Commentary

SELECTION AT MOLECULAR LEVEL IN MAMMALIAN MYOGLOBINS

O. CASTILLO, H. LEHMANN and R. DIAMOND+

Department of Biochemistry, University of Cambridge and +*MRC Laboratory of Molecular Biology, Cambridge, England*

Received 14 September 1978

The amino acid sequence of 43 mammalian myoglobins is now known. References to 40 of them are summarised in [1], and to the remaining 3 given in [2–4]. Sperm whale myoglobin was the first protein for which the three-dimensional structure was related to the amino acid sequence [5,6] and this analysis has recently been further refined [7]. Table 1 indicates the high degree of resemblance in the primary structure of the mammalian myoglobins. Most of the internal residues of the sperm whale molecule are identical with homologous residues in the sequences of other mammalian myoglobins. This suggests a similarity in their three-dimensional structure. In whale myoglobin a salt bridge is found between 18 Glu and 77 Lys. In all known mammalian myoglobins residues 18 and 77 are also Glu and Lys, respectively. Other residues forming salt bridges found in whale myoglobin are also highly conservative. The homologous residues in the other myoglobins are either identical or similar (Lys instead of Arg, or Glu instead of Asp and vice versa), so that they could form salt bridges. For detailed discussion of the conformity of residues, of conservative regions and of the location of salt bridges and hydrogen bonds see [8].

Address correspondence to Professor H. Lehmann, Dept. of Biochemistry, Tennis Court Road, Cambridge CB2 1QW, England

On comparing the 43 myoglobins two contacts deserve particular comment: In whale myoglobin a salt bridge is found between residues 27 Asp and 118 Arg [6,7], and a relationship is noted between residues 31 Arg and 117 Ser, which are in contact via water molecules [7]. Table 2 shows the remarkable fact that not only the sperm whale but all cetaceans have an Asp in position 27 and an Arg in position 118, but that residue 118 is in all other animals a Lys and residue 27 (with two exceptions) a Glu. Thus when the shorter Lys is found instead of the longer Arg, the longer Glu is associated with it instead of the shorter Asp. Figure 1 shows the computer drawing of the 27–118 Asp–Arg salt bridge of sperm whale myoglobin generated according to the specifications of Watson Kendrew [5,6] and Takano [7] by the computer programme BILDER. Using the main chain atoms as points of reference, two other comparable computer drawn pictures indicate that the change from Asp–Arg to Glu–Lys can arrive at nearly the same bond distance. Asp–Lys appears to exist in this position in two myoglobins, those of Galago and Kangaroo,

Table 1
Resemblance of amino acid sequence in 43 mammalian myoglobins

	Myoglobin chain	External residues	Internal residues	Haem contacts	
				External	Internal
No. residues	153	120	33	10	12
Invariant	90	62	28	9	10
Variant	63	58	5	3	
% Invariant	59	52	85	86	

Table 2
Reciprocal mutations in positions 27 and 118, and 31 and 117 in 43 mammalian myoglobins

Helical No.	Sequential No.	MAN	CHIMPANZEE	GORILLA	ORANGUTAN	GIBBON	SIAMANG	BABOON	MACAQUE	WOOLLY MONKEY	SQUIRREL MONKEY	MARMOSET	GALAGO	POTTO	SLOW LORIS	SPORTIVE LEMUR	SPERM WHALE	GRAY WHALE	DOLPHIN	PORPOISE	AMAZON RIVER DOLPHIN	BOTTLE-NOSED DOLPHIN	KILLER WHALE	HORSE	ZEBRA	OX	SHEEP	RED DEER	PIG	HARBOUR SEAL	SEA LION	DOG	BAT-EARED FOX	CAPE FOX	HUNTING DOG	BADGER	TREESHREW	HEDGEHOG	FRUIT BAT	RABBIT	KANGAROO	OPOSSUM	PLATYPUS	ECHIDNA
B8	27	E	E	E	E	E	E	E	E	E	E	E	**D**	E	E	E	**D**	**D**	**D**	**D**	**D**	**D**	**D**	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	**D**	E	E	E
G19	118	K	K	K	K	K	K	K	K	K	K	K	K	K	K	K	**R**	**R**	**R**	**R**	**R**	**R**	**R**	K	K	K	K	K	K	K	K	K	K	K	K	K	K	K	K	K	K	K	K	K	K
B12	31	R	R	R	R	R	R	R	R	**S**	**S**	**S**	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R
G18	117	S	S	S	S	S	S	S	S	**K**	**K**	**K**	N	S	S	S	S	S	S	S	S	S	S	S	A	A	A	S	S	S	S	S	N	S	S	S	S	S	S	S	S	S	S	S	S

A: alanine; D: aspartic acid; E: glutamic acid; K: lysine; N: asparagine; R: arginine; S: serine

but not in the myoglobins of other prosimians and non-eutherians; 27–118 Glu–Arg has never been found. Residues 27 and 118 in the reconstructed ancestral myoglobin chain are Glu and Lys [8]. It thus seems that one observes here natural selection at the molecular level. When the longer Glu changes to the shorter Asp the shorter Lys is replaced by the longer Arg. The Galago and Kangaroo would represent the transitional stage. In Glycera haemoglobin which is a monomer consisting of 147 residues [9] the homologous residues are Asp and Arg as in the cetaceans.

Another example of reciprocal mutations would be the pairing mentioned above of residues 31–117, which usually are Arg and Ser (in some cases Arg and

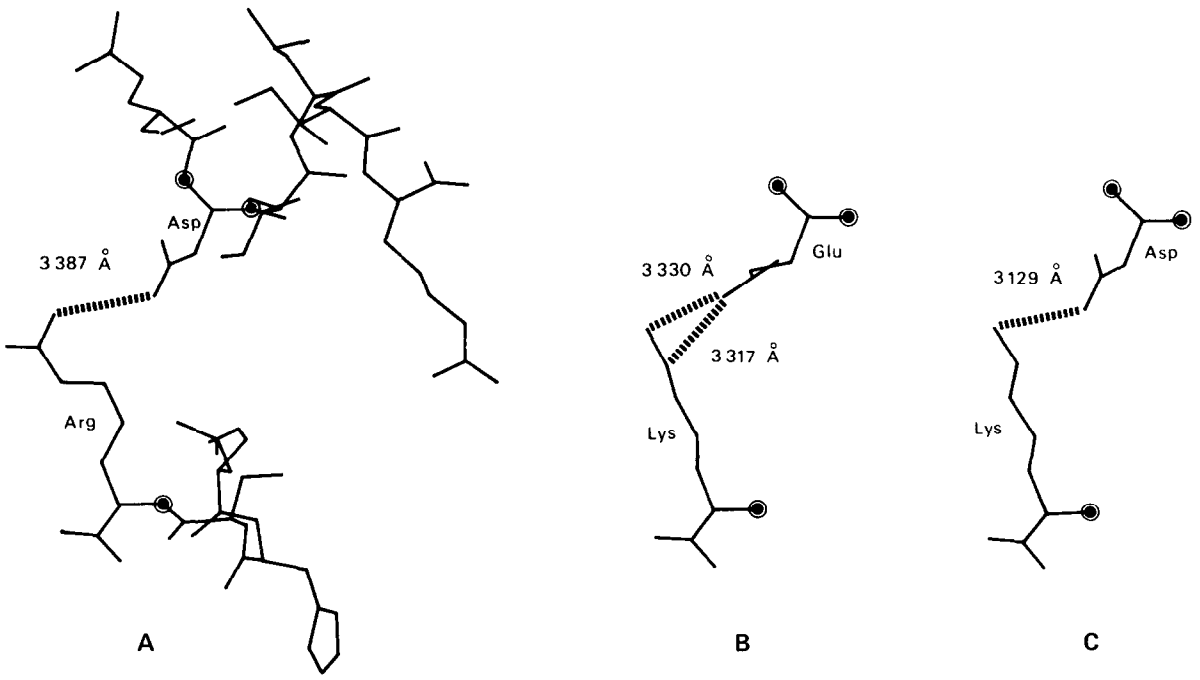

Fig.1. (A) Computer-generated illustration of the salt bridge 27 Asp–118 Arg in whale myoglobin. (•) Main chain atoms. (B) Corresponding salt bridge Glu–Lys. (C) Corresponding salt bridge Asp–Lys. (For details see text.)

Ala or Arg and Asn), but in the three New World Monkeys are Ser and Lys (table 2). Here again the water-mediated contact would be preserved although the nature of the residues is reversed.

A similar reversal can be seen in the myoglobin of Aplysia [10]. It has been pointed out [8] that whereas in all vertebrate myoglobins residue 45 is occupied by a basic amino acid residue which is considered to form a salt bridge with an acid residue at position 60, the homologous side chains in Aplysia are reversed and are 45 Asp and 60 Lys.

We suggest that these relationships indicate natural selection. If a change occurs in the sequence it could affect the overall molecular configuration. Levitt [11] suggests as a possible consequence an alteration in the tertiary structure which would minimise the energy involved. We would propose that the fixing of a second compensating mutation would restabilise the molecule.

References

[1] Castillo, O. and Lehmann, H. (1977) Biochim. Biophys Acta 492, 232–236.
[2] Bruce, F. J., Castillo, O. and Lehmann, H. (1977) FEBS Lett. 78, 113–118.
[3] Jones, L. T., Castillo, O. and Lehmann, H. (1977) Biochim. Biophys. Acta 493, 460–464.
[4] Castillo, O., Jones, L. T. and Lehmann, H. (1978) Biochim. Biophys. Acta 533, 289–292.
[5] Kendrew, J. C. (1962) Brookhaven Symp. Biol. 15, 216–228.
[6] Watson, H. C. (1969) Prog. Stereochem. 2, 299–333.
[7] Takano, T. (1977) J. Mol. Biol. 110, 537–568.
[8] Romero-Herrera, A. E., Lehmann, H., Joysey, K. A. and Friday, A. E. (1978) Phil. Trans. Roy. Soc. B. 283, 61–163.
[9] Imamura, T., Baldwin, T. O. and Riggs, A. (1972) J Biol. Chem. 247, 2785–2797.
[10] Tentori, L., Vivaldi, G., Carta, S., Marinucci, M., Massa, A., Antonini, E. and Brunori, M. (1973) Int. J. Peptide Protein Res. 5, 187–200.
[11] Levitt, M. (1976) J. Mol. Biol. 104, 59–107.