REPTILIAN HEMOGLOBINS: N-TERMINAL SEQUENCE OF AN α-CHAIN OF VIPER (*VIPERA ASPIS*) HEMOGLOBIN

Michel DUGUET, Jean-Pierre CHAUVET et Roger ACHER
Laboratory of Biological Chemistry, 96 Boulevard Raspail, Paris VI, France

Received 16 September 1971

1. Introduction

Although both birds and mammals diverged from ancestral reptiles, it has been found that avian neurohypophysial hormones are similar to those of present-day reptiles and different from those of mammalian species [1]. Because the amino acid sequences of several mammalian hemoglobin chains [2] and of the α-chain of chicken hemoglobin [3] are known, it is of interest to compare these proteins with their reptilian homologues in order to see to what extent the similarity between birds and reptiles exists at the molecular level. Moreover, the structure of a β-chain of a frog hemoglobin has also been determined [4] and because amphibians gave rise to reptiles, the comparison with a reptilian hemoglobin should also be instructive.

2. Methods and materials

2.1. *Isolation of a viper α-chain*

Vipers, after taking of the venom, were decapitated and blood was collected in a sodium citrate solution [5]. Red cells were separated by centrifugation, and hemolysis performed as previously described [5, 6]; hemoglobin was purified by fractionated precipitation with ammonium sulfate and gel filtration on Sephadex G-100 in 0.076 M Tris buffer pH 8.6. The material appears homogeneous by chromatography on anionic (Carboxymethyl-Sephadex at pH 6.0) [7] or cationic (Diethylaminoethyl-Sephadex at pH 8.5) [8] ion-exchangers as well as by electrophoresis on starch gel at pH 7.5 and 8.6 [9] and on cellulose acetate [10] at pH 5.8, 7.4, 8.6, 9.4 and 10.25.

After removal of the heme group by acetone precipitation of the globin [5], the protein was reduced by β-mercaptoethanol at pH 8.6 in 8 M urea and the cysteinyl residues were alkylated with iodoacetic acid [11]. Carboxymethylated globin (100 mg) was subjected to a counter-current distribution in a secondary butanol–0.5 M acetic acid–6% dichloroacetic acid (9:10:1, v/v/v) solvent system under the conditions previously described [5, 6]. After 60 transfers, the protein concentration of each tube was

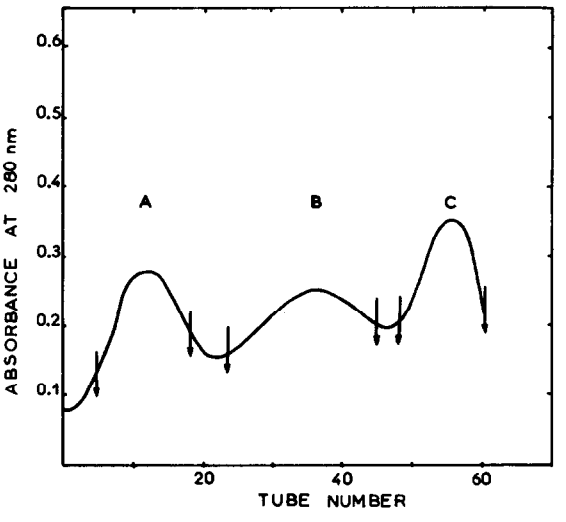

Fig. 1. Counter-current distribution of carboxymethylated globin. The three components A, B and C are collected as shown by arrows.

Table 1
Determination of the *N*-terminal sequence of the α-chain.

Peptide		Sequence	Number of residues
MT_I	T_1	1 7 → → → → → → Val–Leu–Ser–Glu–Asp–Asp–Lys	7
	T_2	8 9 → Asn–Arg	2
MT_{II}	T_3	10 11 → Val–Arg	2
MT_{III}	T_4	12 16 → → → → Thr–Ser–Val–Gly–Lys	5
	T_5	17 31 → Asn (Glu_3, Leu_2, Pro_2, Gly_2, Thr_2, Ser, Tyr) – Arg	15
MT_{IV}		32 Hse	1

→ Determination by Edman degradation [14].
MT: tryptic fragment of maleylated *N*-terminal fragment.
T: tryptic peptide of *N*-terminal fragment.

determined by spectrophotometry at 280 nm. Fig. 1 shows the results. Three peaks, A, B and C representing 10, 35 and 16%, respectively, of the material used for the distribution were observed. They correspond to different molecular species as judged by amino acid composition, N and C terminal sequences and tryptic peptide maps. The three chains have a molecular weight of about 16,000 as determined by gel filtration on Sephadex G-100 (0.1 M acetic acid, 8 M urea), using human and frog α- and β-hemoglobin chains as references. Component A was studied first because it contains a single methionyl residue per mole (in contrast to components B and C which have no methionine) so that cleavage by cyanogen bromide can be used for obtaining two fragments. On the other hand the N and C terminal sequences, the position of the methionine and other features indicated that component A is an α-type chain.

2.2. *Identification of an α-type chain*

The homogeneity of the component A was checked by electrophoresis on cellulose acetate [10]. An N-terminal sequence, Val–Leu–Ser, was determined by Sanger's dinitrofluorobenzene [12] and Edman's phenylisothiocyanate [13] techniques. Carboxypeptidase B hydrolyzed the chain (weight ratio enzyme/substrate 2%; 38°; pH 8.0; 30 min) releasing arginine (0.9 residue) and tyrosine (0.4 residue) so that the C-terminal sequence Tyr–Arg could be deduced. Cyanogen bromide cleavage [14] gave an N-terminal fragment of 32 residues. Methionine is an invariant residue in the position 32 of the α-chain family [2]; this feature and the nature of the N- and C-terminal sequences allow the conclusion that component A is an α-type chain.

The N-terminal fragment was prepared as follows: 100 mg of carboxymethylated chain was dissolved in 2 ml of 70% formic acid and 250 mg of cyanogen bromide were added. After 24 hr at room temp., the

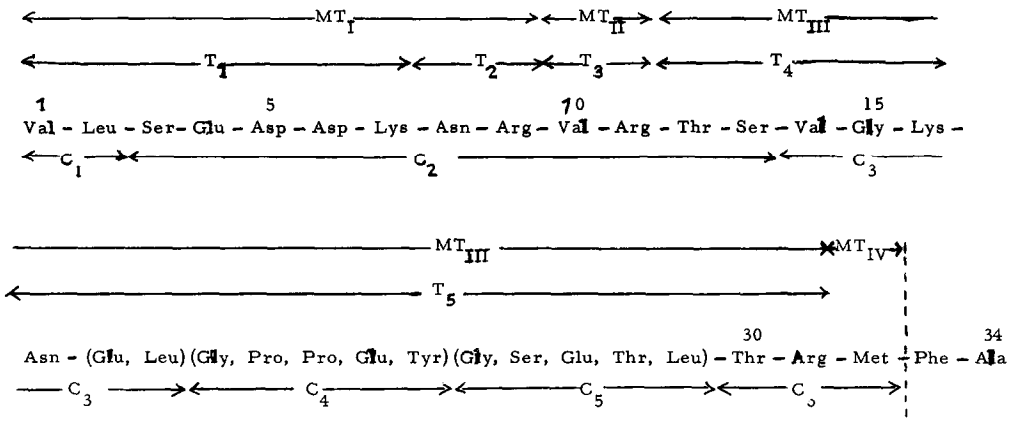

Fig. 2. Alignment of tryptic and chymotryptic peptides of the N-terminal fragment of the component A.

solution was diluted with 10 vol. of water, and the reagents eliminated by 3 consecutive lyophilizations. The 2 fragments of the chain were separated by gel filtration on Sephadex G-50 in 0.1 M acetic acid and subjected to amino acid analysis and determination of the N-terminal sequence. The smaller fragment contained 32 residues and its N-terminal sequence was Val–Leu–Ser–Glu–Asp–Asp. The larger fragment had about 110 residues and its N-terminal sequence was Phe–Ala. It is evident that the smaller polypeptide is the N-terminal fragment of the α-chain and the sequence Met–Phe–Ala (Residues 32–34) can also be deduced.

2.3. *Study of the N-terminal fragment*

The polypeptide contains 3 arginines and 2 lysines; trypsin cleavage (weight ratio enzyme/substrate 1%; 37°; 0.1 M ammonium bicarbonate, pH 8; 3 hr) gives 7 products which were separated by peptide mapping on paper [15]. After acid hydrolysis (6 N HCl; 24 hr; 105°), amino acid analysis was performed [16]. 5 peptides (T_1–T_5) and homoserine were isolated and the sequences of T_1, T_2, T_3 and T_4 determined by Edman degradation [13] as shown in table 1. An additional peptide T_{1+2} was also found in the tryptic digest.

The N-terminal residue of T_5 is asparagine and the amino acid composition is Asn (Glu_3, Leu_2, Gly_2, Pro_2, Thr_2, Tyr, Ser) Arg. The alignment of tryptic peptides was partially determined by blocking lysine residues of the cyanogen bromide fragment with maleic anhydride [17] and isolating overlapping peptides after subsequent tryptic cleavage. Four products can be identified: MT_I : $T_1 + T_2$, MT_{II}: T_3, MT_{III}: $T_4 + T_5$, and MT_{IV} which is homoserine. Chymotryptic hydrolysis of the cyanogen bromide fragment (weight ratio enzyme/substrate 1%; 37°; 0.1 M ammonium bicarbonate, pH 8.0; 4 hr) gave 8 peptides, one of which is a tripeptide Thr–Arg–Hse (see footnote*). Because of the structures of T_2 (Asn–Arg) and T_3 (Val–Arg), it is evident that T_5 has to be placed before the homoserine residue in the sequence and the complete alignment can therefore be deduced (fig. 2). The other chymotryptic peptides confirm the arrangement.

3. Discussion

Tryptic cleavage of the bond Lys-7–Asn-8 was not complete under the conditions used and a peptide T_{1+2} was isolated; this fact does not seem to be explained by the proximity of Arg-9 because trypsin splits completely the bond Lys–Ala in the sequence Lys–Ala–Arg of the oxidized trypsin inhibitor of Kunitz [18]. In contrast, the anionic sequence before the lysyl residue in the sequence Glu–Asp–Asp–Lys–Asn is probably the cause of this slow cleavage, similarly to the slight hydrolysis observed after 90 min in the sequence Glu–Glu–

* Hse: Homoseryl residue.

Asp–Lys–Ala of the γ-chain of human fetal hemoglobin [13]. Chymotrypsin shows the expected specificity except for the cleavage of the Ser–Val bond. Chymotryptic splitting near the carboxyl group of serine is very rare [19] but because of cleavage by elastase of the Ser–Val and Ser–Leu bonds in insulin A-chain and Ser–His in insulin B-chain [20], it is possible that the chymotrypsin used was slightly contaminated by elastase; another explanation could be that chymotrypsin has a slight intrinsic elastolytic activity which is higher when seryl bonds involve a hydrophobic residue such as valine.

It is of interest to note that the invariant residues observed in mammalian α-chains and chicken α-chain [2, 3] at positions 1, 2, 3, 6, 7, 11 and 16 are found in the viper α-chain except for lysine 11 which is replaced by an arginine. In contrast, in amphibian α-chains [6, 7] and carp α-chain [21], the N-terminal valine is replaced by Acetyl–Ala and Acetyl–Ser, respectively.

Acknowledgements

The authors are indebted to Dr. Jacqueline Chauvet for collecting viper blood. They wish to thank Miss Christiane Devaux for her skilled technical assistance.

References

[1] R. Acher, J. Chauvet and M.T. Chauvet, European J. Biochem. 17 (1970) 509.
[2] M.O. Dayhoff, Atlas of Protein Sequence and Structure, The National Biomedical Research Foundation (1969).
[3] G. Matsuda, H. Takei, K.C. Wu, K. Minuzo and T. Shiozawa, 8th International Congress of Biochemistry, Abstracts (1970) 4.
[4] J.P. Chauvet and R. Acher, FEBS Letters 10 (1970) 136.
[5] J.P. Chauvet and R. Acher, Compt. Rend. Acad. Sci. (Paris) 265 (1967) 2084.
[6] J.P. Caffin, J.P. Chauvet and R. Acher, FEBS Letters 5 (1969) 196.
[7] J.P. Chauvet and R. Acher, Intern. J. Protein Res. (1971) in press.
[8] T.H.J. Huisman and A.M. Dozy, J. of Chromatog. 19 (1965) 160.
[9] O. Smithies, Advan. in Prot. Chem. 19 (1964) 1.
[10] J. Kohn, Das Artzliche Laboratorium 10 (1964) 223.
[11] A.M. Crestfield, S. Moore and W.H. Stein, J. Biol. Chem. 238 (1963) 622.
[12] J. Chauvet, G. Nouvel and R. Acher, Biochim. Biophys. Acta 115 (1966) 130.
[13] W.A. Schroeder, J.R. Shelton, J.B. Shelton, J. Cormick R.T. Jones, Biochemistry 2 (1963) 992.
[14] J. Schreiter and W. Witkop, J. Am. Chem. Soc. 86 (1964) 2441.
[15] C. Baglioni, Biochim. Biophys. Acta 48 (1961) 489.
[16] D.H. Spackman, W.H. Stein and S. Moore, Anal. Chem. 30 (1958) 1190.
[17] P.J.G. Botler, J.I. Harris, B.S. Hartley and R. Lederman, Biochem. J. 112 (1969) 679.
[18] R. Acher and J. Chauvet, in: Structure-Function Relationships of Proteolytic Enzymes, eds. P. Desnuelle, H. Neurath and M. Ottesen (Munksgaard, Copenhagen, 1970).
[19] G.L. Neil, C. Nieman and G.E. Hein, Nature 210 (1966) 903.
[20] M.A. Naughton and F. Sanger, Biochem. J. 78 (1961) 156.
[21] K. Hilse, V. Sorger and G. Braunitzer, Hoppe-Seyler's Z. Physiol. Chem. 344 (1966) 166.