. These data confirm what was sugggested by starch gel data; that one major *cervicolor* β has glycine at position 13 (like β^s), whereas the other has serine or cysteine (β^{dmaj} has cysteine). The other two positions at which the two major *cervicolor* β chains differ, 16 and 73, are, however, not positions at which β^s and β^{dmaj} differ (the latter differ at 20 and near the C-terminus, probably at 139 [13]).

In the sequence data on unfractionated cervicolor globin,[24] minor levels of amino acids were found at positions 12 (Asx), 13 (Gly), 16 (Ala), 73 (Glx), 76 (Asn) and 77 (Lys). It is likely, by homology with β^{dmin}, that all of the minor residues but those at positions 12 and 13 are from the *cervicolor* minor β chain. At 12 and 13, it is possible that they are from either the minor α or minor β chains. So it appears that *cervicolor* minor β is probably like β^{dmin} and unlike β^{dmaj} at 16, 73, 76 and 77. *Cervicolor* minor β is also like β^{dmin} in lacking methionine at 109, although this is also true of *cervicolor* major β. There *were* two positions at which minor residues were not found, although they would be expected if *cervicolor* minor β were like β^{dmin}: At 20 and 58, no minor residues of proline were detected. In addition, at 22, neither *cervicolor* major nor *cervicolor* minor β chain has "Glx." At 22, the major chain has alanine; the minor β may have serine or alanine, so that at this position both chains are probably like β^{pmin} (which has alanine). Thus, at seven of the nine positions at which β^{dmin} differs from β^{dmaj} (16, 20, 58, 73, 76, 77 and 109), some statement can be made as to whether or not *cervicolor*

minor β is like β^{dmin}. At five of these positions (16, 73, 76, 77 and 109) *cervicolor* minor β does resemble β^{dmin}.

Discussion

Three alleles at the β-chain locus (*Hbb*) have been identified in laboratory strains of *M. musculus*. *Hbb*[s], a singlet, produces one major β chain, β^{s}, and

```
Beta Chains:
```

	13	14	16	20	22	25	58	73	76	80	104	109	111	121
caroli	Cys	Met	Ala	Pro	Glx	Gly	Pro	Asp	Asn	Asn	**	Met	Leu	Glx
major cervicolor 2	Syr	Leu	Ala	Syr	Ala	Ala	Ala	Asp	*	Asn	?	***	?	?
major cervicolor 3	Gly	Leu	Gly	Syr	Ala	Ala	Ala	Glx	*	Asn	?	***	?	?
βdmaj	Cys	Leu	Gly	Ser	Glx	Gly	Ala	Asp	Asn	*	Arg	Met	Val	Asp

```
Alpha Chains:
```

	8	71	78
caroli	Asn	Ser	Ala
cervicolor	Asn	?	?
musculus	Ser	Gly	Gly

FIGURE 6. A comparison of major α and β chains of *Mus cervicolor*[24] and *Mus caroli*[24] with α chain[23] and β^{dmaj} [11, 18, 24] of *Mus musculus musculus*, at the positions where differences were found. My data on *Mus musculus molossinus* α chain[24] differ from Popp's data[23] on *Mus musculus musculus* α. I found alanine at 70 and glycine at 71, whereas Popp found glycine at 70 and alanine at 71. *Mus caroli* α chain also has alanine at 70.[24] The simplest assumption, which I make in this figure, is that *musculus* α chain has glycine at 71. I have sequence or tryptic peptide composition data on 79% of *caroli* α chain, 42% of *cervicolor* major α chain, 70% of *caroli* β chain, 48% of the major β chain of *cervicolor* mouse 2 and 36% of the major β chain of *cervicolor* mouse 3. Alkaline-urea starch gel electrophoresis of globin chains shows that *cervicolor* mouse 2 has, as its only major β chain, that β which was found in mouse 1[24] (see FIGURE 5). (See legend to FIGURE 1 for symbols.)

no minor chains. *Hbb*[d], a doublet, produces a major β-chain variant, β^{dmaj}, as well as a minor β chain, β^{dmin}. *Hbb*[p], a doublet, produces a major β, β^{pmaj}, which has not been found to differ from β^{dmaj}, as well as a minor β-chain variant, β^{pmin}. Sequence differences among the various chains are shown in FIGURE 1. The *Hbb*[d]-*Hbb*[s] polymorphism is found in every North American or European wild mouse population that was adequately tested by Selander and colleagues;[28] they suggested that selection may be maintaining the polymorphism.

Hemoglobins of wild Asian mice were examined in order to see the extent to which the minor β chain varies among various species. A large degree of variability among minor β chains, or lack of a minor β chain in many species, was to be taken as evidence that there are few selective constraints on the minor β chain, and that it serves no function that contributes to the selective advantage of a mouse possessing it. If this is the case, then it is likely that selection is acting only on the major β-chain polymorphism, for β^s and β^{dmaj}, and that the production by Hbb^d of the minor β, and its lack of production by Hbb^s, does *not* contribute to the maintenaince of the polymorphism.

Variants of the Hbb^d and Hbb^p alleles were found in the Asian *musculus* subspecies *M. musculus molossinus* and *M. musculus castaneus.* Starch gel electrophoresis data indicated that *M. musculus molossinus* has a doublet like Hbb^d, except that its minor β chain differs somewhat from β^{dmin} (although the differences were not revealed by sequence analysis covering 57% of *molossinus* minor β). Starch gel electrophoresis indicated that *M. musculus castaneus* has hemoglobins like a mixture of *molossinus* and p hemoglobins; however, a difference was found between *castaneus* major β and β^{dmaj} (or β^{pmaj}), at position 20. Structural analyses of β^{dmin} from three laboratory strains of *musculus* (Swiss, AKR/J, and C3H/HeJ) showed no variability.[24] These data on the minor β chain of a variety of mice of species *M. musculus,* as well as the electrophoretic study of Selander and colleagues,[28] which reveal no variablity in a large sample of wild mice, are consistent with the hypothesis that the minor β chain fulfills some functionally significant purpose in mice, since little variability was found.

In contrast, hemoglobins from non-*musculus* species differ considerably from those of *M. musculus,* both in major and minor β chains. *M. cervicolor* apparently has one major and one minor β chain, as well as a major and a minor α chain (major and minor α chains have also been demonstrated in *musculus* [15]). There is a polymorphism for the cervicolor major β chain, that resembles the β^s, β^{dmaj} polymorphism of *musculus,* although both *cervicolor* alleles are doublets, each producing a minor β chain as well as the major. At two of the positions at which β^{dmin} differs from β^{dmaj}, the minor β chain of *cervicolor* does not resemble β^{dmin} (at 20 and 58). At five of those positions, however, it does resemble β^{dmin} (at 16, 73, 76, 77 and 109). (For two of those positions, 9 and 80, there is no information.) This conservation of residues in the C-terminal half of the molecule suggests that, if there is a functional differentiation between major and minor β chains, then these C-terminal positions (73, 76, 77 and 109) are the most important in causing that differentiation. A rationale for the functional importance of some of these residues is presented below.

M. caroli, the other non-*musculus* species examined, produces a major β, β^{caroli}, but no minor β. The similarity of β^{caroli} to β^{dmin} in its N-terminal region, and to β^{dmnj} in its C-terminal region, led me to hypothesize that β^{caroli} is a Lepore type of β chain.[9] In other words, I suggested that the gene coding for β^{caroli} might have arisen by means of an unequal crossover between Hbb^{dmaj} (coding for β^{dmaj}) and, on the other chromosome, Hbb^{dmin} (coding for β^{dmin}). This possibility is illustrated in FIGURE 7. Other interpretations of the data are possible, however, which do not assume that β^{caroli} is a Lepore type of β chain.[9]

The fact that *M. caroli* apparently has no minor β chain does not support the hypothesis that the minor β chain of other *Mus* species is functionally

important. If *caroli* hemoglobin were superior to the major hemoglobins of other mice, however, then the fitness of *M. caroli* might not be less than that of other mice, despite the lack of the contribution to its fitness that the minor hemoglobin might be expected to make. I have suggested that β^{caroll} may be functionally different from the other mouse major β chains, because it differs from them at position 14 in having methionine instead of the usual leucine; I proposed that residues of the β chain (such as at 14) that are in contact with tryptophan at position 15 in its hydrophobic pocket may be functionally important.[9] If my hypothesis is correct, the difference at position 14 between

Hypothetical
Ancestral Sequences:

	9	14	16	20	58	73	76	77	80	104	109	111	121
βdmaj	Ala	Leu	Gly	Ser	Ala	Asp	Asn	His	Asn	Arg	Met	Val	Asp
βdmin	Ala	Leu	Ala	Pro	Pro	Glu	Lys	Asn	Asn	Arg	Met	Val	Asp

βcaroli Ala Leu Ala Pro Pro Asp Asn His Asn Arg Met Val Asp

Further Divergence

Present Sequences

βdmaj   Ala Leu Gly Ser Ala Asp Asn His _*_ Arg Met Val Asp

βdmin   Ser Leu Ala Pro Pro Glu Lys Asn Asn Arg Alx Val Asp

βcaroli Ala Met Ala Pro Pro Asp Asn His Asn _**_ Met Leu Glx

FIGURE 7. Unequal crossover. The "Lepore hypothesis" that I proposed as one way to account for the evolution of β^{caroll}.[9] The sequences are shown for only those positions where differences were found between β^{dmaj}, β^{dmin} and β^{caroll}. Positions where divergence is assumed to have occurred following the unequal crossover event are underlined. (See legend to FIGURE 1 for symbols.)

β^{caroll} and the major β chains of other mice might be expected to lead to functional differences between them.

If β^{caroll} is indeed a Lepore type of β chain, then it is possible that selective factors favoring fixation of the crossover chromosome (which lacks the minor β-chain locus) may have been more important than selective factors favoring retention of the minor β-chain locus. For example, if the mutation from leucine to methionine at position 14 took place in the minor β-chain locus, before the unequal crossover event, and if methionine at 14 would have been more advantageous had it occurred in a β chain with C-terminal residues of the

major β chain, then the unequal crossover might have created a selectively favored singlet lacking the minor β-chain gene.

My analysis leaves open the question of whether or not the minor β chain found in most of the mouse species examined serves an important function. Implied in my analysis is the assumption that functional differences between major and minor hemoglobins can lead to selective differences, so that an animal with a minor β-chain locus has a selective advantage over an animal without it. These selective differences are also assumed to be able to account for the maintenaince of the Hbb^d-Hbb^s polymorphism by balancing selection.

A hypothesis offered by Spofford [5] suggests how selective differences deriving from functional differences might account both for the existence of doublet loci and for the maintenaince of the doublet, singlet polymorphism. Spofford assumes, however, that both proteins made by a doublet locus are produced in equal amounts, which is not the case for the Hbb^d locus in mice. So I have generalized Spofford's model to permit its application to those doublet loci which produce the two proteins in unequal amounts, using Spofford's assumptions but a different mathematical formulation.

The most important assumptions are that the enzyme chains produced by the locus in question associate randomly into dimers, and that the "adaptive value W of the genotype is the weighted average of the adaptive values of individuals, real or hypothetical, containing only one of the several dimers formed by the genotype." [5] Although hemoglobin is a tetramer, I am assuming, for the purposes of this analysis of the doublet β-chain locus, that the α chain is the same in every mouse, so that hemoglobin can be formally considered a dimer of β chains. I am also assuming, for the purposes of the analysis of the Hbb^d-Hbb^s polymorphism, that β^s and β^{dmaj} are functionally equivalent, as seems reasonable from the nature of the structural differences between them.

In my mathematical formulation, I define "x" as the relative fitness of the mouse with only $\alpha_2\beta_2^s$ (or $\alpha_2\beta_2^{dmaj}$). I define "nx" as the relative fitness of the hypothetical animal with only the hybrid $\alpha_2\beta^{dmaj}\beta^{dmin}$ (or $\alpha_2\beta^s\beta^{dmin}$). I define the fitness of the hypothetical animal with only $\alpha_2\beta_2^{dmin}$ as unity. Then the fitness W, of an animal with the arbitrary ratio k of β^{dmaj} to β^{dmin}, is $(k^2x + 2knx + 1)/(k^2 + 2k + 1)$. One can then show, by equating dW/dk to zero and stipulating that d^2W/dk^2 be negative, that the fitness W has a unique maximum, as long as $x \geq 1$ and $n > 1$, which occurs when k has the value $(nx - 1)/(nx - x)$. In other words, an animal with the ratio of β^{dmaj} to β^{dmin} equal to $(nx - 1)/(nx - x)$ is the fittest possible animal, as long as $\alpha_2\beta_2^{dmaj}$ is selectively as good as or superior to $\alpha_2\beta_2^{dmin}$ and as long as $\alpha_2\beta^{dmaj}\beta^{dmin}$ is superior to both of them. The more superior $\alpha_2\beta_2^{dmaj}$ is to $\alpha_2\beta_2^{dmin}$, or the more similar $\alpha_2\beta_2^{dmaj}$ is to $\alpha_2\beta^{dmaj}\beta^{dmin}$, the larger the ratio of β^{dmaj} to β^{dmin} will be that optimizes the fitness. This theory shows one way in which the Hbb^d-Hbb^s polymorphism may be maintained by balancing selection: It suggests that the Hbb^d/Hbb^s heterozygote has β chains in an optimum ratio of major to minor chain not obtainable by simple homozygosity and is therefore superior to either homozygote.

This formulation shows that functional differentiation between the major and minor β chains might account for selective differences which maintain the Hbb^d-Hbb^s polymorphism. Using the sequence data, it should be possible to guess whether such functional differences might occur. The criterion of Perutz and Lehmann,[37] that internal changes can lead to functional alterations, suggests that at least one of the sequence differences between β^{dmaj} and β^{dmin}

could lead to functionally important differences between them: At internal position 109, β^{dmaj} has methionine and β^{dmin} has either valine or alanine.

It is also possible that points of difference clustered at positions 73, 76, 77 and 80 might be causes of functional differentiation. Those substitutions are at positions that are potentially able to interact with each other, which suggests the possibility that functionally important interactions may be occurring. The residues, however, are all external. Perutz and colleagues suggest that it is unlikely that position 76 has any functional significance. Their argument is that the residue at position 76, at which alanine is found in humans, "is external, and has no contact with the haem; it is also a site that is highly variable in different species (Asn, Gln, Lys, His and Ser have all been found there), which is against its having any functional importance." [38] These same arguments also hold for positions 73, 77 and 80.

I suggest that this view may be over-simplified. There is evidence that structural variation in at least one of those positions, 73, can lead to functional variation. Bookchin and colleagues showed that Hemoglobin C Harlem, with a β-chain variant that differs structurally from sickle cell hemoglobin in having asparagine rather than aspartic acid at position 73, has an oxygen dissociation curve that differs from that of normal hemoglobin Hb A.[39] Since sickle cell hemoglobin has the same oxygen affinity as Hb A,[40] it is likely that a substitution at position 73 can cause a functional change in human hemoglobin, despite the facts that 73 is external and that evolutionary variation occurs there. Thus, the argument that evolutionary variability implies lack of functional importance may not be completely valid. Indeed, that argument seems to be based on the valid converse assertion, that lack of evolutionary variability at a position implies the functional importance of that position. I feel that a more correct outlook on the meaning of evolutionary variation at a position is that it implies lack of *vital* functional importance of the position. Any mutation in a vital position would probably alter hemoglobin functioning considerably, and it would therefore have a high probability of being harmful. A mutation in a functionally less important position would have a much greater chance of being only slightly deleterious or even slightly beneficial, and thus it could have a significant probability of being incorporated as a substitution in the species.

The Perutz model of hemoglobin structure-function relations therefore does not allow a clear-cut prediction to be made as to whether or not β^{dmin} and β^{dmaj} are likely to differ functionally. It does, however, when applied to α-chain structural differences, suggest that the two α chains produced by the α-chain doublet of mice are likely to differ functionally. Two of the three positions at which those α chains differ are almost invariant in mammalian α and β chains: At positions B6[25] and E11[62], one of the α chains produced by the doublet has the usual residues glycine and valine, respectively, whereas the other α chain has valine and isoleucine.[15] The α chain produced by the singlet α-chain locus of mice generally has the usual residues at 25 and 62,[15] except for one mouse strain that is now extinct.[41]

Perutz suggested that position E11[62] has direct functional importance in hemoglobin.[42] It is a heme contact, and it changes its position relative to the heme upon oxygenation: In deoxy hemoglobin, the valine at E11 blocks the oxygen binding site in the β chain, though not in the α chain, whereas in oxy hemoglobin valine at E11 does not block the binding site. Position 25 is invariant in mammalian globin chains because, according to Kendrew,[43] there

is no room between the B and E helices at that position for a side chain larger than that of glycine.

The fact that one of the mouse α chains has unusual residues at two essentially invariant positions (25 and 62) suggests that the substitutions at those two positions may be interacting. One would expect that one substitution at an invariant position would be harmful, whereas two such substitutions would probably make the hemoglobin totally dysfunctional unless the effects of one substitution compensated for the effects of the other. This impression is reinforced by the fact that grey kangaroo α chain, the only mammalian globin chain besides the unusual α chain of mouse to have other than glycine at B6, has isoleucine at E11 as in the case of that mouse α chain.[44]

It is likely that the unusual mouse α chain, with valine at B6 and isoleucine at E11, would be functionally inferior to the more usual α with glycine at B6 and valine at E11, since no existing mouse strain with a singlet α-chain locus has the unusual α chain as its sole α. Spofford's hypothesis suggests that it is possible that the hybrid hemoglobin containing both α chains is functionally superior to either nonhybrid hemoglobin. A test of this hypothesis could shed light on how hemoglobin evolves functionally.

[NOTE ADDED IN PROOF: The recent study by Nute and colleagues [46] demonstrates the potential functional importance of position β109, which I had suggested in this paper.]

References

1. CEPPELLINI, R. 1959. In Biochemistry of Human Genetics. G. E. W. Wolstenholme & C. M. O'Connor, Eds. : 133. J. & A. Churchill, Lt. London, England.
2. HOLLAN, S. R., J. G. SZELENYI, B. BRIMHALL, M. DUERST, R. T. JONES, R. D. KOLER & Z. STOCKLEN. 1972. Nature 235: 47.
3. ABRAMSON, R. K., D. L. RUCKNAGEL & D. C. SCHREFFLER. 1970. Science 169: 194.
4. HUISMAN, T. H. J., W. A. SCHROEDER, W. H. BANNISTER & J. L. GRECH. 1972. Biochem. Genet. 7: 131.
5. SPOFFORD, J. B. 1972. Brookhaven Symp. Biol. 23: 121.
6. HUTTON, J. J., J. BISHOP, R. SCHWEET & E. S. RUSSELL. 1962. Proc. Nat. Acad. Sci. U.S. 48: 1505.
7. HUTTON, J. J., J. BISHOP, R. SCHWEET & E. S. RUSSELL. 1962. Proc. Nat. Acad. Sci. U.S. 48: 1718.
8. MORTON, J. R. 1966. Genet. Res. Camb. 7: 76.
9. GILMAN, J. G. 1972. Science 178: 873.
10. RIFKIN, D. B., M. R. RIFKIN & W. KONIGSBERG. 1966. Proc. Nat. Acad. Sci. U.S. 55: 586.

11. RIFKIN, D. B. 1968. Ph.D. Dissertation. Rockefeller University. New York, N.Y.
12. POPP, R. A. 1972. *In* Atlas of Protein Sequence and Structure, M. O. Dayhoff, Ed. Vol. 5:D–69. Nat. Biomed. Res. Foundation. Silver Spring, Md.
13. POPP, R. A. Personal communication.
14. EHRENSTEIN, G. VON. 1966. Cold Spring Harbor Symp. Quant. Biol. 31: 705.
15. HILSE, K. & R. A. POPP. 1968. Proc. Nat. Acad. Sci. U.S. 61: 930.
16. SCHAPIRA, G., M. BENRUBI & L. REIBEL. 1969. Biochim. Biophys. Acta 188: 216.
17. MARINKOVIC', D., J. MARTINOVIC' & D. KANAZIR. 1967. Nature 213: 819.
18. STEIN, S., M. G. CHERION & A. MAZUR. 1971. J. Biol. Chem. 246: 5287.
19. DRYSDALE, J. W., P. RIGHETTI & H. F. BUNN. 1971. Biochim. Biophys. Acta 229: 42.
20. GARRICK, L. M. 1973. Arch. Biochem. Biophys. Submitted.
21. BRDIČA, R., A. MASSA, S. CARTA, L. TENTORI & G. VIVALDI. 1972. Life Sci. (II) 11: 895.
22. FLAMM, U., J. S. BEST & G. BRAUNITZER. 1971. Hoppe-Seyler's Z. Physiol. Chem. 352: 885.
23. POPP, R. A. 1972. *In* Atlas of Protein Sequence and Structure, M. O. Dayhoff, Ed. Vol. 5:D–58. Nat. Biomed. Res. Foundation. Silver Spring, Md.
24. GILMAN, J. G. 1973. Ph.D. Dissertation. University of Wisconsin. Madison, Wisc.
25. RICE, N. R. 1972. Brookhaven Symp. Biol. 23: 44.
26. BOYER, S. H., E. F. CROSBY, A. N. NOYES, G. F. FULLER, S. E. LESLIE, L. J. DONALDSON, G. R. VRABLICK, E. W. SCHAEFER, JR. & F. F. THURMON. 1971. Biochem. Genet. 5: 405.
27. GILMAN, J. G. & O. SMITHIES. 1968. Science 160: 885.
28. SELANDER, R. K., S. Y. YANG & W. G. HUNT. 1969. Studies in Genetics V. Publ. 6918. University of Texas Press. Austin, Tex.
29. SMITHIES, O. 1965. *In* Proc. 10th Congr. Intern. Soc. Blood Transfusion. S. Karger, Ed. : 1175. Stockholm, Sweden.
30. WEGMANN, T. G. & J. G. GILMAN. 1970. Develop. Biol. 21: 281.
31. SMITHIES, O., G. E. CONNELL & G. H. DIXON. 1962. Am. J. Hum. Genet. 14: 14.
31. SMITHIES, O., G. E. CONNELL & G. H. DIXON. 1962. Am. J. Hum. Genet. 14: 14.
32. EDMAN, P. & G. BEGG. 1967. Eur. J. Biochem. 1: 80. The instrument used was manufactured by Illitron Division, Illinois Tool Works. Chicago, Ill.
33. SMITHIES, O., D. GIBSON, E. M. FANNING, R. M. GOODFLIESH, J. G. GILMAN & D. L. BALLANTYNE. 1971. Biochemistry 10: 4912.
34. ROSSI-FANELLI, A. & E. ANTONINI. 1958. Biochim. Biophys. Acta 30: 609.
35. GROSS, E. 1967. Methods Enzymol. 11: 238.
36. WHITNEY, J. B. Personal communication.
37. PERUTZ, M. F. & H. LEHMANN. 1968. Nature 219: 902.
38. ANDERSON, N. L., M. F. PERUTZ & G. STAMATOYANNOPOULOS. 1973. Nature New Biol. 243: 274.
39. BOOKCHIN, R. M., R. L. NAGEL & H. M. RANNEY. 1967. J. Biol. Chem. 242: 248.
40. ALLEN, D. W. & J. WYMAN. 1954. Rev. Hematol. 9: 155.
41. RUSSELL, E. S. Personal communication.
42. PERUTZ, M. F. 1970. Nature 228: 726.
43. KENDREW, J. C. 1962. Brookhaven Symp. Biol. 15: 216.
44. DAYHOFF, M. O., Ed. 1972. Atlas of Protein Sequence and Structure. Vol. 5: D–369. Nat. Biomed. Res. Foundation. Silver Spring, Md.
45. BONAVENTURA, J. & A. RIGGS. 1967. Science 158: 800.
46. NUTE, P. E., G. STAMATOYANNOPOULOS, M. A. HERMODSON & D. ROTH. 1974. J. Clin. Invest. 53: 320.

DISCUSSION PAPER:
STRUCTURAL STUDIES OF RAT HEMOGLOBINS [*]

L. M. Garrick, V. S. Sharma,[†] and H. M. Ranney [†]

Department of Medicine
State University of New York at Buffalo
Buffalo, New York 14207

Since we are considering structural aspects of rodent hemoglobins, it seems appropriate to mention some studies on the hemoglobins of the rat. Rat hemoglobins are of particular interest because a severe anemia resembling thalassemia exists in the Belgrade strain.[1] Before we could examine this strain, however, we first characterized the hemoglobins of normal adult Wistar animals.

Six hemoglobin components were isolated from rat hemolysate preparatively by DEAE-cellulose chromatography and analytically by electrophoresis on cellulose acetate.[2] As would be expected from the presence of six hemoglobins, five globin chains were observed, two α chains and three β chains. The chains were isolated from unfractionated hemolysate by CM-cellulose chromatography according to Dintzis,[3] which resolved α from β chains, followed by CM-cellulose chromatography according to Clegg and colleagues,[4] which separated α chains into two components and β chains into three.[2] The subunit composition of the hemoglobins was established by electrophoresis on urea-starch gels.[5] TABLE 1 presents the relationships found. $^{I}\alpha$ is the major α chain and $^{II}\alpha$, the minor. These chains are present in about a 3:1 ratio. $^{II}\beta$ is the major β chain, whereas $^{I}\beta$ and $^{III}\beta$ are both minor chains.

The amino acid sequence of $^{I}\alpha$ was determined and is presented in TABLE 2. There are 22 differences between rat $^{I}\alpha$ and mouse α. Also, 80% of the sequence of $^{II}\alpha$ was established. $^{II}\alpha$ is very similar to $^{I}\alpha$; only one charge difference was found occurring at residue 5, which is Ala in $^{II}\alpha$ and Asp in

TABLE 1

SUBUNIT COMPOSITION OF THE SIX RAT HEMOGLOBINS

Hb Number	Subunit Composition
1	$II_{\alpha_2}III_{\beta_2}$
2	$I_{\alpha_2}III_{\beta_2}$
3	$II_{\alpha_2}II_{\beta_2}$
4	$II_{\alpha_2}I_{\beta_2}$
5	$I_{\alpha_2}II_{\beta_2}$
6	$I_{\alpha_2}I_{\beta_2}$

$^I\alpha$. Two other differences were also found: position 70 in $^{II}\alpha$ was Ala instead of Val and position 73 in $^{II}\alpha$ was Val instead of Leu.

Over 200 rats were examined in our laboratory and all contained six hemoglobins. The presence of two α chains thus probably represents nonallelic α loci, *i.e.*, an α-gene duplication. Structural studies of the three β chains are not yet complete; their presence may reflect a β-gene triplication or may be due to some type of postsynthesis modification.

TABLE 2

PRIMARY SEQUENCE OF RAT I_α

H-Val[a]-Leu-Ser-Ala-Asp-Asp-Lys-Thr-Asn-Ile-Lys-Asn-Asn-Trp-Gly-	15
Lys-Ile-Gly-Gly-His-Gly-Gly-Glu-Tyr-Gly-Glu-Glu-Ala-Leu-Gln-	30
Arg-Met-Phe-Ala-Ala-Phe-Pro-Thr-Thr-Lys-Thr-Tyr-Phe-Asn-His-	45
Ile-Asp-Val-Ser-Pro-Gly-Ser-Ala-Gln-Val-Lys-Ala-His-Gly-Lys-	60
Lys[b]-Val-Ala-Asx-Ala-Leu-Ala-Lys-Ala-Val-Glx-Asx-Leu-Asx-Asx-	75
Leu-Pro-Gly-Ala-Leu-Ser-Thr-Leu-Ser-Asx-Leu-His-Ala-His-Lys-	90
Leu-Arg-Val-Asx-Pro-Val-Asx-Phe-Lys-Leu-Leu-Ser-His-Cys-Leu-	105
Leu-Val-Thr-Leu-Ala-Asx-His-Gly-Pro-Ser-Asx-Phe-Thr-Pro-Ala	120
Met-His-Ala-Ser-Leu-Asp-Lys-Phe-Leu-Ala-Ser-Val-Ser-Thr-Val	135
Leu-Thr-Ser-Lys-Tyr-Arg-OH	141

[a] Residues not underlined were assigned after automatic sequencing.
[b] Residues underlined were assigned after amino acid analysis of the relevant peptide and placed by homology to mouse and rabbit α chains.
[c] Residue not identified.

References

1. SLADIC-SIMIC, D., N. ZIVKOVIC, D. PAVIC, D. MARINKOVIC, J. MARTINOVIC & P. N. MARTINOVITCH. 1966. Hereditary hypochromic microcytic anemia in the laboratory rat. Genetics **53**: 1079–1089.
2. GARRICK, L. M., V. S. SHARMA, M. J. MCDONALD & H. M. RANNEY. Rat haemoglobin heterogeneity: Two structurally distinct α chains and functional behavior of selected components. Biochem. J. In press.
3. DINTZIS, H. M. 1961. Assembly of the peptide chains of hemoglobin. Proc. Nat. Acad. Sci. U.S. **47**: 247–261.
4. CLEGG, J. B., M. A. NAUGHTON & D. J. WEATHERALL. 1966. Abnormal human hemoglobins: Separation and characterization of the α and β chains by chromatography and the determination of two new variants Hb Chesapeake and Hb J (Bangkok). J. Mol. Biol. **19**: 91–108.
5. GARRICK, M. D., R. H. BALZER, JR. & J. P. CHARLTON. 1970. An improved method for electrophoretic characterization of globin chains from hemolysates, purified hemoglobins, and fractions selected from chromatographic separations of chains. Anal. Biochem. **34**: 312–330.