# The Macroheterogeneity of Type I Tyrosine-rich Proteins of Merino Wool

## J. M. Gillespie and M. J. Frenkel

Division of Protein Chemistry, CSIRO, Parkville, Vic. 3052.

## Abstract

The cystine-poor family of tyrosine-rich proteins of wool as their S-carboxymethyl derivatives has been shown by chromatography on quaternary ammonium ethylcellulose (QAE-cellulose) at pH 10.5 to consist of 10 groups of proteins which span a threefold range of tyrosine content and vary in many other amino acids, particularly S-carboxymethylcysteine. Electrophoresis at pH 8.9 revealed that most fractions contain two or more components and by further chromatography of five of these groups at pH 8.3 on DEAE-cellulose 10 components have been isolated. The amino acid composition suggests that the components within each QAE-cellulose group are very closely related.

Isoelectric focusing of this family of proteins revealed up to 17 distinct components with isoelectric points mainly in the region pH 3–6. Although most proteins are least soluble in their isoelectric region, these proteins are insoluble even at pH 9, which is at least 3 pH units above their pI. As the pH is raised above 9 they become increasingly more soluble, and it has been shown that this is due to the progressive ionization of tyrosine at high pH. It is probable that the insolubility of these proteins is related to interactions involving tyrosine.

# Introduction

Proteins that are very rich in the aromatic amino acids and serine and glycine but quite deficient in lysine, histidine, glutamic acid, isoleucine and methionine have been found in every wool sample examined (Zahn and Biela 1968; Gillespie and Darskus 1971; Gillespie and Frenkel 1974). However, the amount present is quite variable (1-12%), apparently depending on the genetic characteristics of the wool and the dietary conditions under which it is produced (Frenkel *et al.* 1974). There are at least two major families of high-tyrosine proteins which differ markedly in their content of cystine and phenylalanine, but both seem to consist of components with molecular weights less than about 10000. One component  $(I_{0.62})^*$  of the type I family [poor in  $\frac{1}{2}$ cystine (<7 residues %), rich in phenylalanine (about 10 residues %)] has already been purified and the amino acid sequence determined (Dopheide 1973; Frenkel *et al.* 1973).

The heterogeneity of these proteins has previously been studied using the S-methyl derivative (Zahn and Biela 1968), the S-carboxymethyl (SCM) form (Gillespie 1972; Frenkel *et al.* 1973) and the S-3-sulphoalanyl form (Brunner and Brunner 1973). The present study represents the first part of an investigation into the heterogeneity of type I high-tyrosine proteins (HT-I) as their SCM derivatives, a feature of the work being the isolation of 10 of the major proteins in substantially purified form. Much of the ambiguity of earlier work has been eliminated by studying purified proteins. The aim

\* A type I high-tyrosine component with relative mobility 0.62 as defined by Frenkel et al. (1973).

of the work is to determine the nature and extent of heterogeneity and the precise relationships that exist between members of this family of proteins. This paper examines the solubility properties of these proteins and their macroheterogeneity.

# **Materials and Methods**

#### Preparation of Type I High-tyrosine Proteins

Wool used in this work was obtained from two Merino sheep (fine non-Peppin strain) which were housed indoors, fed a constant diet, and the wool clipped at monthly intervals. HT-I proteins were isolated from the individual fleeces in the SCM form by the procedure of Gillespie (1972). Proteins isolated from the different sheep were not mixed, although these were found to be identical in the number and proportion of components they contained.

## Chromatography on Quaternary Ammonium Ethylcellulose (QAE-cellulose)

Chromatography of HT-I proteins on a QAE-cellulose (Schleicher & Schuell 2282) column (50 by 3 cm) was carried out with a pumping rate of  $1 \cdot 6$  ml/min; fractions of about 10 ml were collected. Proteins were eluted with an initial linear gradient of NaCl in the starting buffer ( $0 \cdot 05M \beta$ -alanine– $0 \cdot 05M$  NaOH, pH 10  $\cdot 5$ ) to a limiting concentration of  $0 \cdot 5M$  with an eluant volume of 2400 ml, followed by a linear gradient ranging from  $0 \cdot 5$  to  $1 \cdot 0M$  NaCl and an eluant volume of 1200 ml. The special precautions needed in the preparation of the column have already been described (Gillespie 1972).

#### Chromatography on DEAE-cellulose

The major fractions obtained from QAE-cellulose were further fractionated on DEAE-cellulose either in the fibrous form (Eastman 7392) or as the microgranular Whatman DE32. Columns (25 by 2 cm) were packed with the required resin and fractions of approximately 5 ml were collected at a pumping rate of 1.6 ml/min with the fibrous DEAE-cellulose and 0.32 ml/min with the microgranular form. The starting buffer for both types of chromatography was 0.05M glycylglycine–0.025M NaOH–6M urea, pH 8.3, and the salt gradient used in each experiment is shown on the appropriate chromatographic profile (Figs 6a-6d).

### Electrophoresis in Starch Gels and on Cellulose Acetate

Details of the electrophoresis in starch gels at pH 8.6 (Gillespie 1972), and on cellulose acetate at pH 8.9 (Blagrove *et al.* 1974) have been previously given.

#### Isoelectric Focusing

Isoelectric focusing in acrylamide discs was performed in the pH range 3–10 (Karlsen *et al.* 1973). The gels were then stained in a Coomassie brilliant blue–sodium formate–formic acid mixture as described previously (Barrett 1973).

#### Moving-boundary Electrophoresis

HT-I protein dissolved in  $\beta$ -alanine-KOH buffer (I = 0.1) at pH 10.9 was dialysed against this buffer for several days. Moving-boundary electrophoresis was carried out in a Tiselius apparatus (LKB, Stockholm) at a protein concentration of 1.6% with a voltage gradient of 7 V/cm for 3 h.

#### Determination of the Solubility Properties of HT-I Protein

HT-I protein or its iodinated derivative was dissolved in 0.1M ammonia and diluted with water to 0.4%. Aliquots (3 ml) were dialysed with rocking for 2 days at 2°C against 50-ml lots of a mixed buffer system in the pH range 6–12. The buffers were made by titrating 40 ml of a solution containing 0.05M each of piperidine, ethanolamine, tris and imidazole to the required pH with 1M HCl and then adding sufficient 1M NaCl and water to give a final volume of 50 ml and an ionic strength of 0.2. After dialysis the pH values of the buffers were measured, the contents of the dialysis bags were centrifuged at 40 000 g for 1 h, and the protein concentration in the supernatant was measured

spectrophotometrically at 278 nm. For clarification before spectrophotometry, the supernatants were first diluted with nine volumes of acetic acid, or in the case of the iodinated protein with nine volumes of 6M urea–0.3 I phosphate, pH 7. It was found that even after prolonged centrifugation the supernatants, especially from the unmodified protein, were somewhat turbid and this turbidity slowly but progressively increased on standing, suggesting that agreggation leading to insolubilization was very slow. A consequence of this is that the high-tyrosine proteins appear somewhat more soluble than they actually are.

#### Amino Acid Analysis

The samples of protein were hydrolysed *in vacuo* for 22 h at 108°C with 6M HCl containing 2 mm phenol. The hydrolysates were freeze-dried and the content of amino acids determined with a Beckman Spinco 120C amino acid analyser.

#### Relation between pH and Ionization of Tyrosyl Residues

A series of 0.001% solutions of HT-I protein prepared in a buffer system containing 0.025M each of tris,  $\beta$ -alanine and K<sub>2</sub>HPO<sub>4</sub>, and 0.2M KCl, were titrated to the required pH in the range 9–13 with KOH or HCl. The solutions were shaken for 30 min and the final pH was measured, then the extinction of each solution was measured at 245 nm in a 4-cm pathlength cell against the pH 9 solution as blank. The change in absorbance is a measure of the extent of ionization of tyrosyl residues.

#### Iodination of HT-I Protein

HT-I protein (200 mg) was dissolved in 20 ml of 6M urea–0·10M sodium tetraborate that had been titrated to pH 9·9 with NaOH. This was cooled in ice and cold KI<sub>3</sub> solution (0·05M I<sub>2</sub> in 0·24M KI) was slowly run in until the stoichiometric amount to completely diiodinate the tyrosine had been added. The yellow solution was extensively dialysed, first against 0·01M borate and then against several changes of deionized water, and then freeze-dried (Means and Feeney 1971).



Fig. 1. Effect of pH on the solubility in aqueous buffers of HT-I SCM-protein in the unmodified form ( $\bullet$ ) and after iodination ( $\odot$ ). A third curve shows the relation between pH and ionization of tyrosine hydroxyl groups in this family of proteins ( $\Box$ ).

# Results

# Solubility Properties of HT-I Protein

The effect of pH on the solubility of HT-I protein was studied over the pH range 6–12 and the results are plotted in Fig. 1. It can be seen that, as the pH was lowered from the region of 11, there was a very steep fall in the solubility of the proteins so that at pH 10·1 about 50% of the protein was precipitated and at pH 9·6 90% had precipitated. At pH 9 the proteins were virtually insoluble with a solubility of less than 0·01%. The solubility is adequate at pH 10·4 (0·35%) and at values

above this for studies to be made of the properties of these proteins in dilute aqueous salt solutions.

The supernatants corresponding to points between pH 9 and 11.5 were dialysed, freeze-dried and examined electrophoretically in starch gels. It can be seen (Fig. 2) that as the pH is lowered there is a progressive and complete removal of components showing lowest mobility and presumably lowest net charge at the pH of electrophoresis (8.6).





The solubility curve of the iodinated protein (Fig. 1) is shifted about 1 pH unit to the acid side, although the shape of the curve remains essentially unchanged in showing no discontinuity.

# Tyrosyl Titration Curve

The relation between the ionization of tyrosyl hydroxyl groups and pH for HT-I protein is plotted on Fig. 1. It can be seen that there is little ionization below pH 10 but a steep increase occurs as the pH is increased above this value. The curve relating solubility of the unmodified protein to pH parallels the ionization curve but is shifted a little over 0.5 of a pH unit in the acid direction.

# Evidence for Charge Heterogeneity

The HT-I protein was examined electrophoretically at pH 11 by the movingboundary procedure. The resulting pattern (Fig. 3*a*) shows the presence of two major components and a third minor component evident as a shoulder on the slower moving peak. Over the pH range 8.5-11, cellulose acetate electrophoresis gave slightly better resolution than moving-boundary electrophoresis, two major, three minor and several trace bands being visible (Fig. 3b). When electrophoresis on cellulose acetate was carried out at pH 11 in phosphate, carbonate or piperidine buffers rather than the glycine buffer (shown in Fig. 3b) there was no resolution into separate bands.



Starch-gel electrophoresis at pH 8.6 gave somewhat better resolution (Fig. 3c) than either of the previous techniques. The increase in resolution relative to cellulose acetate did not apply when subfractions were studied, and here cellulose acetate was the preferred method.

Isoelectric focusing in the pH range 3–10 gave the best electrophoretic resolution. Up to 17 components were visible on the gel, corresponding to a range of isoelectric points from 3.5 up to about 7 (Fig. 3*d*), with the majority of bands having pI values less than 6. This is in agreement with the approximate value calculated from the amino acid composition, assuming that some of the carboxyl groups are amidated.

# Evidence for the Presence of Proteins of Differing Tyrosine Content

When HT-I protein was chromatographed on QAE-cellulose at pH 10.5, two major and eight minor fractions were obtained (Fig. 4). This degree of heterogeneity is greater than was indicated by electrophoretic studies in this pH region. Ten fractions

were pooled as indicated on the elution profile (Fig. 4) and dialysed against deionized water, then ammonia was added to solubilize the proteins and the solutions were freeze-dried. Complete amino acid analyses of each fraction were performed (Gillespie 1972), although only the level of tyrosine is shown on the profile. It can be seen that the HT-I protein consists of a minimum of 10 components, with tyrosine levels ranging from 7.8 to 20.2 residues %.



Fig. 4. Chromatography of HT-I proteins on QAE-cellulose at pH 10.5 in  $\beta$ -alanine–NaOH buffer of ionic strength 0.05. The linear gradients of NaCl are indicated on the profile. Pooled fractions (A–J) are designated by the open bars.

# Evidence for Heterogeneity within QAE-cellulose Fractions

Each QAE-cellulose fraction (fractions G and H were pooled for electrophoresis and subsequent fractionation) was examined electrophoretically on cellulose acetate at pH 8.9 (Fig. 5). It can be seen that the resolution of the chromatographic separation was extremely good, with very little overlap of components between successive fractions. Each fraction, with the exception of F, contains a number of components, indicating that there is a substantial degree of heterogeneity within each HT-I protein fraction. This electrophoretic examination indicates the presence of at least 35 components, many of which are quite minor.

Four of the QAE-cellulose fractions (C, D, G+H and I) were chromatographed on DEAE-cellulose at pH 8.3 in 6M urea and the elution profiles are shown in Figs 6(a)-6(d). Each profile revealed many peaks and when fractions that were pooled as indicated were examined on cellulose acetate it could be seen (Figs 7*a*-7*d*) that each fraction generally contained only one major component with traces of minor constituents present. Although not completely pure, some of these subfractions have been hydrolysed and their amino acid composition determined (Table 1). This table also shows the composition of fraction F, composed largely of a single component,  $I_{0:62}$ , which has already been studied (Dopheide 1973; Frenkel *et al.* 1973).

From Table 1 it can be seen that there is a striking similarity between subfractions within a QAE-cellulose group; thus differences between subfractions C2, C3 and C4 are less than 1 residue % for most amino acids. It is not surprising that particular

compositional characters noted for each group are retained by the subfractions, e.g. the subfractions of group C have a lower tyrosine content than those of group G + H in accord with the trend shown by the parent QAE-cellulose fractions.

Besides heterogeneity of tyrosine there is also a large variation in most of the amino acids, the ranges being indicated in Table 1. It can be seen that methionine is absent in all of the fractions and that isoleucine is generally not present. The



Fig. 5. Electrophoresis on cellulose acetate in 8M urea at pH 8.9 of the fractions (A–J) obtained from QAE-cellulose chromatography (see Fig. 4). U is the pattern of unfractionated HT-I proteins; G/H =fractions G+H.

 Table 1. Amino acid compositions (residues %) of a type I high-tyrosine protein preparation and 10 purified fractions prepared from it

In the identification of fractions, the prefixed letters (C, D, F, G+H, I) refer to the QAE-cellulose groups (elution curve in Fig. 4) and the numerals refer to components within a group separated on DEAE-cellulose (elution curves in Fig. 6). The range over which each amino acid varies is also indicated

Amino acid	Whole HT-I	Fractions										Range
		C2	C3	C4	D2	D3	D4	F	G + H2	G + H3	I4	
Lys	0.44	0.18	0.15	0.40	0.09	0.06	0.00	0.00	0.62	0.65	0.26	0.0-0.2
His	1.13	1.35	1.66	1.83	1.40	1 · 44	1.35	0.00	1.91	1.71	0.19	0.0-1.9
Arg	5.35	4.90	5.38	5.73	7.43	6.86	6.60	3.40	3.86	4.32	4.64	3 • 4 - 7 • 4
SCMCys	5.98	5.56	6.46	6.53	5.43	6.11	6.56	6.42	1.59	2.44	3 · 34	1.6-6.6
Asp	3.31	4.31	4 · 57	4.53	2.88	3.16	3.61	1.85	1.91	$2 \cdot 04$	4.76	$1 \cdot 9 - 4 \cdot 8$
Thr	3.33	4.81	4.74	4.18	1.47	1.96	2.70	3.41	0.98	1.06	3.28	1.0-4.8
Ser	11.80	12.60	13.70	13.60	12.90	13.40	13.40	12.60	12.80	11.70	8.95	9.0-13.7
Glu	0.57	0.48	0.44	0.86	0.15	0.00	0.32	0.24	1.24	0.82	$1 \cdot 80$	0.0-1.8
Pro	5.28	6.55	6.44	5.56	1.48	2.42	2.90	6.69	2.37	3.26	6.56	1 · 5 – 6 · 7
Gly	27.60	30.80	$25 \cdot 50$	26.80	34.00	32.50	28.90	$26 \cdot 50$	35.90	34.20	$25 \cdot 80$	25.5-35.9
Ala	1.48	1.05	1.04	1 · 14	0.16	0.32	0.32	3.07	0.56	0.97	0.57	$0 \cdot 2 - 3 \cdot 1$
Val	2.09	1.64	1.57	1.66	0.16	0.32	0.64	4.70	0.36	0.48	3.22	0.2-4.7
Met	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Ile	0.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.42	0.0-1.4
Leu	5.46	5.65	6.25	6.53	4.57	4.57	5.47	3.51	7.67	7.66	4.31	3 · 5 - 7 · 7
Tyr	15.00	10.10	11.10	10.40	13.80	13.60	14.30	18.00	19.80	19.70	20.20	$10 \cdot 1 - 20 \cdot 2$
Phe	10.30	10.00	11.00	10.70	13.90	13.40	12.80	9.63	8.29	8.88	9.46	8.3-13.9







# Discussion

Type I high-tyrosine SCM-proteins are very insoluble at pH values below about 9 (Fig. 1). Above this pH these proteins show globulin-like properties, being rather insoluble in low-ionic-strength solutions such as 0.1M ammonia, but becoming much more soluble and giving clear solutions on the addition of small amounts of NaCl. Proteins normally show least solubility in their isoelectric region, but the HT-I proteins are insoluble even at pH values well outside their isoelectric range (pH 3-6). It is likely that this insolubility is due to their high content of aromatic amino acids, for similar solubility properties have been recorded for poly-L-tyrosine (Katchalski and Sela 1953). As these proteins contain virtually no ε-amino groups, the only side chain groups capable of changing their state of ionization in the pH region from 9 to 11 will be the tyrosyl hydroxyl groups, and it is reasonable to conclude that the increased solubility as the pH is raised above 9 is due to the progressively increased ionization of these groups. If this were correct it would be expected that the curves relating solubility and ionization of tyrosine to pH would be coincident but in fact, although they have the same shape, the solubility curve is shifted to a lower pH range. It is probable that the solubility curve is in error because of the extremely slow approach to equilibrium in the precipitation of these proteins. The close correspondence in shape of the two curves leaves little doubt that the ionization of at least some of the tyrosyl residues is responsible for the solubilization of these proteins. This conclusion is supported by the shift in position of the solubility curve of the iodinated protein, which is displaced in the direction of increased solubility at lower pH values, as would be expected from the lower pK (about 8) of diiodotyrosine in an acidic protein (Means and Feeney 1971). The interactions responsible for the insolubility of high-tyrosine proteins can also be broken by 6 or 8M urea and under these conditions the proteins are quite soluble except in their immediate isoelectric range. Whiting and Tanford (1962) found that urea greatly increased the solubility of tyrosine and phenylalanine and attributed this to a reduction in hydrophobic forces. It is likely that the solubilizing effect of urea on the high-tyrosine proteins is due to similar causes.

A definite relation was observed in chromatography on QAE-cellulose at pH 10.5, between the elution position of a component and its tyrosine content. It has been suggested that this is an ion-exchange process in which the negative charge on the

**Fig. 6.** Chromatography of some QAE-cellulose fractions (see Fig. 4) on DEAE-cellulose at pH 8.3 in glycylglycine–NaOH–urea buffer. Linear gradient of NaCl is indicated on each profile. Pooled subfractions are designated by the open bars. (*a*) Fraction C on fibrous DEAE-cellulose (Eastman 7392). (*b*) Fraction D on fibrous DEAE-cellulose. (*c*) Fractions G+H on microgranular DEAE-cellulose (Whatman DE32). (*d*) Fraction I on microgranular DEAE-cellulose.

Fig. 7. Electrophoresis of the subfractions obtained from DEAE-cellulose chromatography of QAE-cellulose fractions (see Fig. 6). Electrophoresis was on cellulose acetate in 8M urea at pH 8.9; the direction (arrow) was the same for all strips. The numbers on (a), (b), (c) and (d) correspond to the subfractions on the respective parts of Fig. 6.

protein and hence its elution position is determined by its content of ionized tyrosyl residues (Gillespie 1972). However, the finding in the present electrophoretic studies that the various components do not differ greatly in charge at pH 10.5-11 casts doubt on this conclusion and suggests a more complex situation in which aromatic interactions between tyrosine and the ion-exchange resin are the dominating influence. Similar interactions appear to be responsible for the extreme retardation of these proteins on Sephadex and Biogel-P (Frenkel *et al.* 1973). Brunner and Brunner (1973) attributed large losses on anion exchangers to aromatic interactions between the protein and the supporting resin. Whatever the reason, however, there is no doubt that this initial separation at pH 10.5 gives a degree of resolution not achieved by any other fractionation procedure, separating 10 groups of components which differ not only in tyrosine content but also in many other amino acid residues, particularly SCM-cysteine. One of these groups, G+H, contains components with far lower SCM-cysteine contents than any yet recorded for proteins from a hard keratin.

Within the QAE-cellulose fractions there is considerable heterogeneity of a relatively minor kind and, although it is difficult to make an accurate estimate, the analytical procedures used in this study would suggest the presence of at least 30–40 components, of which only six could be considered of major significance. The causes of the heterogeneity within groups is being examined in work now in progress. This present work indicates a degree of macroheterogeneity less than was previously thought, making these proteins comparable in this respect with the low-sulphur proteins, which now appear to consist of seven or eight components, and much less complicated than the high-sulphur proteins (Joubert *et al.* 1968; W. G. Crewther, personal communication).

Previous workers using S-methyl (Zahn and Biela 1968) and S-3-sulphoalanyl (Brunner and Brunner 1973) derivatives also found heterogeneity in the high-tyrosine proteins. It is difficult, however, to make comparisons between their work and ours for three reasons; they used different derivatives, their proteins were mixtures of type I, type II and probably membrane proteins (Bradbury *et al.* 1965) and no evidence was provided of the state of electrophoretic homogeneity of their subfractions. Certainly, the correlations found previously (Brunner and Brunner 1973), i.e. a positive relationship between cystine and tyrosine and a negative relationship between tyrosine and phenylalanine, do not apply to the HT-I components discussed in this work. Studies in progress on the type II components have so far also failed to reveal such correlations. These relationships may have arisen previously through the analysis of mixtures of different component species. Nevertheless, previous work does give an appreciation of the complexity of these proteins and the difficulties experienced in their fractionation.

# Acknowledgments

Our thanks are due to Dr C. M. Roxburgh for the amino acid analyses and to the School of Agriculture, University of Melbourne, for housing the sheep at their Mt Derrimut Field Station.

# References

Barrett, A. J. (1973). Biochem. J. 131, 809. Blagrove, R. J., Frenkel, M. J., and Gillespie, J. M. (1974). Comp. Biochem. Physiol. (In press.) Bradbury, J. H., Chapman, G. V., and King, N. L. P. (1965). Proc. 3rd Int. Wool Text. Res. Conf., Paris. Vol. 1, p. 359.

Brunner, H., and Brunner, A. (1973). Eur. J. Biochem. 32, 350.

Dopheide, T. A. A. (1973). Eur. J. Biochem. 34, 120.

Frenkel, M. J., Gillespie, J. M., and Woods, E. F. (1973). Eur. J. Biochem. 34, 112.

Frenkel, M. J., Gillespie, J. M., and Reis, P. J. (1974). Aust. J. Biol. Sci. 27, 31.

Gillespie, J. M. (1972). J. Chromatogr. 72, 319.

Gillespie, J. M., and Darskus, R. L. (1971). Aust. J. Biol. Sci. 24, 1189.

Gillespie, J. M., and Frenkel, M. J. (1974). Comp. Biochem. Physiol. 47B, 339.

Joubert, F. J., Jager, P. J. de, and Swart, L. S. (1968). In 'Symposium on Fibrous Proteins'. (Ed. W. G. Crewther.) p. 343. (Butterworths Australia: Sydney.)

Karlsen, C., Davies, H., Öhman, J., and Andersson, U. (1973). 'LKB 2117 Multiphor. I. Analytical Thin Layer Gel Electrofocusing in Polyacrylamide Gel.' (LKB-Produkter Stockholm, Application Note 75.)

Katchalski, E., and Sela, M. (1953). J. Am. Chem. Soc. 75, 5284.

Means, G. E., and Feeney, R. E. (1971). 'Chemical Modification of Proteins.' pp. 225-6. (Holden Day Inc.: San Francisco.)

Whiting, P. L., and Tanford, C. (1962). J. Biol. Chem. 237, 1735.

Zahn, H., and Biela, M. (1968). Eur. J. Biochem. 5, 567.

Manuscript received 12 June 1974

