STUDIES ON MARSUPIAL PROTEINS

III.* *N*-BROMOSUCCINIMIDE CLEAVAGE OF THE α-CHAIN OF KANGAROO HAEMOGLOBIN AND THE AMINO ACID SEQUENCE OF THE *N*-TERMINAL FRAGMENT

By J. M. Beard† and E. O. P. Thompson†

[*Manuscript received July* 17, 1969]

Summary

The tryptophan contents of the α- and β-chains of haemoglobin from the grey kangaroo *Macropus giganteus* were determined by *N*-bromosuccinimide titration to be 1 and 2 residues respectively. Following cleavage of the tryptophyl bond in the α-chain with *N*-bromosuccinimide in 8M urea solutions at pH 3·5 the *N*-terminal fragment was purified by gel filtration and paper ionophoresis.

The amino acid sequence of this fragment was determined by the "dansyl"–Edman method and confirmed by the isolation and sequencing of tryptic peptides from this fragment. The sequence of the first 16 residues of the α-chain is

Val-Leu-Ser-Ala-Ala-Asp-Lys-Gly-His-Val-Lys-Ala-Ile-Trp-Gly-Lys-

I. Introduction

One approach to the determination of the amino acid sequence of a polypeptide chain is to fragment the chain by chemical methods selective for residues present in relatively small amounts (Witkop 1968). *N*-Bromosuccinimide (NBS) was introduced by Patchornik, Lawson, and Witkop (1958*a*) as a reagent for the estimation of the tryptophan content of proteins by an oxidative reaction that also resulted in cleavage of the peptide chain (Patchornik, Lawson, and Witkop 1958*b*). This reaction has been extensively used for the estimation of the tryptophan content of proteins (Spande and Witkop 1967) and for determining the amino acid following the tryptophan residue (Ramachandran and Witkop 1959). NBS attacks bonds other than those involving tryptophan, e.g. histidyl and tyrosyl bonds, but Funatsu, Green, and Witkop (1964) reported that in the presence of 8M urea the reaction is selective for tryptophyl bonds. Fragments resulting from this selective cleavage have, however, only been isolated from glucagon (Patchornik *et al.* 1960) and bovine globin (Sasakawa 1963).

This reagent has now been applied to the estimation of the tryptophan content of the α- and β-chains of haemoglobin of the kangaroo *Macropus giganteus* and to isolation and sequencing of the *N*-terminal fragment from the α-chain.

* Part II, *Aust. J. biol. Sci.*, 1969, **22**, 1437–54.

† School of Biochemistry, University of New South Wales, P.O. Box 1, Kensington, N.S.W. 2033.

II. Materials and Methods

The α-chain of human globin was prepared by the method of Clegg, Naughton, and Weatherall (1965). The α- and β-II globin chains of kangaroo haemoglobin were prepared as described by Thompson, Hosken, and Air (1969) using a similar method. NBS (Hopkins and Williams Ltd.) was recrystallized from water.

The methods of tryptic digestion, high voltage paper ionophoresis, peptide mapping, and amino acid analysis were the same as previously described (Thompson, Hosken, and Air 1969). Peptide degradation by the "dansyl"–Edman procedure followed the description by Gray (1967) with the modifications described by Air and Thompson (1969).

(*a*) *Determination of Tryptophan*

The tryptophan contents of the α- and β-chains were determined by titration with 0·01M NBS using the procedure of Ramachandran (1962). Both the protein and reagent were dissolved in 8M urea–acetic acid, pH 4·0. In all estimations the amount of globin used was computed from the phenylalanine content of a quantitative amino acid analysis of an aliquot which had been hydrolysed in 6N HCl at 110°C for 24 hr. The hydrolysate was analysed on the 60-cm column of a Beckman 120C amino acid analyser eluted directly with pH 4·25 buffer. In this way problems due to contamination with urea and non-protein impurities which prevented accurate weighing were avoided. Both the α- and β-globins contain 8 residues of phenylalanine.

Alkaline hydrolysis for liberation of tryptophan was done by heating with 3·75N NaOH in an oxygen-free system at 110°C for 16 hr using a Nalgene polypropylene tube sealed in a glass tube. The hydrolysate was acidified with 6N HCl and 1·1M citric acid containing 5% thiodiglycol, and analysed as described by Neumann, Moore, and Stein (1962). A freshly prepared tryptophan standard was used.

(*b*) *Selective Cleavage of α-Globin with NBS*

A solution of 10M urea, deionized on a mixed-bed ion-exchange column, was diluted to 8M with 90% acetic acid to give a pH of 3·5. The α-globin (10 mg/ml) was dissolved in this buffer and incubated with a 10 or 50 molar excess of NBS for 15 min at room temperature. After reaction the excess NBS was destroyed by addition of 10% by volume of 90% formic acid (Shaltiel and Patchornik 1963).

When the *N*-terminal amino acids present in the reaction mixture were to be determined urea was removed by gel filtration through a column of G-10 Sephadex (3 by 37 cm) equilibrated and eluted with 10% formic acid and the whole peptide fraction lyophilized. The *N*-terminal amino acids present were determined as the phenylthiohydantoins (PTH) using the procedure of Blömback *et al.* (1966) after coupling in 60% pyridine–water containing 6% *N*-ethylmorpholine. The PTH-derivatives were examined by thin-layer chromatography on Eastman precoated plates (Jeppsson and Sjöquist 1967) and by regeneration of the amino acids by heating with 6N HCl at 150°C for 24 hr (Van Orden and Carpenter 1964).

For purification of the *N*-terminal fragment the lyophilized fraction was dissolved in 10% formic acid, loaded onto a column of G-50 Sephadex (0·9 by 130 cm), equilibrated, and eluted with 10% formic acid (16 ml/hr). Fractions at equal time intervals were collected and measured at 250 nm (the absorption maximum of the oxidized tryptophan). It was possible to load the reaction mixture directly onto G-50 Sephadex and remove the urea at the same time as fractionating the peptide material. The fractions containing the *N*-terminal fragment were pooled and further purified by paper ionophoresis at pH 6·4 (3000 V, 1 hr). A guide strip was stained with ninhydrin and the two major bands eluted with 0·01M NH_3 and pooled.

(*c*) *Isolation of the Tryptophan-containing Tryptic Peptide from α-Globin*

A tryptic digest of whole α-globin was fractionated by paper ionophoresis at pH 6·4 followed by chromatography in butanol–pyridine–acetic acid–water (15 : 10 : 3 : 12 v/v) as described by Thompson, Hosken, and Air (1969). The tryptophan-containing peptide was detected with Ehrlich's reagent (Smith 1953). Isolation of the peptide from successive one-dimensional separations using a guide strip stained in this way to locate the peptide gave a sample which was recovered after elution with 0·01M NH_3.

III. Results

(*a*) *Tryptophan Content of Kangaroo Globins*

The spectral changes obtained on titration of the globins with NBS were similar to those recorded by other workers both with model peptides (Patchornik *et al.* 1960) and with proteins (Peters 1959). There was a decrease in extinction at 280 nm with an increase at 250 nm due to oxidation of tryptophyl residues to oxindole derivatives. Using the extinction correction factor of 1·31 (Spande and Witkop 1967) the tryptophan contents of the α- and β-II chains of kangaroo globin were 1·1 and 2·1 residues per mole of globin respectively (mean of three separate titrations). Alkaline hydrolysis of the α-chain gave 1·0 residues of tryptophan per mole. These are identical to the content in human globin chains.

During titration the drop in extinction at 280 nm after each addition of NBS was found to be instantaneous and stable for up to 30 min as reported by Rao and Cama (1963). For certain other proteins 10–20 min has been required for complete oxidation (Peters 1959; Kronman, Robbins, and Andreotti 1967).

Table 1

N-terminal residues in *N*-bromosuccinimide (NBS)-treated α-globins

Globins were treated with 10 or 50 moles NBS per mole of protein in 8M urea–acetic acid, pH 3·5 or 4·0. Urea was removed by gel filtration and the protein coupled with phenylisothiocyanate. After cyclization the phenylthiohydantoins were regenerated with 6N HCl at 150°C for 24 hr. Amino acids were quantitatively estimated by automatic amino acid analysis. Values are given as moles per mole of valine, which is *N*-terminal in the globin chains, and are uncorrected for losses during regeneration

Amino Acid	Unoxidized Kangaroo α-Globin	Oxidized Kangaroo α-Globin		Oxidized Human α-Globin	
		10 moles NBS*	50 moles NBS†	10 moles NBS†	50 moles NBS†
Lysine		0·3	0·6		
Aspartic acid			0·4	0·9	0·5
Glutamic acid		0·2	0·3	0·4	0·2
Proline				0·3	0·3
Glycine	0·1	0·6	0·8	1·3	1·5
Alanine		0·5	1·0	1·3	0·8
Valine	1·0	1·0	1·0	1·0	1·0
Isoleucine			0·2		
Leucine		0·2	0·5	0·6	0·4
Phenylalanine			0·4	0·3	0·3

* pH 3·5. † pH 4·0.

(*b*) *Selective Cleavage of the Tryptophyl Bond in Kangaroo α-Globin*

Experiments with the α-chain of human globin and the α-chain of kangaroo globin at pH 4 for the cleavage in 8M urea solutions were followed by *N*-terminal group estimations using the Edman method. The results are given in Table 1 and it is seen that the reaction is not specific for tryptophyl bonds. In addition to glycine, arising from the specific cleavage of the tryptophylglycyl bond, several other *N*-terminal amino acids were found in NBS-treated human α-globin indicating that

non-specific cleavages had also occurred. For isolation of the *N*-terminal fragment from kangaroo α-globin, a slightly lower pH of 3·5 was used. This gave greater selectivity although the yield of *N*-terminal glycine was somewhat reduced.

(*c*) *Purification of* N-*Terminal Fragment*

The fractionation of the NBS-treated α-globin on G-50 Sephadex is shown in Figure 1. The elution was followed at 250 nm as this is the wavelength of maximum

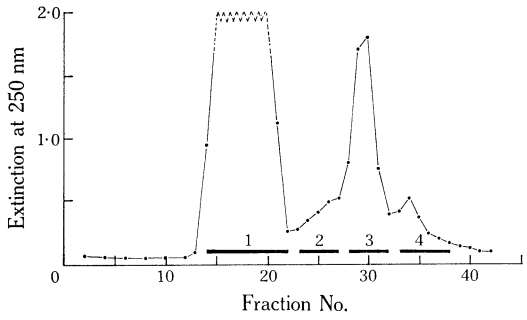


Fig. 1.—Gel filtration in 10% formic acid on a Sephadex G-50 column (138 cm by 1 cm diam.) of a sample of kangaroo α-globin treated with NBS in 8M urea at pH 3·5. The urea was first removed on a column of G-10 Sephadex. Fraction size was *c.* 3 ml, and fractions pooled for analysis are indicated by bars.

absorption of the dioxindole formed after cleavage of tryptophyl bonds. The fractions were pooled as shown. Fraction 1 contained any uncleaved globin and the large

TABLE 2

AMINO ACID COMPOSITION OF THE *N*-TERMINAL FRAGMENT OF KANGAROO α-GLOBIN

Samples hydrolysed in 6N NCl at 110°C for 18–24 hr. Values are given as moles of amino acid per mole of peptide, setting aspartic acid as one residue

Amino Acid	Experiment 1		Experiment 2: Combined Preparation†
	Acid*	Lactone*	
Lysine	1·7	1·8	1·9
Histidine	0·9	1·1	1·0
Aspartic acid	1·0	1·0	1·0
Serine	1·0	1·0	0·9
Glycine	1·3	1·2	1·3
Alanine	3·2	3·1	3·0
Valine	1·3*	1·3*	1·9
Isoleucine	0·8	0·9	0·7
Leucine	0·8	0·8	1·0

* Recovery of *N*-terminal valine low due to staining with ninhydrin for detection and a time of hydrolysis of only 18 hr.

† Combined eluates of lactone and acid forms. Not stained with ninhydrin and hydrolysed 24 hr.

fragment (residues 15–141) resulting from cleavage. Fraction 3 contained the *N*-terminal fragment. The other fractions were mainly unidentified fragments resulting

from non-specific cleavage. Fraction 4 was shown to include free arginine by paper ionophoresis at pH 1·9 and this probably results from cleavage of the ultimate tyrosylarginine bond in the chain.

Paper ionophoresis of fraction 3 gave two major bands. These were eluted separately with 0·01M NH_3 and found to have the same amino acid composition (Table 2) and identical *N*-terminal sequences, Val-Leu-Ser-. In subsequent fractionations the material eluted from these two bands was combined as they represent lactone and free acid forms of the same peptide.

The composition of the *N*-terminal fragment was

$$Asp_1Ser_1Gly_1Ala_3Val_2Ile_1Leu_1Lys_2His_1$$

plus the modified tryptophan residue, which would place the tryptophan residue at position 14 as in human α-globin. The yield of isolated *N*-terminal fragment averaged

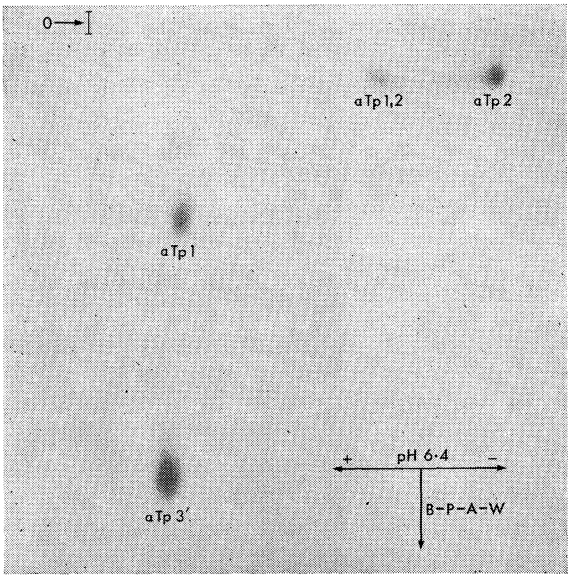


Fig. 2.—Peptide map of the *N*-terminal fragment of the α-chain of kangaroo haemoglobin digested with TPCK-trypsin. Partial splitting of one bond has resulted in three peptides labelled αTp1,αTp2, and αTp1,2 normally present in digests of α-globin. αTp3′ is the *C*-terminal peptide in the acid form of the modified tryptophan residue.

20% estimated from the amino acid analysis of the eluted material. This is considerably less than the yield expected from the *N*-terminal glycine found by direct determination (Table 1) but it is known, e.g. Bennett (1967), that recoveries of peptides by elution from paper ionograms can be relatively low.

(*d*) *Sequence of the* N-*Terminal Fragment*

The sequence determined by the dansyl–Edman procedure was definite for eight consecutive steps with some uncertainty with regard to residue 4, which could have been either histidine or alanine, both of which tend to give a product

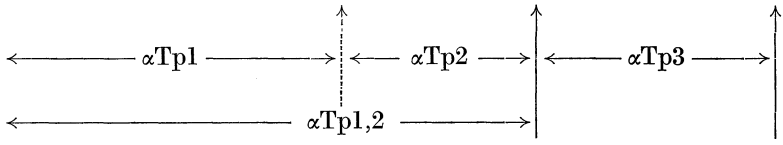
running with the dansylamide spot (Air and Thompson 1969). That it was not histidine was confirmed by subtractive analysis after steps 3, 4, and 5 which showed no significant loss of histidine but losses of alanine for residues 4 and 5. The sequence was

Val-Leu-Ser-Ala-Ala-Asx-Lys-Gly-

The remainder of the sequence and the identification of residue 6 as aspartic acid was based on the tryptic peptides isolated from the *N*-terminal fragment and the tryptophan-containing peptide in tryptic digests of the α-chain. Figure 2 shows a peptide map of a tryptic digest made at 37°C for 4 hr. There is incomplete splitting of a bond involving -Lys-Gly- due to the preceding aspartic acid residue. Thus peptides αTp1, αTp2, αTp1,2 and the *C*-terminal peptide containing the oxidized tryptophan residue were obtained. With shorter times of hydrolysis the combined αTp1,2 peptide predominated. The ionophoretic mobility of αTp1, which was neutral at pH 6·4, indicated that residue 6 was aspartic acid in agreement with the incomplete hydrolysis observed.

Quantitative analyses of each tryptic peptide, including the tryptophan peptide from a trypsin digest of unmodified α-globin, and sequence determination by the dansyl–Edman method gave data consistent with the following sequence:

Val-Leu-Ser-Ala-Ala-Asp-Lys-Gly-His-Val-Lys-Ala-Ile-Trp-Gly-Lys-



IV. Discussion

The *N*-terminal fragment of kangaroo α-chain resulting from cleavage at the single tryptophyl bond was isolated in 20% yield. The extent of cleavage was no more than 60% as indicated by the amount of *N*-terminal glycine found in the NBS-treated globin (Table 1). This is comparable to the yields obtained by other workers on proteins (Ramachandran and Witkop 1967). The conditions used were designed to minimize cleavage at tyrosyl and histidyl bonds. In 8M urea the reactive species is *N*-bromourea and the reaction with tryptophan is some 200-fold faster than that with tyrosine (Funatsu, Green, and Witkop 1964).

It has been shown by Shaltiel and Patchornik (1963) that NBS causes oxidation of histidyl residues but that cleavage of histidyl bonds in proteins requires heating for 1 hr at 100°C. The order of rates of oxidation of reactive side chains is tryptophyl followed by tyrosyl followed by histidyl residues. The use of a lower pH of 3·5 might be expected to favour tryptophyl bond cleavage (Sasakawa 1963). At pH values above 5·5 no cleavage of tryptophyl bonds occurs (Green and Witkop 1964).

A time of reaction of 15 min was found to be sufficient for the maximum yield of the glycine residue in kangaroo α-globin. There was no advantage with either human or kangaroo α-chain in using an excess of reagent greater than 10 moles per

mole of protein so far as selectivity was concerned. However, it is apparent that, even using the best conditions based on previous work, there is a considerable amount of peptide bond splitting apart from the tryptophylglycyl bond. In both human and kangaroo α-globin considerable amounts of *N*-terminal alanine were found. This *N*-terminal residue was not present in a control sample of untreated kangaroo globin, pointing to some -X-Ala- bond splitting. The sequence of human α-globin shows -Tyr-Gly-,-Tyr-Phe-, and -Tyr-Arg- bonds which do not account for the -X-Ala- split; the source of the alanine terminal residues is therefore not known. The peak of material most retarded during gel filtration (Fig. 1, peak 4) was shown to contain free arginine in addition to other peptide material. This could represent splitting of the *C*-terminal tyrosylarginine linkage in the globin chain. There is also evidence of heterogeneity from the elution profile in the gel filtration.

There are two stages in the release of peptide fragments in this reaction. The evidence in the literature and in our determination of tryptophan is that the oxidation step is essentially quantitative. The scission of the peptide bond is a slower reaction and according to Patchornik, Lawson, and Witkop (1958*a*) further amounts of NBS are required. It is therefore difficult to increase the yield at the cleavage step by increasing the amount of reagent or time of reaction without increasing the chances of modifying other reactive side chains followed by cleavage of peptide bonds.

The chemistry of the cleavage reaction suggests that a lactone should be formed initially in the cleavage (Green and Witkop 1964) and that equilibration of this with the free acid form should occur. In agreement with this we isolated two zones of different ionophoretic mobility but identical amino acid compositions and *N*-terminal sequences. After elution with ammonia the peptide mixture, on re-running, gave a single ionophoretic band corresponding to the acid form. Treatment with 2M acetic acid at 37°C for 40 hr did not result in re-formation of detectable lactone. The formation of the free acid under alkaline conditions thus appears to be more rapid than is the case with the *C*-terminal homoserine lactone obtained after cyanogen bromide cleavage (Ambler 1965).

The amino acid sequence of the *N*-terminal fragment of the α-chain of kangaroo haemoglobin was found to be

Val-Leu-Ser-ALA-Ala-Asp-Lys-GLY-HIS-Val-Lys-Ala-ILE-Trp.

The residues in capital letters are different from those in human α-globin. This degree of variability is similar to that observed in α-chains of other species (Eck and Dayhoff 1966).

Provided that only a small number of tryptophan residues are present in a polypeptide chain the selective cleavage of tryptophyl bonds by NBS in 8M urea is a feasible method of isolating fragments for amino acid sequence studies. With increasing numbers of tryptophan residues the complexity of the reaction mixture would increase markedly due to incomplete cleavage of each bond. If a fragment can be readily isolated this approach is useful for allocating amino acids, or peptides from enzyme digests, to particular areas of the molecule. In conjunction with the dansyl–Edman procedure it will give sequence information with less dependence on large numbers of quantitative amino acid analyses.

V. Acknowledgments

We are indebted to Professor G. B. Sharman and his colleagues for samples of kangaroo blood and to Miss W. K. Fisher and Mr. R. G. Whittaker for skilled technical assistance. This work was supported in part by the Australian Wool Board and the Australian Research Grants Committee.

VI. References

Air, G. M., and Thompson, E. O. P. (1969).—*Aust. J. biol. Sci.* **22**, 1437.

Ambler, R. P. (1965).—*Biochem. J.* **96**, 32P.

Bennett, J. C. (1967).—In "Methods in Enzymology". (Ed. C. H. W. Hirs.) Vol. XI. p. 330. (Academic Press, Inc: New York.)

Blömback, B., Blömback, M., Edman, P., and Hessel, B. (1966).—*Biochim. biophys. Acta* **115**, 371.

Clegg, J., Naughton, M. A., and Weatherall, D. J. (1965).—*Nature, Lond.* **207**, 945.

Eck, R. V., and Dayhoff, M. O. (1966).—"Atlas of Protein Sequence and Structure." (National Biochemical Research Foundation: Silver Spring.)

Funatsu, M., Green, N. M., and Witkop, B. (1964).—*J. Am. chem. Soc.* **86**, 1846.

Gray, W. R. (1967).—In "Methods in Enzymology". (Ed. C. W. H. Hirs.) Vol. XI. p. 469. (Academic Press, Inc: New York.)

Green, N. M., and Witkop, B. (1964).—*Trans. N. Y. Acad. Sci.* **26**, 659.

Jeppsson, J. O., and Sjöquist, J. (1967).—*Analyt. Biochem.* **18**, 264.

Kronman, M. J., Robbins, F. M., and Andreotti, R. E. (1967).—*Biochim. biophys. Acta* **147**, 462.

Neumann, N. P., Moore, S., and Stein, W. H. (1962).—*Biochemistry* **1**, 68.

Patchornik, A., Lawson, W. B., Gross, E., and Witkop, B. (1960).—*J. Am. chem. Soc.* **82**, 5923.

Patchornik, A., Lawson, W. B., and Witkop, B. (1958*a*).—*J. Am. chem. Soc.* **80**, 4747.

Patchornik, A., Lawson, W. B., and Witkop, B. (1958*b*).—*J. Am. chem. Soc.* **80**, 4748.

Peters, T. (1959).—*C. r. Trav. Lab. Carlsberg* **31**, 227.

Ramachandran, L. K. (1962).—*J. scient. ind. Res.* C**21**, 111.

Ramachandran, L. K., and Witkop, B. (1959).—*J. Am. chem. Soc.* **81**, 4028.

Ramachandran, L. K., and Witkop, B. (1967).—In "Methods in Enzymology". (Ed. C. H. W. Hirs.) Vol. XI. p. 283. (Academic Press, Inc: New York.)

Rao, G. J. S., and Cama, H. R. (1963).—*Biochim. biophys. Acta* **71**, 139.

Sasakawa, S. (1963).—*J. Biochem., Tokyo* **53**, 188.

Shaltiel, S., and Patchornik, A. (1963).—*J. Am. chem. Soc.* **85**, 2799.

Smith, I. (1953).—*Nature, Lond.* **171**, 43.

Spande, T. F., and Witkop, B. (1967).—In "Methods in Enzymology". (Ed. C. H. W. Hirs.) Vol. XI. p. 498. (Academic Press, Inc: New York.)

Thompson, E. O. P., Hosken, R. W., and Air, G. M. (1969).—*Aust. J. biol. Sci.* **22**, 449.

Van Orden, H. O., and Carpenter, F. H. (1964).—*Biochem. biophys. Res. Commun.* **14**, 399.

Witkop, B. (1968).—*Science, N.Y.* **162**, 318.