THE β-CHAIN OF BADGER HAEMOGLBIN: AMINO ACID COMPOSITION OF THE TRYPTIC PEPTIDES AND THE N-TERMINAL SEQUENCE TO POSITION 42

S. DUCASTAING, I. HOMBRADOS, A. IRON and E. NEUZIL
Biochimie Médicale, U.E.R. III, Université de Bordeaux II, 146, rue Léo Saignat, 33076, Bordeaux, France

and

Y. PLANCKE and K. HAN
Laboratoire de Chimie Biologique, Faculté de Médecine, Place de Verdun, 59045, Lille Cédex, France

Received 2 July 1973

1. Introduction

We report here the partial structure of the β-chain of the haemoglobin of the Badger (*Meles meles*), an animal of the hitherto poorly studied group of carnivores. The tryptic hydrolysate was resolved by electrochromatography and the amino acid composition of 15 peptides is given. The amino acid sequence to position 40 has been determined by analysis of four tryptic peptides, which are compared to the homologous peptides isolated from human β-chain. The sequence is confirmed and extended to position 42 by stepwise degradation of the whole β-chain, using a sequencer.

2. Materials and methods

The haemoglobin was isolated from haemolysates obtained from heparinized blood [1] by conventional techniques, and purified by paper electrophoresis at pH = 8.6 [2]. The globin moiety was precipitated by cold hydrochloric acetone according to the method of Rossi-Fanelli [3]. The α- and β-chains were separated following the technique of Clegg [4], as modified by Adams et al. [5]: the globin dissolved in urea–mercaptoethanol medium was chromatographed on a 3 X 25 cm column of Whatman CM-32 carboxymethyl-cellulose. The separated chains were treated with ethyleneimine, and purified by gel filtration on Sephadex G-25. The S-aminoethyl-β-chain (20 mg) was dissolved in 2 ml of water; 0.4 ml of 0.5 M NH_4HCO_3 was added to give a final pH of 8.5; the digestion was carried out with 0.2 ml of Worthington TPCK trypsin solution (2 mg/ml in water) at 38°C for 2 hr, with occasional stirring. The hydrolysate was submitted to electrochromatography [6] and the peptide map visualized by 3% ninhydrin acetone spray and by specific staining [7].

The different peptides were eluted from the paper and hydrolyzed by 5.6 N HCl, or by aminopeptidase M (Röhm, Darmstadt) when the presence of asparagine or glutamine was suspected. The amino acid composition of the isolated peptides was determined with a Technicon TSM-1 auto-analyser, and their sequences studied by the combined dansyl–Edman technique [8, 9]. Automatic Edman degradation [10] was performed on a PS 100 SOCOSI Sequencer (St Maur, France). The identification of the PTH-amino acids was obtained by thin layer chromatography: the H and modified E systems were routinely used [11]. In some cases, gas chromatography was used to confirm the thin-layer chromatography results [12].

3. Results

3.1. *Isolation of the α- and β-chains*

The elution diagram of the two chains is illustrated in fig. 1; the separated chains were aminoethylated and purified as described above; the overall yield, starting from 400 mg of globin, is 80 mg for the α-chain and 90 mg for the β-chain.

3.2. *Finger-printing of the β-chain peptides and their amino acid content*

The tryptic hydrolysis of the β-chain leads to an almost completely water-soluble hydrolysate, which was resolved by electrochromatography into 17 distinct ninhydrin-positive spots (fig. 2). The presence of a very small insoluble fraction in the hydrolysate favours the idea that the whole β-chain has been split and that all the corresponding peptides are present in the electrochromatogram.

The specific staining gave the results indicated in table 1.

The amino acid composition of the 17 electrochromatographically-separated spots is given in table 2.

Spot 17 corresponds to Lys. The presence of only one Arg or Lys in their respective amino acid content proves that spots 1, 2, 3, 4, 6, 9, 11, 12, 14, 15, and 16 each contain a single peptide. Spot 13, containing

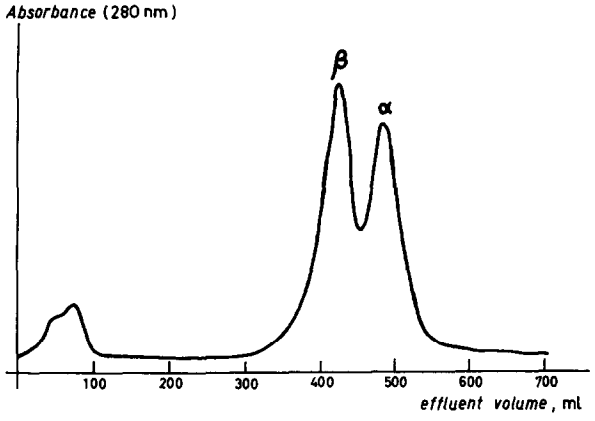

Fig. 1. Carboxymethylcellulose column chromatography of β-chains. Elution was performed with a gradient formed by mixing 500 ml of 0.005 M Na_2HPO_4 with 500 ml of 0.03 M Na_2HPO_4; the two buffers were in solutions of 8 M urea, 0.05 M 2-mercaptoethanol and adjusted to pH 7 with H_3PO_4.

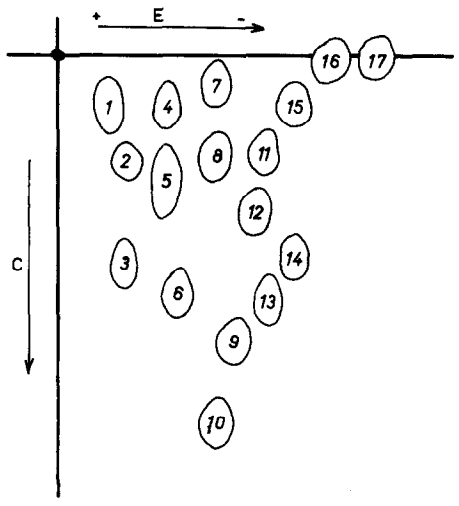

Fig. 2. Electrochromatography of the tryptic hydrolysate of the β-chain (Whatman 3 MM paper): E = 17 hr electrophoresis, 11 V/cm; pyridine–AcOH–H_2O (200:8:1800, vol/vol), pH 6.4; C = 23 hr descending chromatography, isoamyl-alcohol–pyridine–H_2O (6:6:7, vol/vol).

neither Arg nor Lys, appears as a C-terminal dipeptide. Spot 5, containing three Lys, is obviously impure. Paper chromatography using acetic acid *n*-butanol as solvent phase, separated this impure spot into a homogenous peptide 5a (Lys_1, Asp_3, Leu_2) and a mixture which was in turn fractionated by electrophoresis, giving two peptides: 5b (Lys_1 His_1 Asp_2 Glu_1 Pro_1 Val_1 Leu_1 Phe_1) and 5c (Lys_1 Thr_1 Glu_3 Pro_2 Ala_2 Val_1 Tyr_1 Phe_1). The very similar composition of spots 2 and 3 can be related to the same peptide initially present in the native β-chain; spot 2 then appears as an artefact

Table 1
Amino acid characterisation by specific staining.

Amino acid characterised	Procedure	Spot numbers
Arg	Sakaguchi phenanthrenequinone	1, 10
Trp	Ehrlich	9, 10
Tyr	α-Nitroso-β-naphthol	3, 5, 10, 13
His	Pauly	4, 5, 7, 12, 13, 14, 16
Sulphur amino acids	Platinic iodide	2, 3, 7, 10

Table 2
Amino acid composition of the spots eluted from the electrochromatogram.

Spots	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Lys		0.89 [1]	0.96 [1]	1.00 [1]	2.99 [3]	1.08 [1]	0.89 [1]	1.96 [2]	0.89 [1]		1.05 [1]	0.92 [1]		0.94 [1]	1.05 [1]	0.92 [1]	1.03 [1]
A.E. Cys							0.75 [1]			0.41							
SAE-Cys				0.87 [1]	0.94 [1]		0.95 [1]					0.90 [1]	0.96 [1]	1.89 [2]		0.89 [1]	
Arg	1.12 [1]									0.79 [1]							
Asp	2.01 [2]	3.25 [3]	3.01 [3]		4.73 [5]	1.07 [1]	1.06 [1]	0.98 [1]		0.45		0.93 [1]					
Thr		1.12 [1]	0.92 [1]	0.96 [1]	0.79 [1]				0.89 [1]	0.94 [1]	0.98 [1]						
Ser		1.85 [2]	1.96 [2]			1.89 [2]	1.10 [1]	1.85 [2]	1.96 [2]	0.35							
Glu	2.11 [2]	0.91 [1]	0.96 [1]	1.97 [2]	4.31 [4]	1.06 [1]	1.13 [1]	1.00 [1]		1.07 [1]							
Pro		1.85 [2]	1.92 [2]		2.85 [3]					0.96 [1]							
Gly	3.19 [3]	2.10 [2]	2.09 [2]			0.98 [1]		0.95 [1]	1.04 [1]	0.55	0.98 [1]	1.17 [1]		1.04 [1]		1.20 [1]	
Ala	1.10 [1]	1.04 [1]	1.06 [1]	1.02 [1]	1.99 [2]				0.93 [1]		0.92 [1]	4.22 [4]		1.12 [1]		1.11 [1]	
Val	2.93 [3]	1.00 [1]	1.04 [1]	1.09 [1]	1.75 [2]	1.10 [1]		0.95 [1]	1.00 [1]	2.61		2.91 [3]		1.04 [1]	1.05 [1]		
Met			0.74 [1]														
Leu	0.92 [1]	0.97 [1]	0.94 [1]	1.10 [1]	2.86 [3]	2.16 [2]	2.12 [2]	2.05 [2]	1.03 [1]	3.55		1.16 [1]		1.12 [1]			
Tyr			0.76 [1]		0.76 [1]					0.80 [1]			0.92 [1]				
Phe		2.12 [2]	2.12 [2]		1.84 [2]	1.04 [1]		1.07 [1]			1.03 [1]			1.01 [1]			
Total residues of amino acids	13		19	8	27	10	8	11	8		5	12	2	8	2	4	1
Trp									+	+							
N-terminal amino acid	Val	Phe	Tyr	Val		Val	Leu	Lys	Ser	Leu	Gly	Val	Tyr	Val	Val	Ala	Lys

Table

Comparison of the amino acid composition between

Human β-chain		Badger β-chain
Tryptic peptides	Amino acid composition	Spot number
HTp I	Lys_1 His_1 Thr_1 Glu_2 Pro_1 Val_1 Leu_1	4
HTp II	Lys_1 Thr_1 Ser_1 Gly_1 Ala_2 Val_1 Leu_1 Trp_1	9
HTp III	Arg_1 Asp_2 Glu_2 Gly_3 Ala_1 Val_3 Leu_1	1
HTp IV	Arg_1 Thr_1 Glu_1 Pro_1 Val_2 Leu_2 Tyr_1 Trp_1	10a
HTp V	Lys_1 Asp_3 Thr_1 Ser_2 Glu_1 Pro_2 Gly_2 Ala_1 Val_1 Met_1 Leu_1 Phe_3	
THp VI	Lys_1 Val_1	15
HTp VII	Lys_1 His_1 Gly_1 Ala_1	16
HTp VIII	Lys_1	17
HTp IX	Lys_1 His_1 Asp_3 Ser_1 Gly_2 Ala_2 Val_1 Leu_4 Phe_1	6 5a 6 + 5a
HTp X	Lys_1 SAE-Cys_1 His_1 Asp_1 Thr_2 Ser_1 Glu_1 Gly_1 Ala_2 Leu_2 Phe_3	7 11 7 + 11
HTp XI	His_1 Arg_1 Asp_2 Glu_1 Pro_1 Val_1 Leu_1 Phe_1	5b
HTp XIIa	SAE-Cys_1 Asp_1 Gly_1 Val_2 Leu_2	10b
HTp XIIb	Lys_1 His_2 Gly_1 Ala_1 Val_1 Leu_1 Phe_1	14
HTp XIII	Lys_1 Thr_1 Glu_3 Pro_2 Ala_2 Val_1 Tyr_1 Phe_1	5c
HTp XIV	Lys_1 His_1 Asp_1 Gly_1 Ala_4 Val_3 Leu_1	12
HTp XV	His_1 Tyr_1	13

The amino acid differences are underlined.

derived from the original peptide (spot 3): the sensitivity of Met residues is well known, and the observed sensitivity of the Tyr residue may be linked to its N-terminal position. Another similarity in the amino acid composition is observed with spots 6 and 8, the only difference being a further Lys residue in spot 8; since one of the two Lys in spot 8 has been proved to be N-terminal and since, furthermore, the N-terminal amino acid in spot 6 is Val, it appears that the original peptide is spot 8 with the N-terminal sequence Lys-Val, a peptide bond known to be only partially hydrolysed by trypsin. The occurence of both one Lys and one SAE-Cys in spot 7 is an indication of heterogeneity; however only Leu has been characterized by N-terminal amino acid determination. Spot 7 may therefore be considered as due to a single peptide in which the proximity of SAE-Cys and Lys prevents hydrolysis at the SAE-Cys position [1]. The same indication of heterogeneity appears in the case of spot 10, containing also two basic amino acids (Arg and SAE-Cys), despite the fact that Leu has been shown to be the only N-terminal amino acid. Further fractionation of spot 10 by electrophoresis at pH of 4.7 [11], led to two separated peptides: 10a (Arg_1 Thr_1 Glu_1 Pro_1 Val_2 Leu_2 Tyr_1 Trp_1), and 10b (SAE-Cys_1 Asp_1 Ser_1 Gly_1 Val_1 Leu_3). The amino acid composition of the entire contents of spot 10 shown in table 2, is explained by the fact that peptide 10b arises from a limited hydrolysis of a longer Cys-containing peptide, which represents the very small insoluble above-mentioned fraction. The location of tryptic peptides in the β-chain can be derived by comparing their amino acid composition with the amino acid composition of the tryptic peptides isolated from human haemoglobin β-chain. The following tentative sequence is proposed in table 3.

We must notice the entire similarity in composition between badger spots 1, 5a, 10a, 12, 13, 14, 15, 16, 17, and, respectively, human peptides THp III, HTp XIII, HTp IV, HTp XIV, HTp XV, HTp XIIb, HTpVI, HTpVII, and HTpVIII. The differences may concern several, but often only one, amino acids, as is

1
human tryptic peptides and badger fingerprint spots.

Amino acid composition	Proposed name for the tryptic peptides
Lys_1 His_1 Thr_1 Glu_2 Ala_1 Val_1 Leu_1	BTp I
Lys_1 Thr_1 Ser_2 Gly_1 Ala_1 Val_1 Leu_1 Trp_1	BTp II
Arg_1 Asp_2 Glu_2 Gly_3 Ala_1 Val_3 Leu_1	BTp III
Arg_1 Thr_1 Glu_1 Pro_1 Val_2 Leu_2 Tyr_1 Trp_1	BTp IV
Lys_1 Asp_3 Thr_1 Ser_2 Glu_1 Pro_2 Gly_2 Ala_1 Met_1 Val_1 Leu_1 Tyr_1 Phe_2	BTp V
Lys_1 Val_1	BTp VI
Lys_1 His_1 Gly_1 Ala_1	BTp VII
Lys_1	BTp VIII
Lys_1 Asp_1 Ser_2 Glu_1 Gly_1 Val_1 Leu_2 Phe_1	BTp IXa
Lys_1 Asp_3 Leu_2	BTp IXb
Lys_2 Asp_4 Ser_2 Glu_1 Gly_1 Val_1 Leu_4 Phe_1	
Lys_1 SAE-Cys_1 His_1 Asp_1 Ser_1 Glu_1 Leu_2	BTp Xa
Lys_1 Thr_1 Gly_1 Ala_1 Phe_1	BTp Xb
Lys_2 SAE-Cys_1 His_1 Asp_1 Thr_1 Ser_1 Glu_1 Gly_1 Ala_1 Leu_2 Phe_1	
Lys_1 His_1 Asp_2 Glu_1 Pro_1 Val_1 Leu_1 Phe_1	BTp XI
SAE-Cys_1 Asp_1 Ser_1 Gly_1 Val_1 Leu_3	BTp XIIa
Lys_1 His_2 Gly_1 Ala_1 Val_1 Leu_1 Phe_1	BTp XIIb
Lys_1 Thr_1 Glu_3 Pro_2 Ala_2 Val_1 Tyr_1 Phe_1	BTp XIII
Lys_1 His_1 Asp_1 Gly_1 Ala_4 Val_3 Leu_1	BTp XIV
His_1 Tyr_1	BTp XV

Table 4
Comparison of the amino acid sequence of N-terminal segments of human and badger β-chain.

```
                    1                   5                  10                  15
Human β-chain       Val–His–Leu–Thr–Pro–Glu–Glu–Lys–Ser–Ala–Val–Thr–Ala–Leu–Trp–
Badger β-chain      Val–His–Leu–Thr–Ala–Glu–Glu–Lys–Ser–Ala–Val–Thr–Ser–Leu–Trp–
                    ←———————— HTp I ————————→←———————— HTp II ————
                    ←———————— BTp I ————————→←———————— BTp II ————

                   16                  20                  25                  30
Human β-chain       Gly–Lys–Val–Asn–Val–Asp–Glu–Val–Gly–Gly–Glu–Ala–Leu–Gly–Arg–
Badger β-chain      Gly–Lys–Val–Asn–Val–Asp–Glu–Val–Gly–Gly–Glu–Ala–Leu–Gly–Arg–
                    ——→←————————————— HTp III ———————————→
                    ——→←————————————— BTp III ———————————→

                   31                  35                  40                  45
Human β-chain       Leu–Leu–Val–Val–Tyr–Pro–Trp–Thr–Gln–Arg–Phe–Phe–Glu–Ser–Phe–
Badger β-chain      Leu–Leu–Val–Val–Tyr–Pro–Trp–Thr–Gln–Arg–Tyr–Phe–  –  –  –
                    ←———————— HTp IV ————————→ ←————————————————
                    ←———————— BTp IV ————————→ ←————————————————

                   46                  50                  55                  59
Human β-chain       Gly–Asp–Leu–Ser–Thr–Pro–Asp–Ala–Val–Met–Gly–Asn–Pro–Lys.
Badger β-chain      –  –  –  –  –  –  –  –  –  –  –  –  –  –  –Lys.
                    ———————— HTp V ————————————————————→
                    ———————— BTp V ————————————————————→
```

The non identical residues of badger are underlined.

seen in the pairs 3-HTpV, 4-HTpI, 5b-HTp XI, (7+11)-HTp X, 9-HTp II, and 10b-HTp XIIa. In the pair (7+11)-HTp X substitution of a neutral residue by a lysine residue gives rise to an additional splitting which leads to the occurence of two peptides.

3.3. *The N-terminal sequence of the β-chain*

The next step in the elucidation of the primary structure of the β-chain is the determination of the amino acid sequence of the 19 analysed peptides. In this preliminary communication, we present only the results obtained for peptides BTp I, BTp II, BTp III, and BTp IV, (represented by spots 4, 9, 1 and 10a), which can be considered as the homologues of the first four tryptic peptides isolated from the human β-chain. The Edman degradation was carried out up to the last amino acid for peptide BTp I and for peptide BTp II. The stepwise degradation could only be followed up to position 7 for peptides BTp III and BTp IV. The following sequences may be postulated:

BTp I : Val–His–Leu–Thr–Ala–Glx–Glx–Lys.
BTp II: Ser–Ala–Val–Thr–Ser–Leu–Trp–Gly–Lys.
BTp III: Val–Asx–Val–Asx–Glx–Val–Gly–(Gly–Glx–Ala–Leu–Gly)–Arg.
BTp IV: Leu–Leu–Val–Val–Tyr–Pro–Trp–(Thr–Glx)–Arg.

The final confirmation of the above sequence, and the extension to position 42, was achieved through the use of the sequencer, and gave the results shown in table 4.

4. Conclusion

A comparison of the sequence of the badger haemoglobin β-chain up to position 42 with the corresponding fragment of the human haemoglobin β-chain shows only three one-base mutations due to transversion:

A_2 (position 5) Pro → Ala.
A_{10} (position 13) Ala → Ser.
C_7 (position 41) Phe → Tyr.

The mutation of position 5 is not of the conservative type.

The partial primary structure reported in this communication thus extends the comparative studies of the β-chain of haemoglobin to another zoological class, the carnivores. Since sequence studies undertaken in parallel on badger myoglobin are in progress [13, 14], we should be able in the near future to make broader phylogenetic comparisons of haemoglobins and myoglobins at the amino acid sequence level.

Acknowledgements

The animals, breeded in Centre d'études biologiques des animaux sauvages 79-CHIZE (France), have been kindly supplied by Professor Canivenc. We also thank Mrs Annick Larrue for her skillful technical help. The financial grant for the sequencer work was supplied by I.N.S.E.R.M. (Professor Dautrevaux).

References

[1] Lehmann, H. and Huntsman, R.G. (1966) Man's Haemoglobins, Vol. 1, p. 331, North-Holland, Amsterdam.
[2] Ducastaing, S. and Neuzil, E. (1969) Compt. Rend. Soc. Biol. 163, 2140.
[3] Rossi-Fanelli, A., Antonini, E. and Caputo, A. (1958) Biochim. Biophys. Acta 30, 608.
[4] Clegg, J.B., Naughton, M.A. and Weatherall, D.J. (1966) J. Mol. Biol. 19, 91.
[5] Adams, H.R., Wrighstone, R.N., Miller, A. and Huisman, T.H.J. (1969) Arch. Biochem. Biophys. 132, 223.
[6] Biserte, G., Plaquet-Schoonaert, T., Boulanger, P. and Paysant, P. (1960) J. Chromato. 3, 25.
[7] Easley, C.W., Zegers, B.J.M. and De Vijlder, M. (1969) Biochim. Biophys. Acta 175, 211.
[8] Gray, W.R. and Hartley, B.S. (1963) Biochem. J. 89, 59.
[9] Hartley, B.S. (1970) Biochem. J. 119, 805.
[10] Edman, P. and Begg, C. (1967) Eur. J. Biochem. 1, 80. .
[11] Needleman, S.B. (1970) Protein Sequence Determination, p. 223, Springer Verlag, Berlin.
[12] Pisano, J.J. and Bronzert, T.J. (1969) J. Biol. Chem. 244, 5597.
[13] Tetaert, D., Han, K.K., Dautrevaux, M., Ducastaing, S., Hombrados, I. and Neuzil, E. (1973) FEBS Letters 29, 38.
[14] Han, K.K., Tetaert, D., Ducastaing, S. and Dautrevaux, M., 9th Internatl. Congr. of Biochem., Stockholm, July, 1973, in press.