N-TERMINAL CHAIN ELONGATION AS EVIDENCE FOR DUPLICATION OF MYOGLOBIN IN THREE SOUTH AMERICAN MONKEYS

A.E. ROMERO HERRERA and H. LEHMANN

Medical Research Council Abnormal Haemoglobin Unit, University of Cambridge, Department of Biochemistry, Cambridge, Great Britain

Received 20 February 1973

1. Introduction

The question whether mammalian myoglobin consists of major and minor components has not been satisfactorily resolved. For example, the claim that in the newborn the adult myoglobin was accompanied by a foetal myoglobin has not been substantiated: it is now known that this was in fact foetal haemoglobin. One minor haemoglobin differing from myoglobin and haemoglobin electrophoretically, occasionally found in human muscle extract, is the result of the removal of the C-terminal residue of arginine of the α-chain of human haemoglobin (des-Arg Hb A or Hb Köelliker) [1]. Another "minor fraction" results from the attachment of small amounts of myoglobin to another protein. Evidence has now come forward, however, that in three different South American monkeys, whose myoglobin has been described elsewhere [2], Humboldt's woolly monkey (*Lagothrix lagothricha*), the squirrel monkey (*Saimiri sciureus*), and the common marmoset (*Callithrix jacchus*), there are indeed two different myoglobins present. They do not differ from each other in electrophoretic mobility and could not be separated by chromatography. However, on peptic digestion of the insoluble tryptic core, it was found that there were two different tryptic peptides, one representing residues 1–16 and another one with residues 1–16 plus two additional residues at the N-terminus.

2. Material and methods

800 g of muscle from one woolly monkey, 200 g from one squirrel monkey and 160 g pooled from four marmosets were the materials used. The muscle was minced, homogenised with 1.5 mM KCN (1 g tissue per ml), and filtered (Whatman No. 1.). The residue was extracted twice more as above and the combined filtrates were mixed with saturated $(NH_4)_2SO_4$ (4:6, v/v) and, after standing at room temp. for 1 hr, the precipitate was filtered off (Whatman No. 1.). The filtrate was concentrated by pressure filtration using an Amico PM30 membrane, applied to a Sephadex G-25 column (2.5 cm × 30 cm) and eluted with 0.05 M Tris-HCl buffer, pH 8.5. The eluate was then submitted to DEAE Sephadex A-50 column chromatography [3] (see fig. 1). In the case of the squirrel monkey and the marmosets, the myoglobin fraction was purified further by repeated discontinuous paper electrophoresis (Whatman No. 4.) for 40 hr at 150 V with Tris–EDTA–boric acid buffer (25.2:2.5:1.9, g/ℓ) pH 9.1 at the anode, and barbital buffer (sodium diethylbarbiturate–diethylbarbituric acid, 5.15:0.92, g/ℓ) pH 8.6 at the cathode [4]. To establish the purity of the preparation a strip of paper was stained with Light Green and the excess stain washed out with 2.5% acetic acid. The myoglobin band was cut out, eluted from the paper with 1.5 mM KCN at 4°, and concentrated to an approx. 10% solution in a Sartorius collodion bag. The purity of the myoglobin was confirmed by fingerprinting of an aliquot as described below, to ascertain whether only peptides belonging to myoglobin were present. In the case of the woolly monkey, the

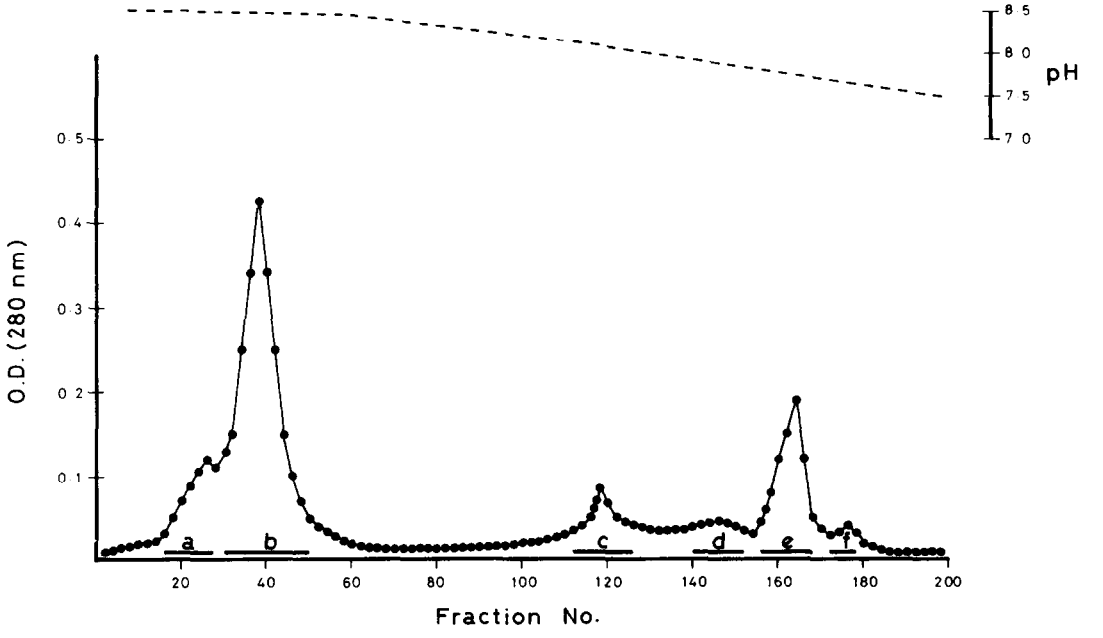

Fig. 1. Woolly monkey: elution pattern of the muscle extract on DEAE Sephadex A-50 medium column (2.5 cm × 85 cm) at 4°, using a linear gradient pH 8.5 – 6.0 (0.05 M Tris-HCl buffer). Fractions of 12 ml were collected at a flow rate of 40 ml/hr and pooled as indicated by the bars. a,b,c: Myoglobin plus other contaminating proteins; d: protein other than myoglobin; e,f: major and minor haemoglobin components.

myoglobin was purified further after removal of the haem (see below), by column chromatography on CM23-cellulose [5] (see fig. 2). It was then lyophilised and dissolved in 70% formic acid to which an excess of cyanogen bromide was added [6], incubated at room temp. for 24 hr and dried in a rotary evaporator. The unreacted myoglobin and the CNBr fragments were separated by gel filtration on Sephadex G-75 (see fig. 3).

The haem group was removed from 10% solutions of purified myoglobin by adding it drop-wise to 20 vol of concentrated HCl in acetone (1.5%, v/v) at −20° [7]. After centrifugation the superantant was removed; the precipitated globin was then washed 3 times with acetone (−20°) to eliminate traces of acid, and dried under nitrogen. The globin was then digested with trypsin and the soluble peptides divided into aliquots of 1.5 mg and 5 mg (starting material) and separated by two dimensional high-voltage paper electrophoresis [9] (Whatman No. 3, pH 6.5) and chromatography [10] to give diagnostic and preparative "fingerprints". The first were developed with 0.2% and the second with 0.02% ninhydrin in acetone (w/v). The diagnostic fingerprints were stained for methionine [11], arginine [12], histidine [13], tyrosine [14] and tryptophan [15]. In addition diagnostic fingerprints were chlorinated to exclude blocked N-terminals.

The insoluble core which remained after tryptic digestion was washed twice with water and resuspended in 0.1 M HCl (1 ml for each 10 mg of starting material) and hydrolysed with 0.05 ml of pepsin solution (2 mg/ml in water w/v) at 37° for 3 hr. The reaction was stopped by lyophilization and the resulting peptides fingerprinted as described above. For amino acid analysis, peptides were eluted from preparative fingerprints [16] with 6 N HCl and hydrolysed at 108° for 24 hr; for sequential dansyl-Edman degradation [17] they were eluted with 0.5 M NH_4OH and dried in a rotary evaporator. Purified woolly monkey myoglobin (fig. 3, peak I – the whole molecule) was submitted to dansylation [18], and after acid hydrolysis, the dansyl-derivatives were purified by high voltage paper electrophoresis, pH 4.4 and pH 2.0, eluted, and identified using polyamide thin-layer chromatogtography

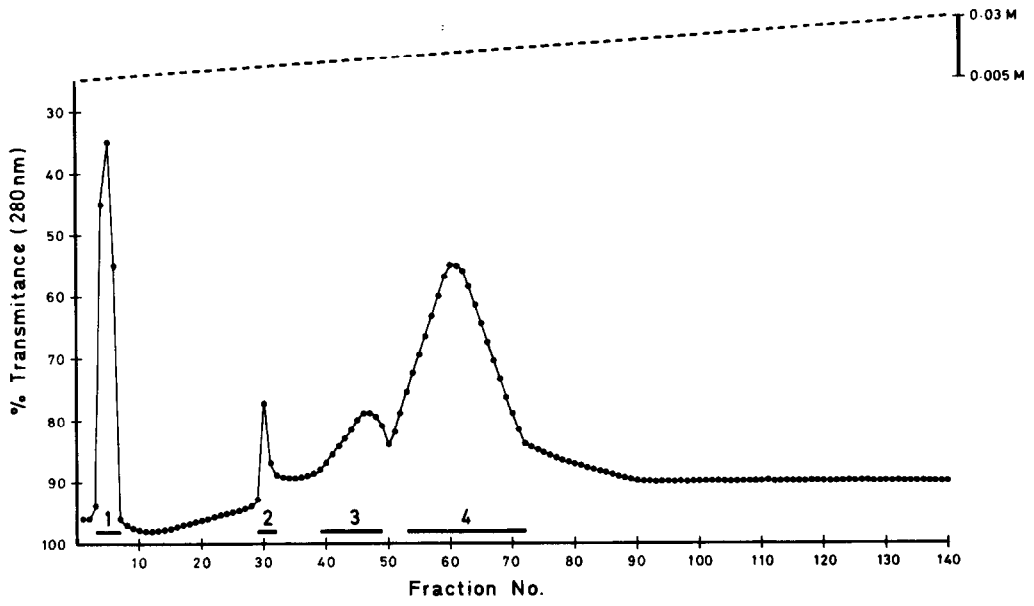

Fig. 2. Woolly monkey: elution pattern of the main myoglobin fraction (peak b, fig. 1) after removal of the haem on CM23-cellulose column (1.5 cm × 28 cm) at room temp. Buffers were prepared in 8 M urea–2 mercaptoethanol (0.05 M–0.003 M Na_2HPO_4–2 H_2O) and adjusted to pH 6.7 with H_3PO_4. 5 ml fractions were collected at a flow rate of 60 ml/hr, and fractions indicated by bars were pooled. 1,2,3: Proteins other than myoglobin; 4: myoglobin.

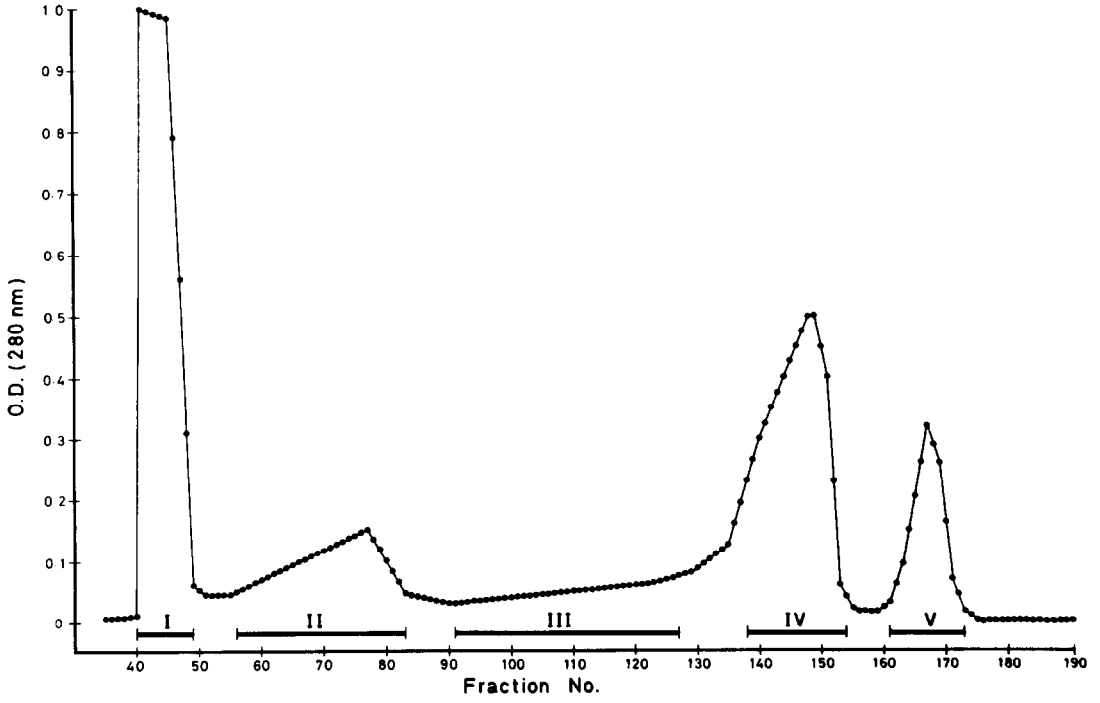

Fig. 3. Woolly monkey: elution pattern of the myoglobin CNBr peptides on Sephadex G-75 medium column (2.5 cm × 90 cm × 2) in 0.2 M acetic acid at 4°. Fractions of 5 ml were collected at a flow rate of 30 ml/hr and pooled as indicated by the bars, freeze-dried and fingerprinted after tryptic digestion. I: residues 1–153 (the whole unreacted molecule); II: residues 1–31; III residues 56–131; IV residues 1–55; V: residues 132–142 and 143–153.

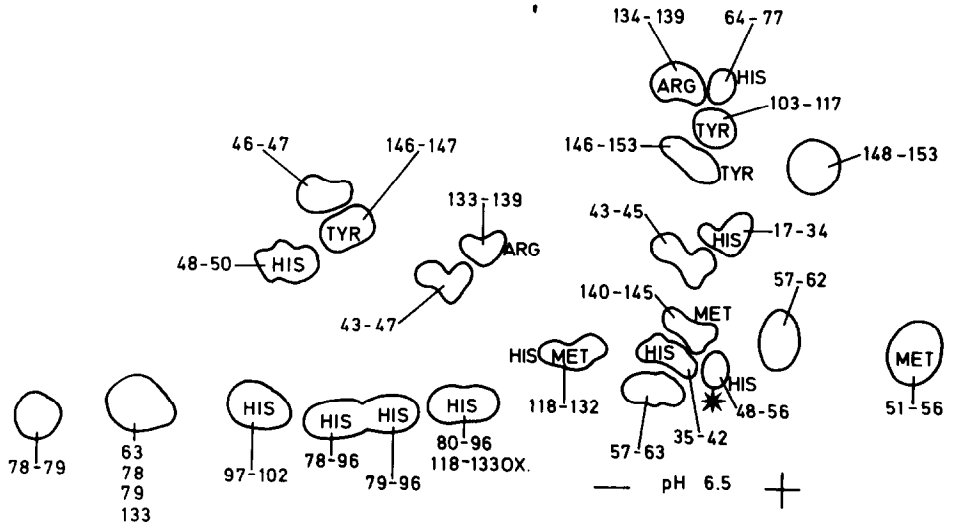

Fig. 4. Woolly monkey: fingerprint of the myoglobin soluble tryptic peptides. The staining reactions are indicated. * is the point of application. The fingerprints of the tryptic peptides from the myoglobins of the squirrel monkey and marmoset were similar.

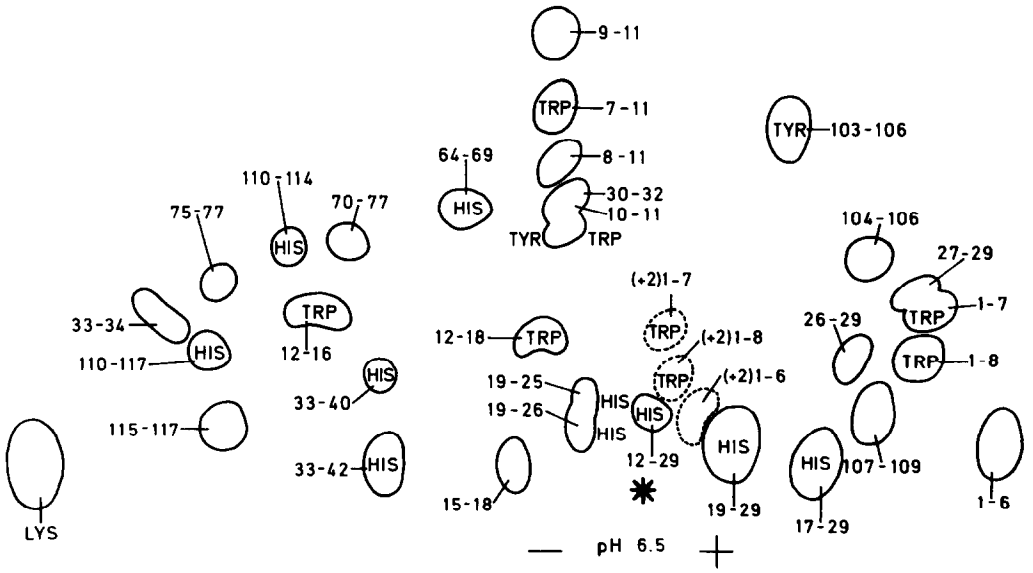

Fig. 5. Woolly monkey: fingerprint of the peptic peptides from the insoluble and partially soluble core of myoglobin left after tryptic digestion. The staining reactions are indicated. * is the point of application. The broken circles indicate the position of the 3 new peptides (see text). The fingerprints of the peptic peptides from the insoluble and partially soluble tryptic core of the squirrel monkey and marmoset were similar.

Table 1

	Lagothrix lagothricha (woolly monkey)							*Callithrix jacchus* (common marmoset)						
	1–6	(+2)1–6	1–7	(+2)1–7	1–8	(+2)1–8		1–6	(+2)1–6	1–7	(+2)1–7	1–8	(+2)1–8	
Aspartic acid	1.07 (1)	1.06 (1)	1.15 (1)	0.91 (1)	1.10 (1)	1.06 (1)		0.99 (1)	1.06 (1)	1.02 (1)	1.02 (1)	1.08 (1)	0.97 (1)	
Serine	1.00 (1)	1.08 (1)	0.97 (1)	0.89 (1)	1.03 (1)	0.93 (1)		1.05 (1)	0.94 (1)	1.00 (1)	0.93 (1)	1.06 (1)	0.90 (1)	
Glutamic acid	1.08 (1)	1.08 (1)	1.06 (1)	1.25 (1)	1.97 (2)	1.82 (2)		1.00 (1)	1.05 (1)	1.04 (1)	1.05 (1)	1.90 (2)	2.13 (2)	
Glycine	1.75 (2)	1.94 (2)	1.63 (2)	1.82 (2)	1.75 (2)	2.01 (2)		2.01 (2)	2.02 (2)	1.90 (2)	2.02 (2)	1.92 (2)	2.00 (2)	
Leucine	1.10 (1)	0.99 (1)	1.19 (1)	1.11 (1)	1.15 (1)	0.96 (1)		0.95 (1)	1.02 (1)	1.04 (1)	1.08 (1)	1.04 (1)	1.00 (1)	
Phenylalanine		0.44 (1)		0.50 (1)		0.43 (1)			0.60 (1)		0.55 (1)		0.58 (1)	
Lysine		0.85 (1)		1.02 (1)		0.82 (1)			0.91 (1)		0.90 (1)		1.00 (1)	
Tryptophan			* (1)	* (1)	* (1)	* (1)				* (1)	* (1)	* (1)	* (1)	
Yield in μmoles per one residue	0.044	0.021	0.050	0.008	0.058	0.016		0.040	0.021	0.098	0.024	0.015	0.018	

The expected number of residues are in parentheses.

* Peptides fluoresced under the UV light and stained with Ehrlich reagent.

3. Results

Figs. 4 and 5 show two fingerprints from the woolly monkey. The first accounts for the soluble tryptic peptides and the second for peptic peptides from the insoluble tryptic peptides residues 1–16 and 64–77, and the partially soluble tryptic peptides residues 17–34, 17–42 and 103–117. Three additional peptides, two of them staining for tryptophan were present in the peptic fingerprints of the three South American monkeys which had not been found in the other primates so far investigated. On amino acid analysis (see table 1) it was seen that they were companion peptides to the peptic peptides residues 1–6, 1–7 and 1–8. They differed from these in each case by additional residues of phenylalanine and lysine. Dansyl-Edman degradation showed these to be at the N-terminus (see table 2). The proportions of the minor fractions were 8% in the woolly monkey and 14% in the marmoset. Although there was not enough material to determine the proportion in the squirrel monkey, it seemed to be of about the same order. When 3 mg of the whole myoglobin molecule of the woolly monkey were dansylated, two N-terminal residues were identified: DNS-glycine and DNS-phenylalanine in approximately the same proportion mentioned above.

4. Discussion

Although the minor myoglobin fractions contain an additional residue of lysine which alters the mobility of the peptic peptides, the whole myoglobin molecule consistently formed a single fraction, by the numerous methods described above. One possible explanation is that somewhere along the polypeptide chain another mutation has occurred changing a neutral charge to a negative one, or a positive one to neutral. Close investigation has so far not yielded any evidence for this, and one has to consider a mutation which would not alter the fingerprint or the amino acid analysis, for instance if one of the lysines in the sequences 61–64 (Leu–Lys–Lys–His), 76–80 (Leu–Lys–Lys–Lys–Gly) or 131–134 (Met–Lys–Lys–Ala) has been deleted, this would be difficult to establish with the low proportion of the minor component. Further investigation is still proceeding.

Table 2

Residue no.			1	2	3	4	5	6	7	8	
Woolly monkey Mb major fraction			$\overrightarrow{\text{Gly}}$	$\overrightarrow{\text{Leu}}$	$\overrightarrow{\text{Ser}}$	$\overrightarrow{\text{Asp}}$	$\overrightarrow{\text{Gly}}$	Glu ↑	Trp ↑	Gln ↑	... etc.
Woolly monkey Mb minor fraction	$\overrightarrow{\text{Phe}}$	$\overrightarrow{\text{Lys}}$	$\overrightarrow{\text{Gly}}$	$\overrightarrow{\text{Leu}}$	$\overrightarrow{\text{Ser}}$	(Asp	Gly)	Glu ↑	Trp ↑	Gln ↑	... etc.
Marmoset Mb major fraction			$\overrightarrow{\text{Gly}}$	$\overrightarrow{\text{Leu}}$	$\overrightarrow{\text{Ser}}$	$\overrightarrow{\text{Asp}}$	$\overrightarrow{\text{Gly}}$	Glu ↑	Trp ↓↑	Gln ↑	... etc.
Marmoset Mb minor fraction	$\overrightarrow{\text{Phe}}$ ↓	$\overrightarrow{\text{Lys}}$	$\overrightarrow{\text{Gly}}$	$\overrightarrow{\text{Leu}}$	$\overrightarrow{\text{Ser}}$	(Asp	Gly)	Glu ↑	Trp ↓↑	Gln ↑	... etc.

→ Dansyl-Edman degradation. ↑: Peptic hydrolysis. ↓: Chymotryptic hydrolysis.

In the marmoset chymotryptic digestion was performed in addition, and a peptide containing residues 1–7 plus one of lysine was obtained.

These observations suggest that there are two different molecules of myoglobin in some South American monkeys. The most likely explanation is a mutation of the initiating methionine codon (AUG) [19] to one for lysine (AAG); presumably the third codon from the N-terminus is again one for methionine. Nothing is known of the myoglobin mRNA, but in the duck the mRNA for haemoglobin consists of about 650 nucleotides [20]. Assuming that globin biosynthesis is generally similar, of this number, 465 nucleotides would code for the 153 amino acids of myoglobin, plus the initiating and terminating codons and another 100 for poly A. This leaves approx. 85 nucleotides unaccounted for, and the two additional residues at the N-terminus described here may indicate the nature of nucleotides in the mRNA which precede those involved in the biosynthesis of myoglobin.

Acknowledgements

We want to thank Dr. Ralph Heywood of the Huntingdon Research Centre, England, for primate material. A.E.R.H. holds a scholarship from the "Instituto Nacional de Energia Nuclear", (Mexico) and the British Council.

References

[1] H.R. Marti, D. Beale and H. Lehmann, Acta Haemat. 37 (1967) 174.
[2] A.E. Romero Herrera and H. Lehmann, submitted for publication to Biochim. Biophys. Acta.
[3] T.H.J. Huisman and A.M. Dozy, J. Chromatogr. 19 (1965) 160.
[4] J.L. Graham and B.W. Grunbaum, Am. J. Clin. Pathol. 39 (1963) 567.
[5] J.B. Clegg, M.A. Naughton and D.J. Weatherall, J. Mol. Biol. 19 (1966) 91.
[6] E. Gross and B. Witkop, J. Biol. Chem. 237 (1962) 237.
[7] V.M. Ingram and A.O.W. Stretton, Biochim. Biophys. Acta 62 (1962) 456.
[8] K. Sick, D. Beale, D. Irvine, H. Lehmann, P.T. Goodall and S. MacDougal, Biochim. Biophys. Acta 140 (1967) 231.
[9] V.M. Ingram, Biochim. Biophys. Acta 28 (1958) 539.
[10] C. Baglioni, Biochim. Biophys. Acta 48 (1961) 392.
[11] G. Toennies and J.J. Kolb, Anal. Chem. 23 (1951) 823.
[12] J.B. Jepson and I. Smith, Nature 172 (1953) 1100.
[13] H.K. Berry, H.E. Sutton, L. Cain and J.S. Berry, Univ. Texas Publ. 5169 (1951) 30.
[14] R. Archer and C. Crocker, Biochim. Biophys. Acta 9 (1952) 704.
[15] I. Smith, Nature 171 (1953) 43.
[16] F. Sanger and H. Tuppy, Biochem. J. 49 (1951) 463.
[17] W.R. Gray, in: Methods in enzymology, Vol. XI, ed. C.H.W. Hirs (Academic Press, New York and London, 1967) p. 469.
[18] K. Narita, in: Protein sequence determination, ed. S.B. Needleman (Chapman & Hall Ltd., London; Springer Verlag, Berlin, Heidelberg, New York, 1970) p. 47.
[19] A.R. Hunter and R.J. Jackson, European J. Biochem. 19 (1971) 316.
[20] R.E. Pemberton and C. Baglioni, J. Mol. Biol. 65 (1972) 531.