PHYLOGENY OF HEMOGLOBINS: AMINO ACID SEQUENCE OF RESIDUES 35 TO 92 OF A VIPER *(VIPERA ASPIS)* HEMOGLOBIN α-CHAIN

Michel DUGUET, Jean-Pierre CHAUVET and Roger ACHER
Laboratory of Biological Chemistry, 96 Bd Raspail, 75006-Paris, France

Received 12 May 1974

1. Introduction

In previous papers separation of the polypeptide chains of a viper hemoglobin by counter-current distribution [1] and determination of the amino acid sequence of the first 34 residues of an α-chain [2] have been described. We report now the sequence analysis of the middle part of the α-chain, namely the sequence of the next 58 residues. Because about 65% of the primary structure of the viper hemoglobin α-chain are presently known (92 residues located out of 141–144), a comparison can be made with homologous parts of the mammalian, avian, and fish α-chain known to date.

2. Results

2.1. *Purification of a C-terminal fragment of α-chain (CN_{II})*

Cyanogen bromide cleavage [1] applied on the whole globin leads to a mixture containing intact β-chain and two fragments, N-terminal (32 residues) and C-terminal (about 110 residues), of α-chain [2,3]. Isolation of the N-terminal fragment (CN_I) by gel filtration on a column of Sephadex G-50 has previously been described [3] and partial purification of the large C-terminal fragment (CN_{II}) can be carried out in the same way. 100 mg of cyanogen bromide treated globin are passed through a 4.5 × 120 cm column of Sephadex G-50 Fine equilibrated with 0.5 M acetic acid, and 3 ml-fractions are collected. Spectrophotometry at 280 nm shows three peaks corresponding to intact chains (tubes 60–70), CN_{II} (tubes 70–90) and CN_I (tubes 110–125). Fractions containing CN_{II} are pooled and contaminating intact chains are removed by selective precipitation. 400 mg of material purified by gel filtration are dissolved in 40 ml of 0.5 M acetic acid and the pH is adjusted to 8.7 with 5 N NaOH. The precipitate, consisting essentially of intact chains, is removed by centrifugation (16 000 *g*, 20 min) and the supernatant solution is freeze-dried (250 mg). Final purification of CN_{II} is carried out by gel filtration on a 2.6 × 90 cm column of Sephadex G-50 equilibrated with 0.1 M acetic acid. Samples of 30 mg are passed through the gel, 6 ml-fractions are collected and absorbance at 280 nm is determined. Tubes 37 to 47 are pooled and the material is freeze-dried.

From 1 g of globin, 130 mg (about 10 μmol) of fragment CN_{II} is recovered. Homogeneity of the polypeptide is checked by determining the N-terminal sequence; Edman degradation on paper strip applied under conditions previously described [4] gives the amino acid sequence Phe–Ala–Ala–His–Pro–Thr.

2.2. *Preparation of citraconylated tryptic peptides of CN_{II}*

Amino acid composition of CN_{II}, determined with samples of 40 nmol by using a Spinco 120 B automatic analyzer [5], indicates that the polypeptide contains about 110 residues 9 of which are lysine and 3 or 4 arginine. Blockage of lysine residues, carried out by citraconic anhydride [6] limits trypsin cleavage to arginine bonds. 120 mg of CN_{II} are dissolved in 10 ml of water and the pH is adjusted at 9.0 with NH_4OH. Eight additions of 40 μl of citraconic anhydride are made at 15 min intervals, the pH being maintained between 8.2 and 8.3 with 5 N NaOH. The mixture is kept for 90 min at room temp and reagents are removed

Table 1
Peptides of the fragment CN_{II} CT_I

Peptide		Sequence	Number of residues
T_6		33 → → → → → 40 Phe–Ala–Ala–His–Pro–Thr–Thr–Lys	8
T_7		→ → → → → → → → Thr–Tyr–Phe–Pro–His–Phe–Asp–Leu	16
	T_7Ch_1	41 → 42 → Thr–Tyr	2
	T_7Ch_2	43 → → → 46 → Phe–Pro–His–Phe	4
	T_7Ch_3	47 → 48 → Asp–Leu	2
	T_7Ch_{3-4}	→ → → Asp–Leu–Ser (Ser, Gly, Ser, Pro, Asn, Leu) Lys	10
	T_7Ch_4	49 → → → → → → 56 Ser–Ser–Gly–Ser–Pro–Asn–Leu–Lys	8
T_8		→ → → Ala–His–Gly–Lys	4
T_{8-9}		57 → → → 61 Ala–His–Gly–Lys–Lys	5
T_9		Lys	1
T_{10}		→ → → → → → → → Val–Ile–Asp–Ala–Leu–Asp–Asn–Ala	21
	$T_{10}Ch_1$	62 → → → → 66 → Val–Ile–Asp–Ala–Leu	5
	$T_{10}Ch_2$	67 → → → → → 73 Asp–Asn–Ala–Val–Glu–Gly–Leu	7
	$T_{10}Ch_3$	74 → → → → → 80 Asp–Asp–Ala–Val–Ala–Thr–Leu	7
	$T_{10}CH_{3-4}$	74 → → → → → → 82 Asp–Asp–Ala–Val–Ala–Thr–Leu–Ser–Lys	9
	$T_{10}Ch_4$	81 → 82 Ser–Lys	2
T_{11}		83 → → → → → → → 90 Leu–Ser–Asp–Leu–His–Ala–Gln–Lys	8
T_{12}		91 → 92 Leu–Arg	2

→ Determination by Edman degradation [4]. T: tryptic peptides of the uncitraconylated CN_{II} CT_I. T_7Ch: Chymotryptic peptides of T_7, etc . . .

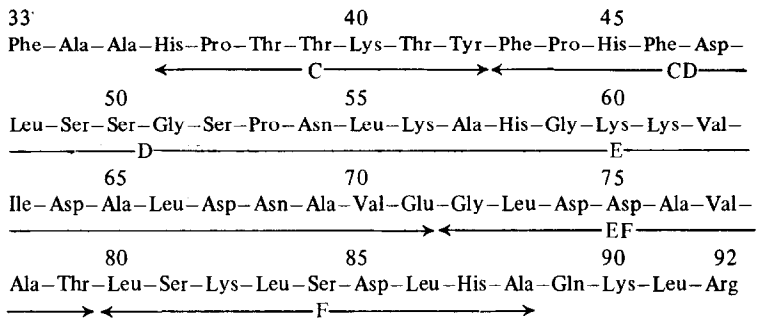
by gel filtration on Sephadex G-25 in 0.1 M ammonium bicarbonate.

The citraconylated polypeptide is cleaved by trypsin under conditions previously described [3] (weight ratio enzyme/substrate 1%, 37°C, 0.1 M ammonium bicarbonate, pH 8.0, 3 hr). The mixture of peptides is passed through a column of Sephadex G-50. Several peaks are observed, the first corresponds to intact CN_{II} and the second to the largest citraconylated tryptic peptide of CN_{II}. Determination of the N-terminal amino acid sequence Phe–Ala–Ala reveals that this peptide (CN_{II}–CT_I) is pure and represents the N-terminal portion of CN_{II}.

2.3. *Determination of amino acid sequence of CN_{II} CT_I*

Amino acid composition of CN_{II} CT_I (5 nmol) indicates that the peptide contains about 60 residues six of

Table 2
Amino acid sequence of the middle-part of the viper α-chain



which are lysine. Citraconyl groups are removed by acid treatment. 42 mg (6.5 μmol) of peptide are dissolved in 4 ml of water and the pH is adjusted to 2.7 with 5 M acetic acid. The mixture is stirred under nitrogen at 37°C for 7 hr, then freeze-dried. The uncitraconylated peptide (5 μmol) is subjected to tryptic hydrolysis and the tryptic peptides are separated by chromatoelectrophoresis on paper under the previously described conditions [4]. Peptide mapping shows eight spots seven of which are peptides and one free lysine. Amino acid compositions are determined and amino acid sequences of the five shorter peptides are established by using Edman degradation (table 1). The tryptic peptides are numbered from T_6 to T_{12} because of their position in the α-chain. Peptides T_7 and T_{10}, which contain 16 and 21 residues respectively, are split by chymotrypsin (weight ratio enzyme/substrate 1%; 0.1 M ammonium bicarbonate pH 8.0; 37°C; 2 hr) and chymotryptic peptides are separated by chromatoelectrophoresis. T_7 and T_{10} give four and five chymotryptic peptides, respectively, which are sequenced by Edman degradation (table 1). The complete amino acid sequences of T_7 and T_{10} can be deduced from N-terminal sequences and overlapping chymotryptic peptides.

The alignment of tryptic peptides has been determined through overlapping peptides isolated from a chymotryptic digest of the uncitraconylated CN_{II} CT_I fragment. The chymotryptic peptide Ala-33–Tyr-42 gives the alignment T_6–T_7, the peptide Asp-47–Leu-66 gives the alignment T_7–$T_{8-9}T_{10}$ and the peptide Ser-81–Leu-83 gives the alignment T_{10}–T_{11}. The amino acid sequence from residue 33 to residue 92 is shown in table 2.

3. Discussion

When the sequence of residues 33–92 of the viper hemoglobin α-chain is compared to the homologous part of the human α-chain one can observe 32% of substitutions instead of 50% found between the homologous N-terminal parts. Clearly the middle part of the chain is less variable and this can be explained by the number of residues interacting with heme or involved in the contacts between α_1 and β_2-chains.

According to Perutz [6], the sequence 33–92 is arranged in helices C, E, and F and bends CD, EF and FG. In this sequence eleven residues: 39 (C_4), 42 (C_7), 43 (CD_1), 45 (CD_3), 46 (CD_4), 58 (E_7), 62 (E_{11}), 83 (F_4), 86 (F_7), 87 (F_8) and 91 (FG_3) are in contact with the heme and nine of these are invariant. Five residues: 38 (C_3), 41 (C_6), 42 (C_7), 91 (FG_3) and 92 (FG_4) are involved in the contacts between α_1 and β_2-chains and four are invariant. By contrast, in the 32-residue N-terminal sequence, no residue is involved in the contacts between α_1 and β_2-chains and one in the contacts α-chain–heme.

The comparison of the sequence of the first 92 residues of viper hemoglobin α-chain reveals 35 substitutions (38%). The percentage is clearly higher than for chicken α-chain (25%) and lower than for carp α-chain (50%).

Because mammalian, avian and reptilian lines are supposed to have diverged from a common vertebrate stock approximately 300 million years ago, it is surprising to find a great difference between chicken and viper in the percents of substitutions when human α-

chain is taken as the reference. It might be possible that the rate of hemoglobin evolution was not uniform in all the vertebrate lines.

Acknowledgement

The authors wish to thank Miss Christiane Devaux and Miss Monique Bourdin for their skilled technical assistance.

References

[1] Gross, E. and Witkop, B. (1962) J. Biol. Chem. 237, 1856–1860.
[2] Duguet, M., Chauvet, J. P. and Acher, R. (1971) FEBS Letters 18, 185–188.
[3] Duguet, M., Chauvet, J. P. and Acher, R. (1973) FEBS Letters 29, 10–14.
[4] Chauvet, J. P. and Acher, R. (1972) Biochemistry 11, 916–927.
[5] Spackman, D. H., Stein, W. H. and Moore, S. (1968) Anal. Chem. 30, 1190–1206.
[6] Dixon, H. B. F. and Perham, R. N. (1968) Biochem. J. 109, 312–313.