PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

The <latex>α</latex>-Chain-Termination Mutants and Their Relation to the <latex>α</latex>-Thalassaemias

D. J. Weatherall and J. B. Clegg

Phil. Trans. R. Soc. Lond. B 1975 **271**, 411-455
doi: 10.1098/rstb.1975.0061

THE α-CHAIN-TERMINATION MUTANTS AND THEIR RELATION TO THE α-THALASSAEMIAS

By D. J. WEATHERALL† and J. B. CLEGG†
Department of Haematology, University of Liverpool

(*Communicated by P. M. Sheppard, F.R.S. – Received 15 October 1974*)

[Plates 22–24]

CONTENTS

† Present address: Nuffield Department of Clinical Medicine, University of Oxford, Radcliffe Infirmary, Oxford OX2 6HE.

The structure, synthesis, genetic transmission, clinical associations and distribution of the elongated α-chain haemoglobin variants has been described. The data indicate that the most likely molecular basis for these common abnormal haemoglobins is a single base substitution in the α-chain termination codon. Because these variants are produced inefficiently they give rise to the clinical picture of α-thalassaemia. When these findings are taken together with recent work regarding the molecular basis for other forms of α-thalassaemia it is possible to build up a fairly complete picture of the molecular pathology of the α-thalassaemias.

1. Introduction

The human haemoglobins all have a similar basic structure consisting of four globin chains each associated with a haem molecule. The principal haemoglobin (Hb) in adults is Hb A which has 2 α chains and 2 β chains ($\alpha_2\beta_2$), while in foetal life the major haemoglobin is Hb F in which the β chains are replaced by γ chains ($\alpha_2\gamma_2$). Until recently it was thought that the inherited disorders of haemoglobin production could be classified into two fairly distinct groups. One consisted of haemoglobin variants with structural alterations in one of the globin chains, leading to reduced stability or abnormal function of the molecule. The other group was the thalassaemias which are characterized by a reduced rate of synthesis of one or more of the globin chains resulting in imbalanced globin-chain production, intracellular precipitation of the partner chain which is produced in excess, and hence in disordered erthropoiesis. However it has become apparent in the last few years that this simple division of the haemoglobinopathies is not as clear cut as was thought previously; we know now that there are several structural haemoglobin variants which are associated with a markedly reduced rate of globin-chain synthesis and hence with the clinical picture of thalassaemia.

The thalassaemias are the commonest inherited disorders of haemoglobin synthesis and produce an extensive public-health problem in many parts of the world. They are classified broadly into the α- and β-thalassaemias depending on whether α- or β-chain synthesis is defective (Ingram & Stretton 1959; Weatherall & Clegg 1972*a*). The β-thalassaemias are a heterogeneous group of disorders which, in the more severe forms, produce a clinical picture which was first described by Thomas Cooley in 1925 and which is known in the haematological literature as Cooley's anaemia (Cooley & Lee, 1925) or thalassaemia major. Although inefficiently-synthesized haemoglobin variants like the Hb Lepores or K Woolwich (Gerald & Diamond 1958; Baglioni 1962; Lang, Lehmann & King-Lewis 1974) can produce the clinical picture of β-thalassaemia, these are relatively uncommon and in most cases no evidence for a structural alteration in the β chain has been demonstrated. In every form of β-thalassaemia the reduced rate of β-chain production results in an excess of α chains which are unstable and precipitate in the red cell precursors. It is the intracellular inclusions which are produced by this process which are responsible for the disordered erythropoiesis common to all these conditions (Fessas 1963; Nathan & Gunn 1968).

The α-thalassaemias also show considerable heterogeneity. The reduced rate of α-chain production leads to an excess of γ chains in foetal life and of β chains in adult life, and these aggregate to produce the abnormal tetramer haemoglobins, Bart's (γ_4) and H (β_4) respectively (Rigas, Koler & Osgood 1955; Gouttas, Fessas, Tsevrenis & Xefteri 1955; Hunt & Lehmann 1959). These are viable molecules and, unlike the excess α chains of β-thalassaemia, do not precipitate extensively in the bone marrow; hence in the α-thalassaemias erythroid-precursor proliferation

is not as abnormal as in the β-thalassaemias. However Hbs Bart's and H are relatively unstable and precipitate in red cells as they age producing rigid intracellular inclusions. Hence α-thalassaemia is associated with a variable degree of haemolytic anaemia.

Two major clinical types of α-thalassaemia have been recognized, the Hb Bart's hydrops syndrome and Hb H disease. The former condition, which is a common cause of stillbirth throughout Southeast Asia, is characterized by death *in utero* or immediately after delivery, associated with the clinical picture of hydrops fetalis. Affected infants have a thalassaemic blood picture and a haemoglobin pattern that consists almost entirely of Hb Bart's with smaller amounts of Hb H and Hb Portland ($\zeta_2\gamma_2$) (Lie-Injo & Jo 1960; Weatherall, Clegg & Wong 1970; Capp, Rigas & Jones 1970); the latter is a normally-occurring minor haemoglobin of foetal life. Biosynthetic studies have shown that these babies make no σ chains whatever (Weatherall *et al.* 1970). Haemoglobin H disease is consistent with survival into adult life and is characterized by a variable degree of anaemia and splenomegaly associated with the presence of 5–30 % Hb H in the peripheral blood (Gouttas *et al.* 1955; Rigas *et al.* 1955). In this condition there is a variable degree of globin-chain imbalance with an excess of γ- and β- over α-chain production in the range of 2–3/1 (Weatherall, Clegg & Naughton 1965; Clegg & Weatherall 1967).

Genetic data, collected principally in Thailand, suggests that the Hb Bart's hydrops syndrome and Hb H disease result from the interaction of two different α-thalassaemia genes designated α-thalassaemia 1 (α-thal. 1) and α-thalassaemia 2 (α-thal. 2) (Wasi *et al.* 1969; Wasi 1973). These studies suggest that the Hb Bart's hydrops syndrome results from the homozygous state for the α-thal. 1 gene which must therefore result in a total absence of α-chain synthesis. The α-thal. 2 gene, on the other hand, causes only a partial reduction in α-chain production. Haemoglobin-H disease results from the doubly heterozygous state for both α-thal. 1 and α-thal. 2. The heterozygous states for either the α-thal. 1 or α-thal. 2 genes alone produce only mild haematological changes in adult life but are characterized by the presence of approximately 5 and 2 % of Hb Bart's respectively in the neonatal period.

In 1968 Sofroniadou and her colleagues noted trace amounts of an electrophoretically slowly-migrating haemoglobin variant in red-cell lysates obtained from a Greek patient with Hb H disease. These workers partially characterized this haemoglobin as an α-chain variant and called it Hb Athens. In 1969 Wasi and his colleagues, reviewing a large number of cases of Hb H disease seen in Thailand, noted independently that some patients had trace amounts of a haemoglobin with similar properties to Hb Athens and called it Hb Thai; this variant was also shown to have an abnormal α chain by Dr E. R. Huehns of University College Hospital, London. Independent studies of patients with Hb H disease in Malaysia by Lie-Injo, Lopez & Lopez (1971) also found trace amounts of slowly-migrating haemoglobins in patients with Hb H disease and these were called the 'abnormal X components of Hb H disease' and a similar observation was made by Todd (1971) in Hong Kong. The significance of these four sets of independent observations did not become apparent until after 1971. In that year Dr P. F. Milner of Kingston, Jamaica, sent us blood samples from individuals with Hb H disease from a Chinese family who lived in Constant Spring, a suburb of Kingston. Each of the affected family members carried trace amounts of a haemoglobin variant with similar electrophoretic properties to those of Hbs Athens, Thai and X.

The clinical, genetic, chemical and biosynthetic characterization of this variant together with similar studies on Hbs Thai, Athens, X and related haemoglobins form the subject of

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

the present paper. These studies underline the heterogeneity of the slow-moving haemoglobin components of Hb H disease, and when taken together with recent evidence regarding the molecular basis of the Hb Bart's hydrops syndrome now provide a relatively clear picture of the molecular pathology of the different forms of α-thalassaemia.

Preliminary data and brief summaries of some aspects of these studies have been presented previously (Milner, Clegg & Weatherall 1971; Clegg, Weatherall & Milner 1971; Weatherall & Clegg 1972*b*; Fessas *et al.* 1972; Clegg & Weatherall 1974*b*; Clegg *et al.* 1974; Lie-Injo, Ganesan, Clegg & Weatherall 1974).

2. Individuals studied

Venous blood samples were collected into standard acid–citrate–dextrose or heparin anti-coagulant (Dacie & Lewis 1970). Bone-marrow samples were obtained by iliac-crest puncture. Blood was transported to Liverpool by air freight either as whole blood on ice, or as washed, frozen red cells on dry ice. In experiments involving [^{14}C] or [^{3}H]leucine incorporation the cells were washed three times, frozen and transported on dry ice.

Samples containing slowly-migrating haemoglobin variants were obtained from a variety of population groups. The initial studies were performed on a large kindred with Hb H disease from Constant Spring, Kingston, Jamaica. Bone marrow and further blood samples were obtained from two members of this family on a subsequent visit to Jamaica. Peripheral-blood samples were obtained from several members of the original family with Hb Athens, and from patients with Hb H disease and slowly-migrating haemoglobins from Hong Kong, Thailand and Malaysia. Blood samples were also obtained from members of a Malaysian family in which there was a child apparently homozygous for the abnormal minor (X) components of Hb H disease, and from another Greek family in which similar haemoglobin variants were found in several members who had evidence of a mild thalassaemia-like disorder.

The sources of all these samples are given in detail in the acknowledgements at the end of this paper.

3. Methods

(*a*) *Haematological investigations*

Haematological studies followed standard techniques (Dacie & Lewis 1970). Red-cell counts were carried out either manually or with a Coulter Model S electronic cell counter. Reticulocyte counts and Hb H inclusion preparations were made by incubating red cells with brilliant cresyl blue for 30 min. Previously described methods were used for the estimation of alkali-resistant or unstable haemoglobins (Dacie & Lewis 1970; Weatherall & Clegg 1972*a*)

(*b*) *Haemoglobin electrophoresis*

Haemoglobin solutions were prepared for electrophoresis as described by Weatherall & Clegg (1972*a*). All samples were analysed by starch-gel electrophoresis with a tris–EDTA–borate buffer system (pH 8.5) or a phosphate buffer system (pH 7.0) (Weatherall & Clegg 1972*a*). The relative proportions of the various haemoglobin components were estimated by quantitative cellulose-acetate electrophoresis (Weatherall *et al.* 1971).

(*c*) *Isolation of slowly-migrating haemoglobins*

Although it was possible to isolate the abnormal globin chains of the variants by urea chromatography of whole-cell globin (see later section) material obtained in this way was found to be contaminated with normal globin chains and for this reason it became necessary to carry out a preliminary purification step. After trying several systems it was found that chromatography on Amberlite IRC-50 gave the most satisfactory results.

Amberlite IRC-50 was suspended in 'developer 2' (Allen, Schroeder & Balog 1958), the fines were removed by repeated washings and the pH of the suspension was adjusted to 7.18. Columns (3 cm × 30 cm) were washed for several days at 4 °C with developer 2 until the effluent and influent pHs were identical. Red-cell lysates containing 1–2 g haemoglobin were dialysed for 24 h at 4 °C against developer 2 and then applied to the Amberlite columns at the same temperature. Under these conditions Hbs H and Bart's are not retained; after they were eluted completely the columns were warmed to 24 °C at which temperature Hbs A and A_2 were eluted. The slowly-migrating components, which remained on the column either in the form of a tight band or more diffuse zone, were eluted by adjusting the Na^+-ion concentration of the developer 2 to approximately 0.5 M by the addition of NaCl. The relative proportions of each of the fractions were calculated by measuring the volumes and absorbance at 540 nm. Each fraction was concentrated *in vacuo* and then dialysed against developer 2 or, in some cases, against water; their purity was then assessed by starch-gel electrophoresis with the tris–EDTA–borate buffer, pH 8.5. Preliminary experiments showed that an almost 100 % recovery of haemoglobin was obtained by using this chromatographic procedure.

(*d*) *Globin-chain separation*

In some experiments globin was prepared by the acid–acetone technique from a whole-cell lysate without any further purification steps. Such material will be referred to hereafter as 'whole-cell globin'; in addition to globin this material contained red-cell membrane proteins and other non-haemoglobin material. Globin was prepared by a similar technique from various purified fractions obtained by Amberlite chromatography. In this case the haemoglobin fractions were dialysed against water before acid–acetone precipitation. Globin-chain separation was carried out on CM-cellulose columns with an 8 M urea–mercaptoethanol–phosphate buffer system, pH 6.8, as described by Clegg, Naughton & Weatherall (1966). After chromatography the individual fractions were pooled, dialysed against 0.5 % formic acid, and then freeze dried. If necessary, the globin chains were converted to *S*-β-aminoethylcysteine (AE) derivatives immediately after chain separation (Clegg *et al.* 1966).

(*e*) *Cyanogen bromide cleavage*

AEα chains were dissolved in 70 % HCOOH to a concentration of 10 mg/ml and treated with a fivefold mass excess of CNBr at 24 °C for 24 h after which time solvent and excess reagent were removed by freeze drying. The freeze-dried material was dissolved in a small volume of 20 % HCOOH and fractionated on a 2 cm × 70 cm column of Sephadex G-100 in 5 % HCOOH.

(*f*) *Enzymic digests of* α *chains and* α*-chain peptides*

Tryptic and chymotryptic digests of α chains were made in 1 % NH_4HCO_3, pH 8, at an enzyme/substrate ratio of 1/100 for 3 h at 37 °C. Digestion was stopped by freeze drying.

The dry digests were dissolved in a small volume of water and freeze dried again to ensure complete removal of NH_4HCO_3.

Digestion of peptides with chymotrypsin, thermolysin, and carboxypeptidase A was carried out as described by Ambler (1963) and Ambler & Medway (1968).

(*g*) *Fractionation of enzyme digests*

The initial fractionation was made by high-voltage paper electrophoresis at pH 6.5 or 4.7. This was followed by descending chromatography in *n*-butanol : acetic acid : water : pyridine; 15:3:12:10 (by vol.) (BAWP) and if necessary by further electrophoretic runs at pH 3.5 or pH 2.1. Details of techniques and buffers are given by Milstein (1966). Peptides were eluted from papers with 0.1 M NH_4OH or 0.1 M CH_3COOH. Peptide maps for diagnostic and analytical purposes were made according to Clegg *et al.* (1966). Papers were stained with 0.02 % ninhydrin in acetone, and when required for amino acid analysis peptides were eluted from the stained papers with 6 M HCl containing 2 mg/ml phenol to minimize destruction of tyrosine (Sanger & Thompson 1963).

Specific stains for individual amino acids in peptides were used on ninhydrin-stained peptide maps as described previously (Clegg *et al.* 1966).

(*h*) *N-terminal analysis and N-terminal sequence determination of peptides*

These were performed by the 'dansyl' and 'dansyl'-Edman methods (Gray 1967). Dansyl(DNS)-amino acid derivatives were identified by thin-layer chromatography on polyamide sheets as described by Woods & Wang (1967). Whenever there was sufficient peptide material available, N-terminal sequence determinations by the 'dansyl'-Edman method were corroborated by amino acid analysis of an aliquot of the peptide after each cycle of degradation.

(*i*) *Haemoglobin synthesis*

The relative rates of globin-chain synthesis in peripheral-blood samples were estimated by previously reported methods (Weatherall *et al.* 1965; Weatherall, Clegg, Na-Nakorn & Wasi 1969). The blood was centrifuged to remove plasma, and the red cells were washed three times in balanced saline solution (NaCl 0.13 M, KCl 0.005 M, $MgCl_2.6H_2O$, 0.0074 M) after which they were suspended in an incubation mixture similar to that described by Lingrel & Borsook (1963) but modified for the use of human red cells (Weatherall *et al.* 1969). After the addition of labelled amino acids in quantities which will be described for each individual experiment the cells were incubated for periods ranging from 5 min to 2 h, and then washed three times in balanced saline, and frozen. In some experiments individual haemoglobins were then isolated by Amberlite chromatography while in others whole-cell lysates were converted to 'globin' without further purification and the constituent globin chains separated by CM-cellulose chromatography as described above. After chromatography the incorporated radioactivity was determined by pipetting 0.2 ml aliquots from each fraction into 1 ml of water and 10 ml of Bray's solution (Bray 1960) and counting in a Packard liquid scintillation spectrometer. When it was required to determine the specific activity of whole haemoglobin the purified haemoglobin fractions obtained by chromatography were converted to globin and the specific activity determined by dissolving the globin in 0.5 % formic acid and determining the absorbance at 280 nm and incorporated radioactivity. Appropriate corrections were made

for the number of leucine residues in different haemoglobins and globin chains and for the different optical properties of the α, β and γ chains at 280 nm (Schwartz 1969).

In experiments in which haemoglobin synthesis was studied in blood samples with a very low reticulocyte count approximately 6 ml of red cells were centrifuged at 35000 *g* for 1 h and the top 1 ml of reticulocyte-rich cells was removed for synthesis studies.

Bone-marrow samples were obtained by iliac-crest puncture and collected into heparinized saline after which they were incubated in an identical way to that described for the peripheral blood. Prior to purification an appropriate amount of haemoglobin solution prepared from the peripheral blood was added to the washed bone-marrow pellet as a source of carrier.

4. Results

The initial isolation and characterization of what became known as Hb Constant Spring was carried out on blood samples obtained from a large family with several members with Hb H disease from Kingston, Jamaica. The family pedigree is shown in figure 1.

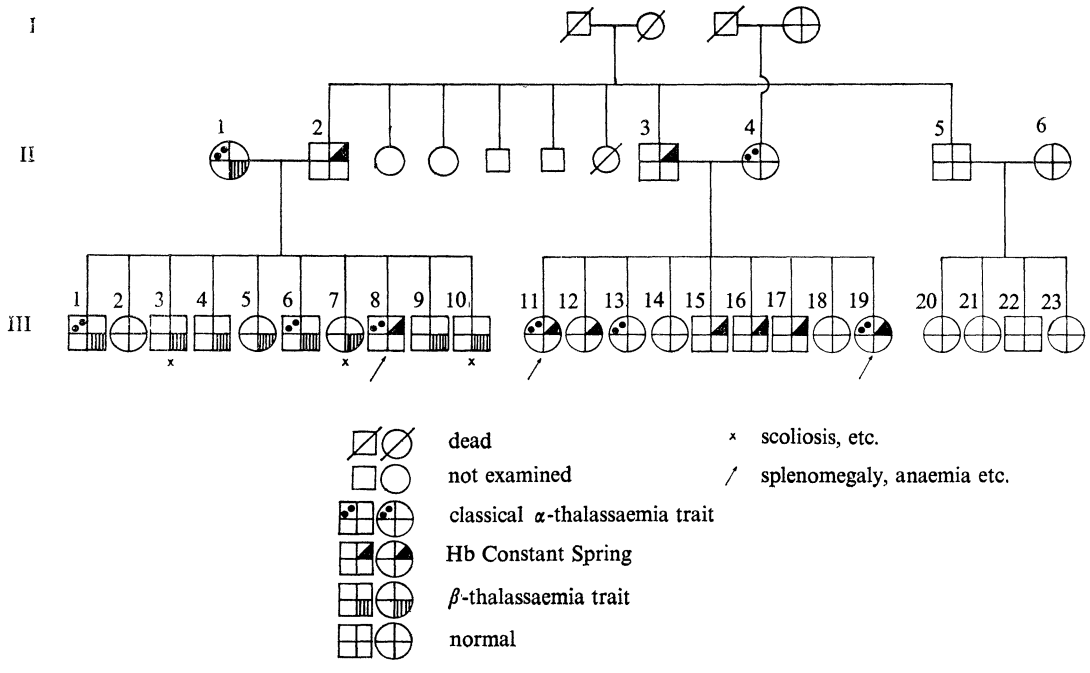


Figure 1. Pedigree of family C from Constant Spring, Jamaica. The arrows indicate the three individuals heterozygous for both Hb CS and α-thal. 1 who have haemoglobin H disease. The family was ascertained through individual III.19.

(*a*) *Electrophoretic characterization*

In family C (figure 1) three siblings (III.8 and III.11 and III.19) showed haemoglobin patterns typical of Hb H disease on starch-gel electrophoresis with the tris–EDTA–borate system, pH 8.5. However, in addition to Hbs A, A_2, H and Bart's small amounts of two slowly-migrating components were observed between the origin and the Hb A_2 zone (figure 2, plate 22). When structural studies indicated subsequently that these components were derived from a unique α-chain variant, which we named 'Hb Constant Spring' (Hb CS), they were designated Hbs CS_1 and CS_2 respectively. This electrophoretic pattern was found to be

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

TABLE 1. HAEMATOLOGICAL AND ELECTROPHORETIC FINDINGS IN MEMBERS OF FAMILY C (FIGURE 1)

(All haematological and electrophoretic data were collected in Jamaica except for individual III.5 who was in England. Blood was collected from the latter and examined in Liverpool. β4, Hb H; γ4, Hb Bart's.)

pedigree designation		age (years)	Hb g/dl	10^6 r.b.cs per μl	p.c.v. (%)	m.c.h.c. (%)	m.c.h. pg	m.c.v. fl	retics. (%)	mean cell fragility (% NaCl)	Hb H inclusions	starch-gel electrophoresis pattern	quantitative cellulose–acetate electrophoresis (% Hb) β4+γ4	A_2	CS	alkali-resistant Hb (%)	probable genotype
I	4	62	14.1	4.24	41	34	33	97	0.8	0.38	neg	A+A_2	.	2.2	.	0.8	normal
II	1	50	12.8	5.84	42	30	22	73	0.4	0.37	+	A+A_2	.	2.3	.	0.7	α-thal.? β-thal.
	2	54	14.0	5.20	42	33	27	81	1.1	0.41	neg	A+A_2+CS	.	2.0	< 1	0.9	Hb CS trait
	3	45	14.0	5.00	43	33	28	87	1.0	0.42	neg	A+A_2+CS	.	2.1	< 1	1.3	Hb CS trait
	4	40	12.8	5.64	41	31	23	74	1.8	0.35	+	A+A_2	.	1.9	.	0.8	α-thal. trait
	5	38	14.8	5.14	44	33	29	87	1.0	0.42	neg	A+A_2	.	2.5	.	0.6	normal
	6	36	13.6	.	40	34	.	.	0.8	.	neg	A+A_2	.	2.0	.	0.5	normal
III	1	28	14.0	5.85	43	32	24	74	0.2	0.34	+	A+A_2	.	5.4	.	0.7	α-thal. β-thal.
	2	26	13.3	4.30	39	34	31	91	1.0	0.38	neg	A+A_2	.	1.4	.	0.4	normal
	3	25	12.5	6.34	40	31	20	64	0.8	0.32	neg	A+A_2	.	5.5	.	1.9	β-thal. trait
	4	23	13.0	6.40	41	31	20	64	0.5	0.31	neg	A+A_2	.	6.1	.	1.6	β-thal. trait
	5	22	12.4	.	39	32	.	.	4.0	.	neg	A+A_2	.	4.9	.	2.1	β-thal. trait
	6	21	14.5	6.66	46	31	22	69	0.8	0.36	+	A+A_2	.	5.9	.	2.0	α-thal. β-thal.
	7	18	11.3	5.38	36	31	21	68	1.6	0.35	neg	A+A_2	.	5.1	.	2.9	β-thal. trait
	8	16	9.6	4.98	39	24	19	73	6.0	0.36	+++	β4+γ4+A+A_2+CS	18.6	.	1.7	1.9	α-thal. Hb CS
	9	15	11.9	6.05	39	30	20	65	1.0	0.36	neg	A+A_2	.	5.5	.	1.0	β-thal. trait
	10	9	11.1	5.65	36	31	20	64	0.6	0.28	neg	A+A_2	.	5.4	.	1.2	β-thal. trait
	11	22	9.7	4.44	40	24	22	87	5.5	0.32	+++	β4+γ4+A+A_2+CS	13.2	.	2.5	2.4	α-thal. Hb CS
	12	20	12.5	4.54	39	32	28	84	1.2	0.41	neg	A+A_2+CS	.	2.3	< 1	0.8	Hb CS trait
	13	19	13.1	5.83	43	30	22	74	1.6	0.34	+	A+A_2	.	2.2	.	0.8	α-thal. trait
	14	18	14.2	4.25	41	34	33	96	0.8	0.40	neg	A+A_2	.	3.0	.	0.6	normal
	15	16	13.1	4.60	39	33	28	86	1.0	.	neg	A+A_2+CS	.	.	< 1	0.2	Hb CS trait
	16	14	13.5	4.74	41	33	28	86	1.4	.	neg	A+A_2+CS	.	.	< 1	0.3	Hb CS trait
	17	12	12.6	5.06	42	30	26	80	1.0	0.37	neg	A+A_2+CS	.	.	< 1	0.4	Hb CS trait
	18	10	13.3	4.83	42	31	27	87	0.5	0.38	neg	A+A_2	.	3.0	.	0.4	normal
	19	8	8.5	4.52	34	25	19	77	12.0	0.37	+++	β4+γ4+A+A_2+CS	18.8	.	1.7	4.2	α-thal. Hb CS
	20	10	14.1	.	41	34	.	.	0.5	.	neg	A+A_2	.	2.5	.	0.7	normal
	21	8	13.3	.	40	33	.	.	0.3	.	neg	A+A_2	.	2.1	.	0.7	normal
	22	5	13.8	.	42	33	.	.	0.2	.	neg	A+A_2	.	3.0	.	0.9	normal
	23	3	13.0	.	38	34	.	.	0.4	.	neg	A+A_2	.	2.8	.	0.8	normal

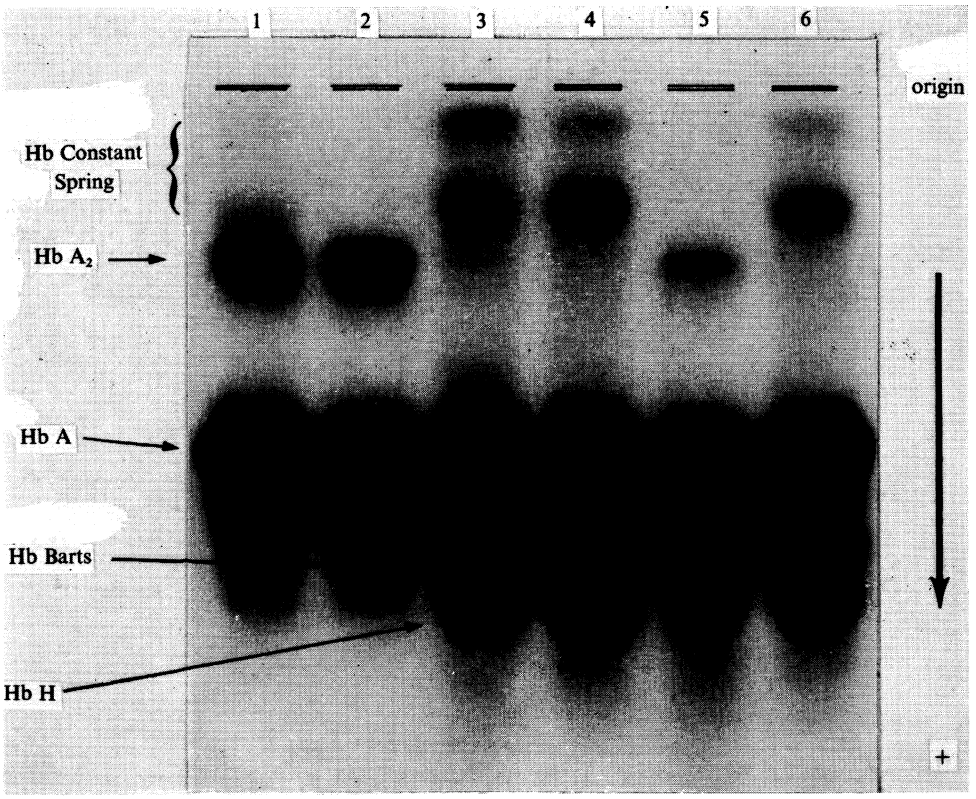


FIGURE 2. Haemoglobin analysis of members of family C. The lysates shown are from the following (left to right): 1, an Hb-CS heterozygote; 2, a normal individual; 3 and 4, siblings heterozygous for α-thal. 1 and Hb CS; 5, an individual with Hb H disease without Hb CS (control not from family C); 6, the third sibling heterozygous for Hb CS and α-thal. 1. Vertical descending starch-gel electrophoresis, tris–EDTA–borate buffer system, pH 8.5, benzidine stain.

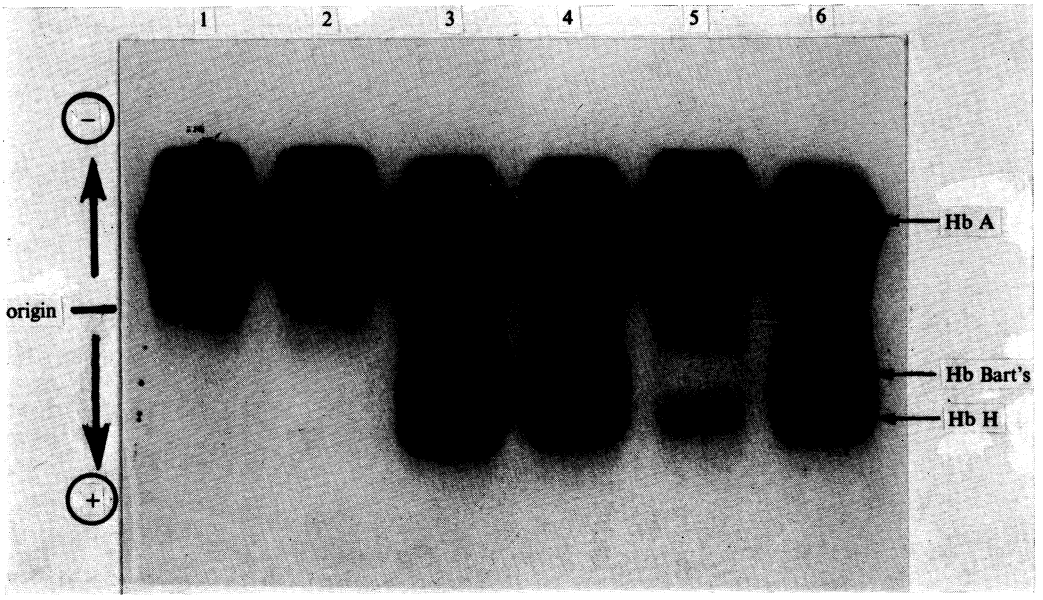


FIGURE 3. Haemoglobin analysis of members of family C. The lysates are from the following (left to right): 1 and 2, normal individuals; 3 and 4, individuals heterozygous for α-thal. 1 and Hb CS; 5, an individual with Hb H disease without Hb CS (control); 6, an individual heterozygous for α-thal. 1 and Hb CS. The anodal migration at pH 7.0 is characteristic of Hbs H and Bart's and Hb CS did not separate under these conditions. Vertical descending starch-gel electrophoresis, phosphate buffer system, pH 7.0, benzidine stain.

(*Facing p.* 418)

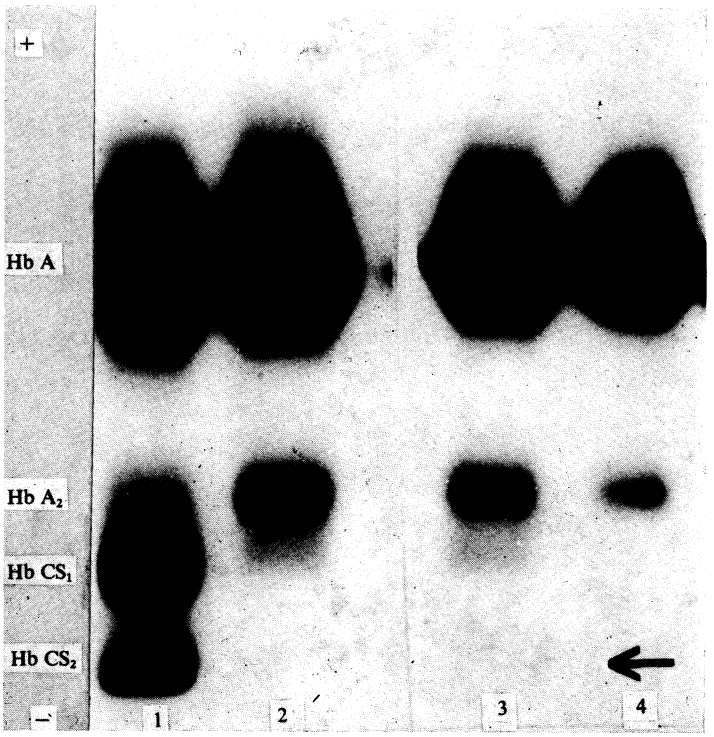


FIGURE 4. Haemoglobin analysis of lysates from members of the family with an Hb-CS homozygote. The lysates are from the following (left to right): 1, the child homozygous for Hb CS; 2 and 3, the heterozygous parents; 4, a normal individual. In addition to Hbs CS_1 and CS_2 there were some faint bands visible behind the origin which were not strong enough to photograph. Vertical descending starch-gel electrophoresis, tris–EDTA–borate buffer system, pH 8.5, benzidine stain.

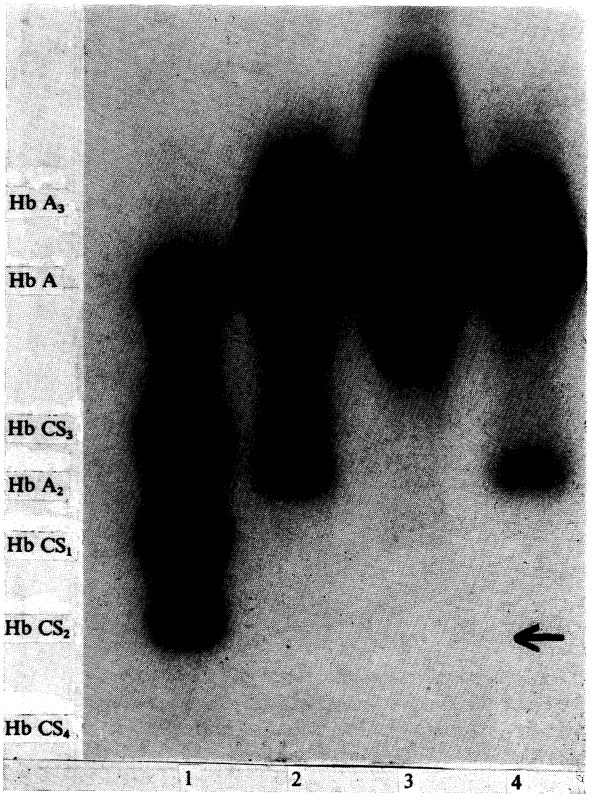


FIGURE 5. For description see opposite.

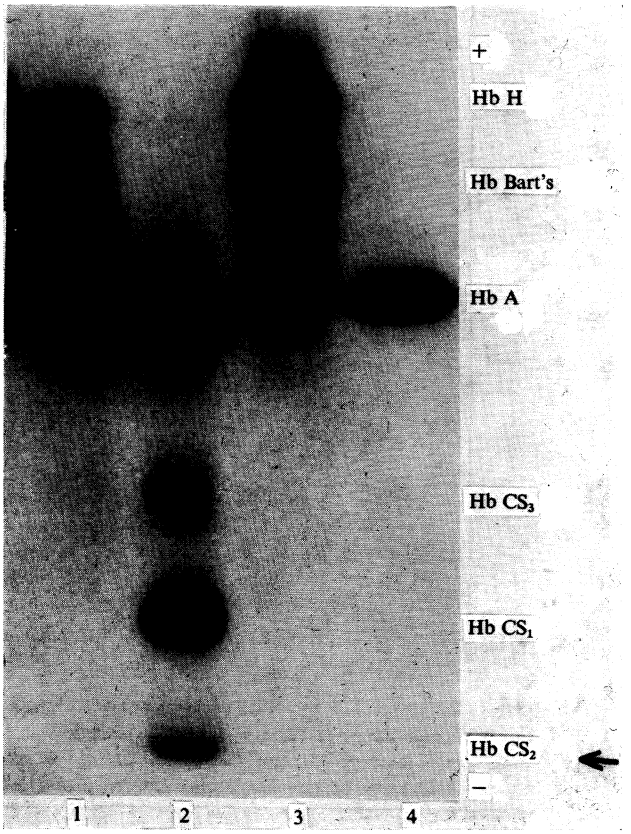


FIGURE 6. For description see opposite.

reproducible in fresh lysates prepared from the blood of these siblings, although when the lysates were kept a few days there was an increase in the concentration of the more cathodal band (Hb CS_2) with the appearance of an extra band (or bands) migrating just behind the origin (Hb CS_4), and a component which migrated slightly more rapidly than Hb A_2 at pH 8.5 was also observed. This fraction, which was only resolved completely from Hb A_2 by Amberlite chromatography, we called Hb CS_3. Starch-gel electrophoresis with a phosphate-buffer system, pH 7.0, showed a pattern typical of Hb H disease and the minor components did not separate from Hb A (figure 3, plate 22).

On cellulose-acetate electrophoresis the two main Hb CS fractions did not separate but migrated as a single band. However they were well resolved from Hb A_2 on this system and thus by eluting the Hb CS zone and the other haemoglobins separately it was possible to quantitate the various haemoglobin fractions. The data are summarized in table 1. The combined levels of Hbs CS_1 and CS_2 ranged from 1.7 to 2.5 % of the total haemoglobin but the levels of Hb A_2 were too low for accurate quantitation. Hbs H and Bart's were eluted as a single component from the same strips and their combined levels ranged from 13.2 to 18.8 %.

Five family members appeared to be carriers of Hb CS without having Hb H disease. On starch-gel electrophoresis the red-cell lysates of these individuals showed Hbs A and A_2 and trace amounts of Hbs CS_1 and CS_2. It was impossible to obtain accurate estimations of the level of CS components, but repeated cellulose-acetate electrophoretic analysis gave values ranging between 0.5 and 0.9 % and the level never exceeded 1 % of the total haemoglobin; the level of Hb A_2 was normal (table 1). Genetic studies, described in a later section, provided clear evidence that these individuals are heterozygous carriers of Hb CS.

The electrophoretic findings in a Malay child apparently homozygous for Hb CS, and his parents, are shown in figure 4, plate 23. Both parents were heterozygous for Hb CS and the child showed relatively large amounts of Hbs CS_1 and CS_2; quantitative cellulose-acetate electrophoresis gave values of 2.1, 4.5 and 2.2 % for Hbs A_2, CS_1 and CS_2 respectively. Trace amounts of Hbs CS_3 and CS_4 were also seen on starch-gel electrophoresis but did not separate well enough on cellulose-acetate electrophoresis for quantitation. Starch-gel electrophoresis, pH 7.0, showed trace amounts of a component with the electrophoretic mobility of Hb Bart's.

All the alkaline starch-gel electrophoretic runs of haemolysates from individuals with Hb H disease with Hb CS, and from the child homozygous for Hb CS showed traces of a

Description of plate 23

Figure 5. Electrophoretic analysis of fractions obtained by Amberlite chromatography of red-cell lysates from an individual homozygous for Hb CS. The fractions are as follows (left to right): 1, the material retained on the column at 24 °C and eluted in 0.2–0.5 M NaCl; the Hb-$CS_{1,2,3}$ components are shown clearly while Hb CS_4 appeared as a faint band at the origin; 2, the fraction eluted at 24 °C showing Hbs A and A_2; 3, the fraction which was not retained at 4 °C showing Hbs Bart's, A_3 and F. In these overloaded conditions Hb Bart's did not separate from Hb A_3 but was seen on starch-gel electrophoresis, pH 7.0; 4, a normal adult lysate to act as a control showing the positions of Hbs A and A_2.

Figure 6. Electrophoretic analysis of fractions obtained by Amberlite chromatography of a lysate prepared from an individual with Hb H disease and Hb CS. The fractions from left to right are: 1, a marker with Hbs A, H and A_2; 2, the material which was retained on the column at 24 °C showing Hbs CS_1, CS_2 and CS_3; 3, the fraction unretained at 4 °C showing Hbs H and Bart's with a small amount of Hb A; 4, the fraction which eluted at 24 °C showing predominantly Hb A.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

fraction migrating just cathodally to Hb A (figure 2). This did not separate completely from Hb A and it was not obvious whether it was Hb F or another Hb CS component. Analysis of the various fractions prepared by Amberlite chromatography provided clear evidence that this is foetal haemoglobin.

(b) Isolation of Hb CS

A variety of electrophoretic and chromatographic techniques were tried in initial attempts to isolate Hb CS. In preliminary experiments in which globin prepared directly from red-cell lysates was used, adequate analytical separations of the αCS chains were obtained, but the very small amounts of αCS as compared with αA chains required heavy overloading of preparative columns and the resulting αCS chains were then too impure for structural studies. For this reason a preliminary purification step was necessary and it was found that Amberlite IRC-50 chromatography under the conditions described in the Methods section gave the best yield of Hb CS.

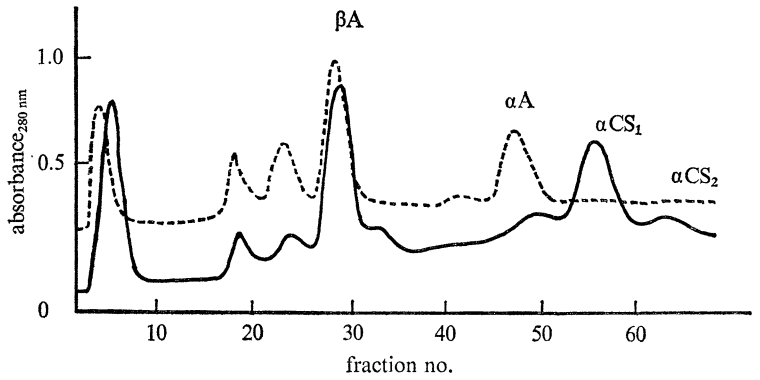


FIGURE 7. Comparative analytical chain separations on CM-cellulose in 8 M urea–mercaptoethanol–phosphate buffer, pH 6.8, of globin prepared from purified Hbs CS and A. The continuous line represents Hb CS and shows a predominance of αCS_1 and a small amount of αCS_2 chain. The broken line is the Hb A profile.

Representative electrophoretic runs showing the various fractions isolated from lysates of individuals with Hb H disease and Hb CS, and from the Hb CS homozygote, are shown in figures 5 and 6, plate 23. The material eluted at 4 °C consisted of Hbs H and Bart's with variable amounts of Hb F. In addition many non-haemoglobin proteins including carbonic anhydrases B and C were present in this fraction. The material which eluted at 24 °C contained Hbs A and A_2 thus confirming the presence of the latter in the individuals with Hb H disease and Hb CS, a fact which was not always obvious on standard starch-gel electrophoresis where the Hb A_2 appeared to merge into the Hb CS_1 band (figure 2, plate 22). The relative amounts of the Hb CS fractions obtained by Amberlite chromatography varied according to the state of the starting material (figures 5 and 6, plate 23). In some cases good yields of Hbs CS_1 and CS_2 were obtained; in other experiments Hb CS_1 was obtained in low yield with relatively large amounts of Hbs CS_2 and CS_4. Hb CS_3 was always present, again in variable amounts. In addition, particularly in lysates from the Hb CS homozygote, several minor fractions were observed which migrated more cathodally than Hb CS_4 on alkaline starch-gel electrophoresis. It should be noted that all the Hb CS components obtained by Amberlite chromatography were benzidine positive and no non-haemoglobin proteins appeared to elute with Hb CS in this system.

For further analysis the fractions obtained by Amberlite chromatography were dialysed for 24 h against water and then converted to globin by acid–acetone precipitation. The globin chains were then separated by CM-cellulose chromatography with the 8 M urea-mercapto-ethanol-phosphate buffer system pH, 6.8. Typical chain-separation profiles are shown in figure 7. Depending on the purity of the Hb CS fractions a relatively small peak eluting in the position of normal α-chains was observed. In addition there were always two, and sometimes three, abnormal α-chain components which eluted after the normal αA chains. The yield of these fractions, which were called αCS_1, αCS_2 and αCS_3, varied; in some cases the yield of αCS_1 was considerably greater than that of αCS_2 although in others more αCS_2 than CS_1 was obtained (figure 7). It seems likely that some of this variability resulted from alteration of the αCS chains during the preparative steps. When whole-cell globin was chromatographed the αCS_1 and αCS_2 fractions were usually present in the quantities expected from the level of Hbs CS_1 and CS_2 in fresh red-cell lysates (figure 8). However as mentioned above the αCS chains prepared in this way were contaminated with αA chains. The αCS_3 component corresponded to Hb CS_3 which migrated close to Hb A_2 on starch-gel electrophoresis. No obvious αCS-chain fraction corresponding to Hb CS_4 was seen in these chain separations.

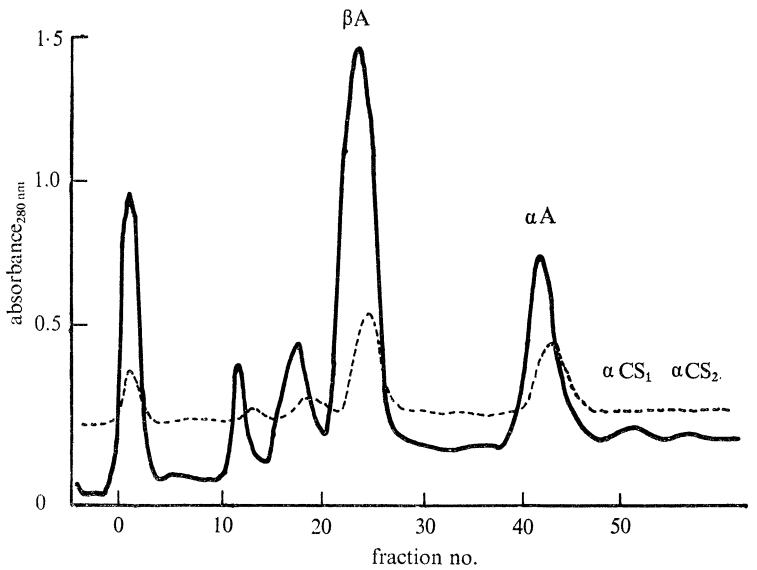


FIGURE 8. Elution profile obtained on separating the globin chains of an individual heterozygous for Hb CS and α-thal. 1 by CM-cellulose chromatography in an 8 M urea-mercaptoethanol-phosphate buffer system, pH 6.8. The αCS_1 and CS_2 fractions are shown as small peaks just after the main αA zone. Normal globin (dashed line) was run as a control.

(*c*) *Structural analysis*

Preliminary structural investigations were carried out with purified AEαCS_1. A tracing of a typical peptide map of AEαCS_1 chain is shown in figure 9. Comparison with a control tryptic digest of AEαA showed that all the normal tryptic α-chain peptides were present in the AEαCS_1 chain digest. In addition 3 additional major peptides, W1, X and Y, were noted, together with a fourth peptide Z which appeared in low and variable yields.

The soluble peptides 1–14 (excluding 12b and 13) from the AEαCS_1 chain digest had normal compositions. The insoluble 'core' peptide (12b.13) was isolated in good yield from

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

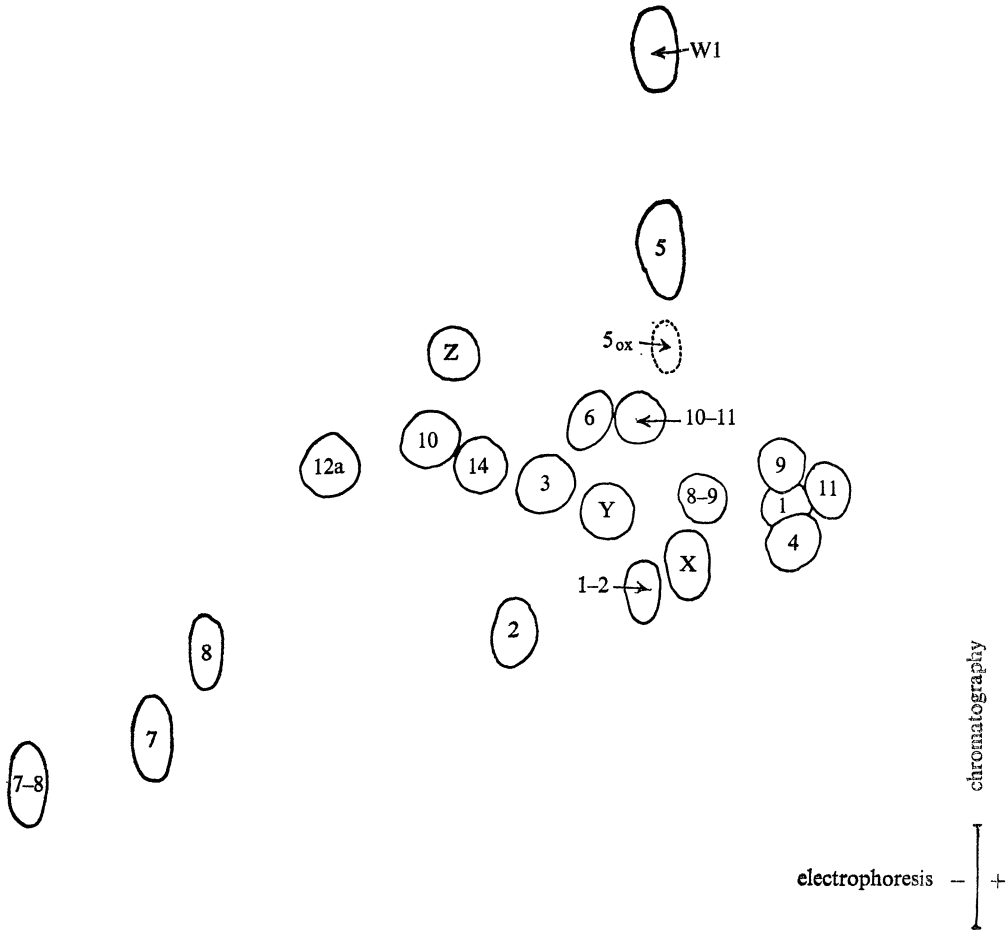


FIGURE 9. Tracing of a fingerprint of the tryptic peptides of the αCS_1 chain. (Peptide W2, unique to the αCS_2 chain, runs midway between W1 and Z.)

TABLE 2. AMINO-ACID COMPOSITIONS OF THE FOUR TRYPTIC PEPTIDES UNIQUE TO THE αCS_1 CHAIN

	W1	X	Y	Z
$CySO_3$	.	.	.	.
Asp	.	.	.	.
Thr	.	.	.	.
Ser	(1.11) 1	(0.98) 1	(0.90) 1	(0.85) 1
Glu	(1.20) 1	(0.77) 1	(0.95) 1	.
Pro	(2.00) 2	(1.82) 2	.	(0.80) 1
Gly	.	(0.97) 1	.	.
Ala	(0.98) 1	(4.24) 4	(1.00) 1	(0.91) 1
Val	(0.84) 1	(2.00) 2	.	.
Met	.	.	.	.
Ileu	.	.	.	.
Leu	(4.02) 4	.	.	(2.82) 3
Tyr	.	.	.	.
Phe	(2.00) 2	.	.	.
Lys	.	.	.	.
His	(0.95) 1	.	.	(1.00) 1
Arg	(0.90) 1	(0.96) 1	(1.07) 1	(1.02) 1
Trp	–	–	+	–

a tryptic digest by precipitation at pH 6.5 followed by heating at 100 °C for 2 min. The precipitate was digested with chymotrypsin and the peptides so obtained fractionated in the pH 4.7 BAWP peptide-mapping system. The peptide map was identical to one obtained from a control digest of normal AEα-chain peptide 12b.13. In addition, the amino-acid compositions of the chymotryptic peptides from AEαCS_1 chain 12b.13, representing the sequence from residues 106 to 138, were identical to those isolated from similar digests of αA12b.13.

From the amino-acid compositions of peptides W1, X, Y and Z given in table 2, it was evident that none of them resembled other α-chain peptides, and in fact there are no similarities to any known human globin-chain peptides. It appeared therefore that the αCS_1 chain contained extra peptide material in comparison to normal α chains. Inasmuch as the normal α-chain sequence was fully represented by its tryptic peptides and there was no evidence from the peptide maps and amino-acid compositions of an interruption in the sequence of residues 1 through 141 it seemed likely that the additional peptides were probably located at the N- or C-terminal end of the αCS chain.

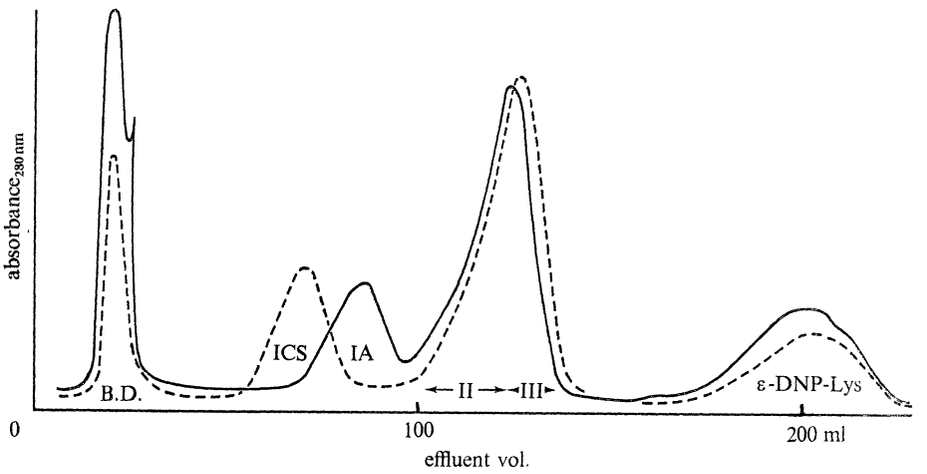


FIGURE 10. Elution profiles of CNBr fragments of αA and αCS_1 chains after chromatography on Sephadex G-100 in 5% HCOOH. Solid line, αA; dotted line, αCS_1. Blue Dextran and ε-DNP-Lysine were added as molecular mass markers.

Location of the extra peptides. The normal human α chain contains methionine residues at positions 32 and 76. Reaction with cyanogen bromide ought therefore to produce 3 fragments of 32, 44 and 65 residues. Trial experiments showed that these could be fractionated on Sephadex G-100 in 5% formic acid (figure 10). Fractions were pooled as indicated and then freeze dried, digested with trypsin and peptide maps prepared. From these digests it was obvious that the first symmetrical peak (fraction I) comprised the 76-residue C-terminal fragment. The leading edge of the second peak (fraction II) gave a fairly pure sample of the 44-residue internal fragment, while the trailing edge (fraction III) contained only the 32-residue N-terminal fragment. A similar experiment was carried out on 10 mg of a CNBr-digest of AEαCS_1 chain (figure 10). The elution pattern was similar to that for the αA digest except that fraction I eluted earlier, suggesting that it was longer than the corresponding fragment in αA. The three fractions were pooled as indicated, dried and digested with trypsin. Fractions IICS and IIICS gave peptide maps similar to those obtained from fractions II and III in the

αA CNBr digest. Fraction ICS contained the C-terminal peptides 10, 11, 12a, 12b.13 and 14 and in addition the three major characteristic peptides of the αCS_1 chain, W1, X and Y indicating therefore that the extra peptides of the αCS_1 chain were attached to the C-terminal end of a normal α chain. These indications that the α chain of Hb CS was longer than a normal α chain were confirmed by SDS-gel electrophoresis of purified Hb CS globin samples (Clegg *et al.* 1971).

Amino-acid sequences of the αCS_1-chain peptides W1, X, Y and Z. The amino-acid compositions of these peptides are given in table 2. Peptide Z was usually present in low and variable yields and it seemed likely that this might be a partial digestion product, possibly derived from W1 since all the residues present in Z could be accounted for by the composition of W1. Furthermore, the amino-acid composition of αCS_1, based on residues not contained in peptides W1, X, Y or Z (i.e. Asp, Thr and Tyr) indicated that αCS_1 contained approximately 30 extra residues, about the total of residues contained in peptides W1, X and Y. This also suggests that Z was a partial digestion product. All four of these peptides were isolated by preparative two dimensional paper electrophoresis at pH 4.7 followed by chromatography in BAWP. Purity was confirmed by amino-acid analysis before any subsequent manipulations were performed.

Peptide X. The sequence of the first 9 residues, determined by the 'dansyl'-Edman method and corroborated at each stage by amino-acid analysis of the remaining peptide, was found to be Glx.Ala.Gly.Ala.Ser.Val.Ala.Val.Pro. Inasmuch as X was basic at pH 6.5 the N-terminal residue must be Gln rather than Glu. The identification of successive N-terminal residues became equivocal after the 8th stage of degradation.

A sample of X was therefore digested with thermolysin and the peptides fractionated by electrophoresis at pH 3.5 followed by chromatography in BAWP. Three major peptides were obtained, one of which gave a positive stain for arginine and had a composition $Pro_2Ala_2Val_2Arg$. The N-terminal amino acid was Val. Six cycles of degradation gave the sequence Val.Ala.Val.Pro.Pro.Ala.Arg. The N-terminal sequence of this peptide overlaps with the N-terminal sequence of X to give a complete sequence for X of Gln.Ala.Gly.Ala.Ser.Val.Ala.Val.Pro.Pro.Ala.Arg.

Peptide Y. This peptide gave a grey colour with ninhydrin and a positive tryptophan stain but had no detectable N-terminal amino acid after acid hydrolysis of the 'dansyl'-peptide. A sample of dansyl-Y was therefore digested with chymotrypsin and the digestion products fractionated by thin-layer chromatography (Woods & Wang 1967). A spot was observed running in the position of DNS-tryptophan, which was not present in an undigested control. Four cycles of degradation and corroborative amino-acid analysis on another sample of Y gave the sequence (Trp).Ala.Ser.Glx.Arg. Since Y is basic at pH 6.5 the Glx must be Gln and sequence of Y is Trp.Ala.Ser.Gln.Arg.

Peptides W1 and Z. The N-terminal sequences of both these peptides, determined by the 'dansyl'-Edman method and corroborated by amino-acid analysis, was Ala.Leu.Leu.Pro. The next cycle failed to produce any detectable DNS-amino acid and equivocal results were also obtained from amino-acid analysis. However at the next stage DNS-Leu was obtained in good yield. Further degradations gave equivocal results. Two further attempts were made to carry the sequence determination through the 5th stage of degradation but in each case no identifiable DNS-amino acid was observed. Good yields of DNS-Leu were again obtained at the 6th cycle.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

A sample of W1 was digested with chymotrypsin and the peptides fractionated by electrophoresis at pH 6.5, followed by chromatography in BAWP. One acidic and two basic peptides were obtained in good yield. Their compositions are given in table 3. They account for the entire amino-acid composition of W1.

W1chyB1. The N-terminal sequence of W1chyB1 was Ala.Leu.Leu so this is obviously the N-terminal fragment of W1 which must therefore have the partial sequence Ala.Leu.Leu.Pro-(His or Ser).Leu.(His or Ser). Digestion of an aliquot with carboxypeptidase A released free leucine and histidine as the only detectable amino acids. A partial sequence for the N-terminal end of W1 was thus established as Ala.Leu.Leu.Pro.(Ser).Leu.His.

Table 3. Amino-acid compositions of peptides derived from W1 by chymotrypsin digestion

	W1chyB1	W1chyB2	W1chyA1	W1
$CySO_3$	.	.	.	.
Asp	.	.	.	.
Thr	.	.	.	.
Ser	(0.99) 1	.	.	1
Glu	.	.	(0.93) 1	1
Pro	(0.91) 1	(0.71) 1	.	2
Gly	.	.	.	.
Ala	(1.00) 1	.	.	1
Val	.	.	(0.87) 1	1
Met	.	.	.	.
Ileu	.	.	.	.
Leu	(2.99) 3	.	(0.60) 1	4
Tyr	.	.	.	.
Phe	.	(0.90) 1	(1.00) 1	2
Lys	.	.	.	.
His	(1.03) 1	.	.	1
Arg	.	(1.00) 1	.	1
Trp	.	.	.	.

W1chyB2 had N-terminal arginine. This was confirmed by amino-acid analysis of the remaining peptide after the first degradation cycle. Subsequent cycles of 'dansyl'-Edman degradation established the complete sequence as Arg.Pro.Phe.

W1chyA1. Four cycles of 'dansyl'-Edman degradation and amino-acid analysis gave the sequence Leu.Val.Phe.Glu for this peptide.

Peptide Z has the same N-terminal sequence as W1 and a composition which is identical to that of W1chyB1 except for the presence of arginine. This suggests that the sequence of Z is Ala.Leu.Leu.Pro.Ser.Leu.His.Arg. The C-terminal arginine thus provides an overlap with W1chyB2 to give the sequence Ala.Leu.Leu.Pro.Ser.Leu.His.Arg.Pro.Phe. Finally, since the compositions of the three peptides W1chyB1, W1chyB2 and W1chyA1 account for all the residues in W1, W1chyA1 must be C-terminal, leading to the overall sequence Ala.Leu.Leu.Pro.Ser.Leu.His.Arg.Pro.Phe.Leu.Val.Phe.Glu. Corroboration of this arrangement comes from the sequence of peptide W2 (see below).

Order of peptides W1, X and Y. 20 mg of AEαCS1 chain was digested with chymotrypsin. A peptide map, when compared with a control αA digest revealed the presence of three new Arg-positive spots (one of which was also Trp-positive) and the loss of the free arginine derived from the cleavage of the Tyr.Arg bond at 140.141 in the normal αA chain. The remainder of the digest was fractionated by electrophoresis at pH 6.5. The Arg-positive

regions were located on a guide strip and the bands cut out and subjected to descending chromatography in *n*-butanol–acetic acid–water (3:1:1 by vol.) for 24 h. The Arg-positive bands thus obtained were again located with a guide strip and the peptides finally purified by electrophoresis at pH 2.1.

Of the 5 peptides, two had compositions corresponding to residues 30–33 and 92–98 of the αA chain. The compositions of the three remaining Arg-positive peptides are given in table 4.

TABLE 4. AMINO-ACID COMPOSITIONS OF CHYMOTRYPTIC PEPTIDES UNIQUE TO THE αCS_1 CHAIN

	αCSchyA	αCSchyB	αCSchyC
$CySO_3$	.	.	.
Asp	.	.	.
Thr	.	.	.
Ser	(0.72) 1	(0.80) 1	.
Glu	(0.94) 1	(0.90) 1	.
Pro	.	(1.84) 2	(0.81) 1
Gly	.	(1.11) 1	.
Ala	(2.12) 2	(3.77) 4	.
Val	.	(2.00) 2	.
Met	.	.	.
Ileu	.	.	.
Leu	(1.17) 1	.	.
Tyr	.	.	.
Phe	.	.	(1.00) 1
Lys	.	.	.
His	.	.	.
Arg	(1.00) 1	(1.71) 2	(1.00) 1
Trp	−	+	−

αCSchyA had an N-terminal sequence Ala.Ser.Glx and Leu was assumed to be C-terminal on the basis of chymotrypsin specificity so the overall sequence was Ala.Ser.Glx.(Arg.Ala).Leu.

αCSchyB had the same composition as αCS X except for the presence of an additional arginine residue and tryptophan, which was assumed to be C-terminal on the basis of chymotrypsin specificity (peptide X is not split by chymotrypsin). The N-terminal residue was arginine so its sequence must be Arg.(X).Trp.

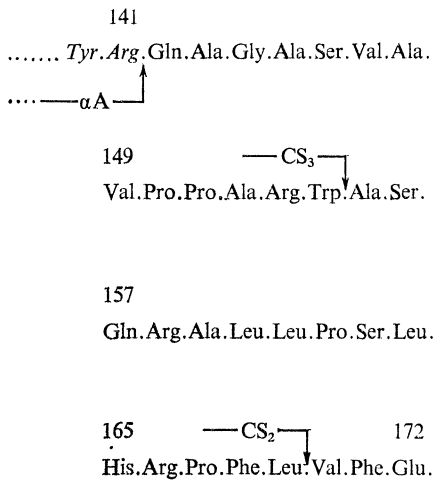


FIGURE 11. Amino-acid sequence of the C-terminal residues of the αCS chains.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

These two peptides establish the overlaps between peptides X, Y and W1 and the C-terminal Arg (residue 141) of the αA chain. The sequence of αCSchyA overlaps the 4 C-terminal residues of Y with the N-terminal dipeptide of W1 and is compatible only with the ordering Y-W1. Similarly, the composition and N-terminal Arg of αCSchyB indicate that this peptide overlaps the C-terminal Arg residue of α14, X and the N-terminal Trp of Y. The ordering of the C-terminal peptides of αCS1 is therefore α14-X-Y-W1 and the complete sequence is as shown in figure 11.

TABLE 5. AMINO-ACID COMPOSITION OF PEPTIDE W2 FROM THE αCS_2 CHAIN

$CySO_3$	.
Asp	.
Thr	.
Ser	(0.78) 1
Glu	.
Pro	(1.70) 2
Gly	.
Ala	(0.82) 1
Val	.
Met	.
Ileu	.
Leu	(4.14) 4
Tyr	.
Phe	(0.88) 1
Lys	.
His	(1.07) 1
Arg	(1.00) 1
Trp	.

Structure of αCS_2 and αCS_3 chains. Peptide maps of AEαCS_2 chain were identical to those of the AEαCS_1 chain (figure 9) with one exception; peptide W1 was absent and had been replaced by a more basic peptide W2. Its composition is given in table 5, and it can be seen that this differs from W1 in having one residue fewer of Glu, Val and Phe. It had an N-terminal sequence Ala.Leu.Leu.Pro, identical with that of W1. Since the three residues which constitute the difference between W1 and W2 comprise the C-terminal tripeptide sequence of W1 (and the αCS_1 chain) it was concluded that the αCS_2 chain is a truncated version of αCS_1, terminating at Leu 169 (figure 11).

On peptide maps of the AEαCS_3 chain only αCSX of the 3 major peptides characteristic of αCS_1-chain digests was visible. When αCS_3-chain peptide maps were stained for tryptophan a spot was observed in the neutral region with the same chromatographic mobility as a free tryptophan marker. This spot was not present in control digests of AEαA and AEαCS_1 chains. It was concluded that the αCS_3 chain is a second truncated version of αCS_1, terminating at Trp 154 (figure 11).

We have not yet managed to isolate an α-chain fraction related to Hb CS_4 in order to determine its structure.

(*d*) *Haemoglobin synthesis*

The relative rates of globin-chain synthesis were studied in reticulocytes obtained from several members of family C from Jamaica and from the original family with Hb Athens from Greece. Two members of family C also allowed us to obtain iliac-crest bone-marrow samples to examine haemoglobin synthesis in erythroid precursors.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

TABLE 6. RADIOACTIVITY INCORPORATED INTO THE GLOBIN CHAINS AFTER INCUBATION OF RETICULOCYTES OF INDIVIDUALS HETEROZYGOUS FOR α THAL. 1 AND HB CS WITH VARIOUS RADIOACTIVE AMINO ACIDS FOR TIME PERIODS RANGING FROM 5 TO 120 MIN

(Approximately 1 ml of cells were incubated with 100 μCi of [^{3}H]amino acids after which the cells were washed, lysed and whole-cell globin prepared by acid-acetone precipitation. The globin chains were separated by CM-cellulose chromatography with an 8 M urea–mercaptoethanol–phosphate buffer system and the radioactivity in each fraction was determined.

source of patients with Hb H + Hb CS	isotope	incubation time/min	total radioactivity count per min			ratio
			γ+β	αA	αCS_1 †	α/(β+γ)
Constant Spring (family C)	[^{3}H]leucine	5	10446	5181	56	0.50
		120	154546	64704	751	0.42
Constant Spring (family C)	[^{3}H]arginine	15	2850	1436	14	0.51
Athens	[^{3}H]leucine	5	2217	945	12	0.42
		120	28632	14846	32	0.53

† The radioactivity incorporated into the αCS_1 chain is probably an overestimate because of contamination with small amounts of highly radioactive αA chain.

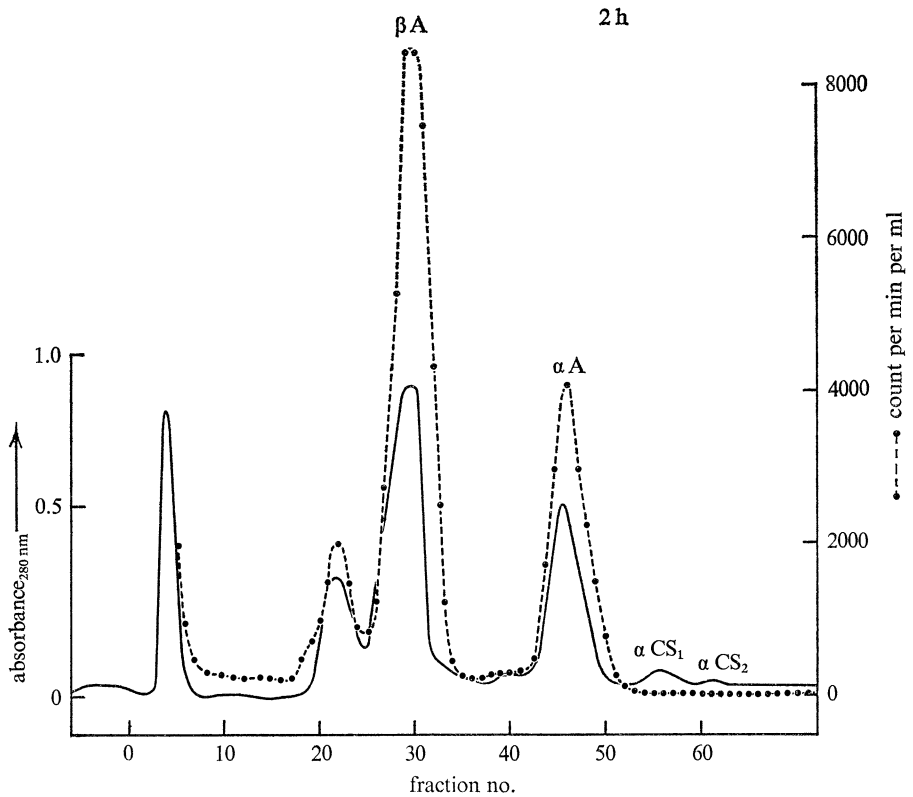


FIGURE 12. Haemoglobin synthesis in reticulocytes from an individual heterozygous for Hb CS and α-thal. 1. The cells were incubated with 100 μCi of [^{3}H]leucine for 2 h and then washed, lysed and whole-cell globin prepared from the lysate without further purification. The chains were separated on CM-cellulose chromatography with the 8 M urea–mercaptoethanol–phosphate buffer system, pH 6.8, and the radioactivity in each fraction determined. There is considerably more radioactivity incorporated into the βA chain than into the αA chain and virtually no radioactivity incorporated into the αCS_1 and CS_2 fractions.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Globin-chain synthesis in individuals with Hb H disease and Hb CS. Reticulocytes from individuals with Hb H disease and Hb CS from family C and the Hb Athens family were incubated with 100 μCi [^{3}H]leucine or [^{3}H]arginine for periods ranging from 5 min to 2 h. The cells were washed, lysed and whole-cell globin precipitated by the acid–acetone method and the radioactivity in globin chains was then determined (table 6 and figure 12). The total radioactivity incorporated into the β and γ chains exceeded that incorporated into the combined αA and αCS chains by a factor of approximately 2/1, a ratio almost identical to that found in Hb H disease in the absence of Hb CS (Weatherall *et al.* 1965; Clegg & Weatherall 1967). The ratios did not alter during incubation periods ranging from 5 min to 2 h suggesting that instability of Hb CS is not a major factor in causing the overall globin-chain imbalance. The αCS chains had low levels of incorporated radioactivity in these experiments; a finding which is discussed in detail in a later section.

TABLE 7. HAEMOGLOBIN SYNTHESIS IN RETICULOCYTES OF AN INDIVIDUAL HETEROZYGOUS FOR BOTH α-THAL. 1 AND Hb CS

Reticulocytes were incubated with 100 μCi of [^{3}H]leucine. The cells were then washed and lysed, the membranes removed by centrifugation, the lysate dialysed against developer 2 and Hbs A, H and CS separated by Amberlite IRC-50 chromatography (see text). The purity of each fraction was checked by starch-gel electrophoresis after which they were converted to globin and the constituent chains separated by chromatography on CM-cellulose columns in an 8 M urea–mecaptoethanol–phosphate buffer system. The chains were dialysed free of urea in 0.5% formic acid and their specific activities determined. The amounts of radioactivity incorporated into the αCS chains together with their low absorbance precluded accurate specific activity determination.

incubation time	specific activity count per min per mg			ratio
min	βA chain	βH chain	αA chain	αA/βA
60	401	3249	2515	6.27

In one experiment the cells from an individual with Hb H disease and Hb CS were incubated for 1 h with [^{3}H]leucine after which they were washed, lysed, and the lysate dialysed for 24 h against developer 2 and the haemoglobin was then separated on Amberlite IRC-50 under the conditions described in § 3. Hbs H and A were isolated, concentrated *in vacuo*, converted to globin and the specific activities of their individual globin chains determined after separation by CM-cellulose chromatography in 8 M urea (table 7). It will be seen that the specific activity of the βH chain exceeded that of the βA chain indicating that the relative turnover rates of Hbs H and A in the cells are different, with Hb A surviving longer in the peripheral blood. In Hb A the specific activity of the α chain exceeded that of the β chain by a factor of about sixfold. These findings have been noted before in patients with Hb H disease (Clegg & Weatherall 1967) and indicate that their red cells contain a relatively large pool of β chains; newly synthesized radioactive α chains combine with unlabelled β chains from this pool to form Hb A, so accounting for the unequal labelling of its chains.

Thus both in terms of the total degree of chain imbalance and in the labelling of the various intracellular pools the biosynthetic findings in individuals with Hbs H and CS are identical to those of individuals with Hb H disease without this variant.

Globin-chain synthesis in Hb CS heterozygotes. The relative rates of α- and β-chain production were measured in peripheral blood samples obtained from 8 individuals from the Jamaican family who are heterozygous for Hb CS. In order to obtain adequate numbers of reticulocytes

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

their red cells were centrifuged as described in §3 and the reticulocyte-rich top 1 ml of cells were used for incubation. The incorporation of [^{3}H]leucine into the α and β chains was determined after a 1 h incubation as described in §3; the α/β-chain-production ratio ranged from 0.72 to 1.03 (mean 0.85). A comparison of these ratios with those of a normal control group and of a group of α-thal. 1 and 2 carriers studied both in the authors' laboratory and by Dr E. Schwartz in Philadelphia is shown in figure 13. The Hb CS heterozygotes have a degree of globin-chain imbalance almost identical to that of the α-thal. 2 carriers. The heterozygous state for Hb CS is thus associated with a small but significant overall reduction in α-chain production and hence with a pattern of haemoglobin biosynthesis indistinguishable from a mild form of α-thalassaemia.

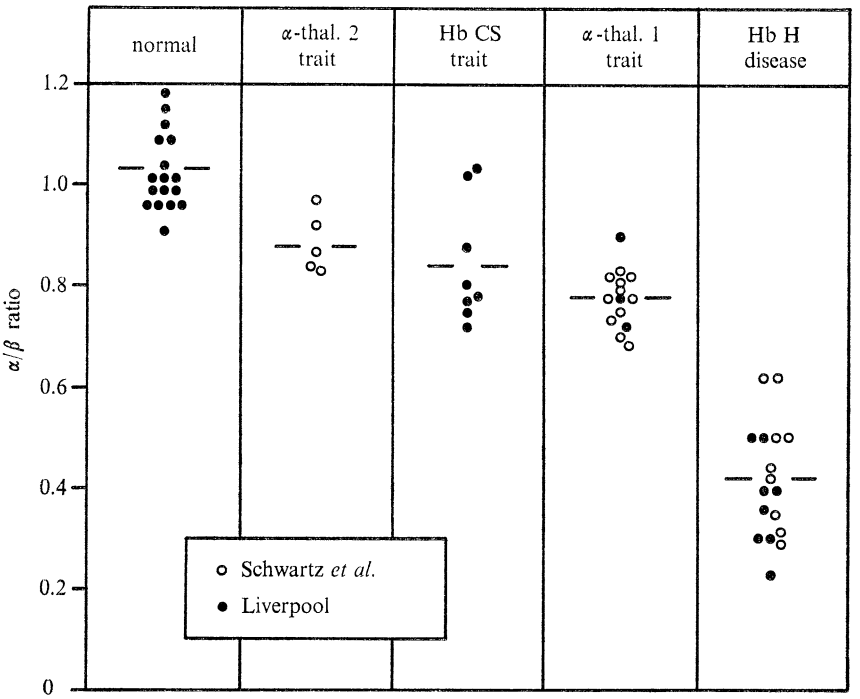


FIGURE 13. A comparison of the rates of synthesis of α and β chains (α/β ratio) in reticulocytes obtained from normal individuals and those with different forms of α-thalassaemia. The horizontal bars represent the mean for each group.

The synthesis of Hb CS during erythroid maturation. In biosynthetic experiments in which whole-cell globin from individuals with Hb H disease and Hb CS was fractionated by CM-cellulose chromatography in 8 M urea, there was very little radioactive label incorporated into the αCS chains at any period of incubation (see figure 12).

Several experiments were performed to try and estimate the total amount of radioactivity incorporated into the αA and αCS chains, or their relative specific activities, after incubation of reticulocytes for various times with radioactive amino acids. In several experiments whole-cell globin was chromatographed on CM-cellulose and the specific activities of the γ, β and αA and αCS chains were determined; in another experiment the individual haemoglobins were first isolated by Amberlite chromatography before chain separation. The results are difficult to evaluate because of the very low amounts of αCS chains present and hence their very low absorbancies after chromatography. Furthermore it is possible that the αCS chains

may have been contaminated with small amounts of highly-labelled αA chains which eluted just before the αCS chains on CM-cellulose. The data summarized in table 8 indicate that the total radioactivity incorporated into the αCS chains is always lower than would be expected from the proportions of Hb CS in the peripheral blood. It can also be seen from this table that the specific activity of the αCS_1 and αCS_2 chains is considerably lower than that of the αA chains, with the ratio αA/αCS ranging from 7.9 to 17.6. Similar results were obtained when the specific activities of the individual peptides of the α chains of Hbs A and CS were measured (see later section). These findings indicate that Hb CS is not unstable and that the αCS chains are not turning over more rapidly than αA chains in the peripheral blood since, if this were the case, they would have a higher specific activity than the αA chains.

TABLE 8. HAEMOGLOBIN SYNTHESIS IN RETICULOCYTES OF AN INDIVIDUAL HETEROZYGOUS FOR α-THAL. 1 AND Hb CS

The reticulocytes were incubated with 50 μCi of [^{3}H]leucine for the times indicated and then the cells were washed, lysed and whole-cell globin prepared from the lysate. The globin chains were separated by CM-cellulose chromatography in an 8 M urea–mercaptoethanol–phosphate buffer system, pooled, dialysed free of urea against 0.5% formic acid and their specific activities determined. Because of the small amounts of αCS_1 and αCS_2 and the small quantities of radioactivity incorporated into these fractions, the estimated specific activities of these components probably is too high; even allowing for this error the specific activities of αCS_1 and αCS_2 are considerably less than that of αA indicating that the synthesis of Hb CS has declined early during erythroid maturation.

incubation time	specific activity count per min per mg					
min	γ chain	β chain	αA chain	αCS_1 chain	αCS_2 chain	$\alpha A/\alpha CS_1$
5	868	1603	1484	188	240	7.9
120	9743	34644	25020	2317	1417	17.6

Rather they indicate that the αCS chains are not being synthesized in reticulocytes in the same relative proportions that they are present in the peripheral blood, i.e. that there is a relative decline in αCS-chain synthesis during erythroid maturation. This was confirmed when the relative rates of synthesis of αCS and αA chains were measured in bone marrow from an individual with both Hb H disease and Hb CS.

An iliac-crest bone-marrow sample was obtained and the cells were washed and incubated with [^{3}H]leucine under the same conditions as those described for the peripheral blood. After incubation for 1 h, the cells were washed, frozen and thawed, and whole-cell globin fractionated on CM-cellulose. The results are shown in figure 14. It can be seen that there are significant amounts of radioactivity incorporated into the αCS fractions; indeed the specific activities of the αA and αCS chains are almost identical indicating that αCS is synthesized in the marrow at a rate commensurate with its levels in the peripheral blood.

The assembly of the αCS globin chains. The results described above indicate that the α chains of Hb CS are produced in low amounts relative to αA chains. A further series of experiments was carried out to try and account for this disparity. One possibility was that the elongated αCS chain is assembled or released more slowly than the αA chain. An estimate of the assembly times of the αCS chain was made by incubating reticulocytes from an individual with Hb H disease and Hb CS for 15 min with [^{3}H]arginine at 37 °C. The αA and αCS chains were isolated, fingerprinted and the arginine-containing peptides were eluted from the fingerprint. A portion of each peptide was used to determine the [^{12}C]arginine concentration

and the remainder was used for determination of the incorporated radioactivity. The result of this experiment is shown in table 9 in which it can be seen that the specific activities of the αA-chain peptides 4, 10 and 14 are very similar, as would be expected after a 15 min incubation period when there should be virtually no difference in radioactive labelling between the N- and C-terminal ends of normal α chains (Clegg *et al.* 1968). Also the specific activities of αCS chain peptides 4, 10, 14, X and Y were also similar indicating that there was no gradient of radioactivity between the N- and C-terminal ends of the αCS chain. These results show

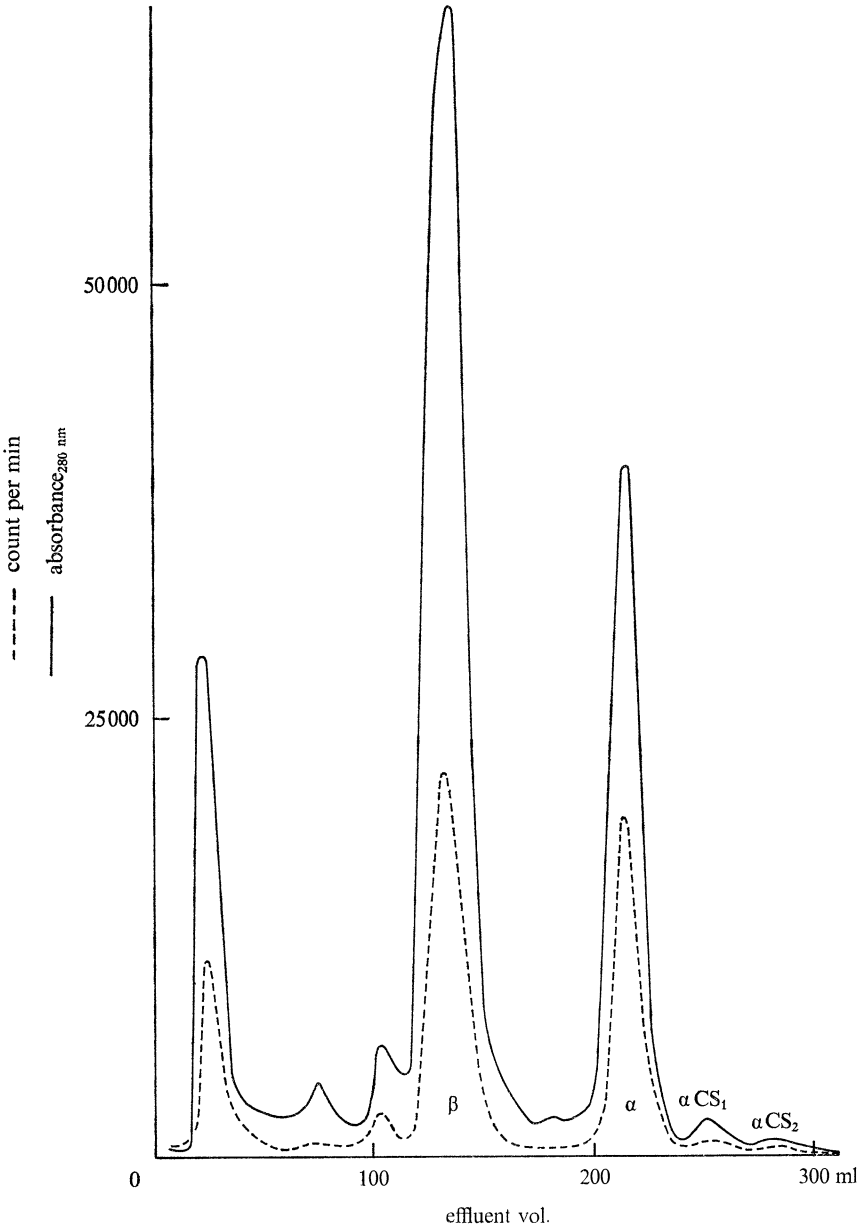
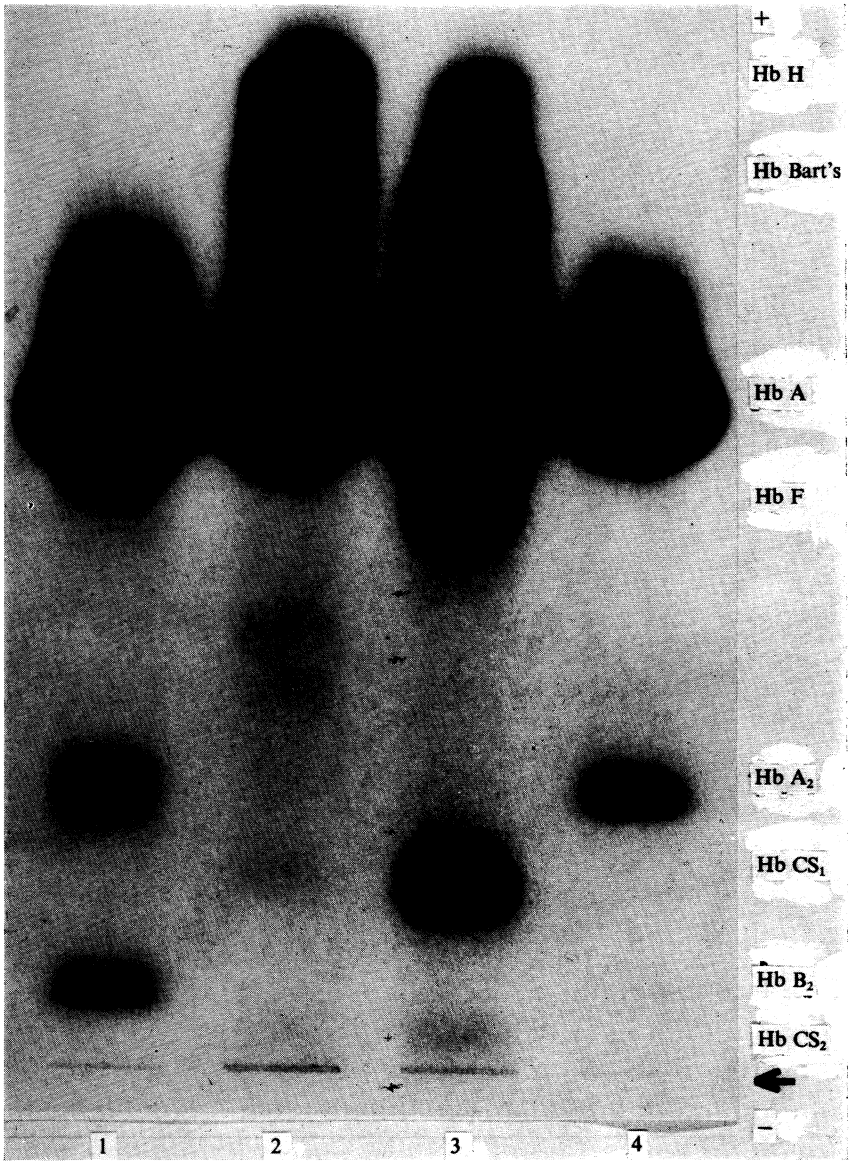


FIGURE 14. Haemoglobin synthesis in the bone marrow of an individual heterozygous for α-thal. 1 and Hb CS. Whole-cell globin, prepared after incubating bone marrow as described in the text, was fractionated on CM-cellulose with 8 M urea–mercaptoethanol–phosphate buffer, pH 6.8. Solid line, radioactivity; dotted line, absorbance at 280 nm. The αCS-chain peak contained 5–6% of the total α-chain radioactivity; this is probably an overestimate due to contamination with the highly radioactive αA chain.



RE 15. The electrophoretic pattern of Hb-CS-like fractions from different racial groups. Left to right: 1, a lysate from an individual heterozygous for Hb B_2, a δ-chain mutant, shown as a control; 2, a lysate from an individual heterozygous for α-thal. 1 and Hb CS from Thailand; 3, a lysate from an individual heterozygous for Hb CS and α-thal. 1 from Malaysia; 4, a normal adult lysate as a control. Vertical descending starch-gel electrophoresis, tris–EDTA–borate system, pH 8.5, benzidine stain.

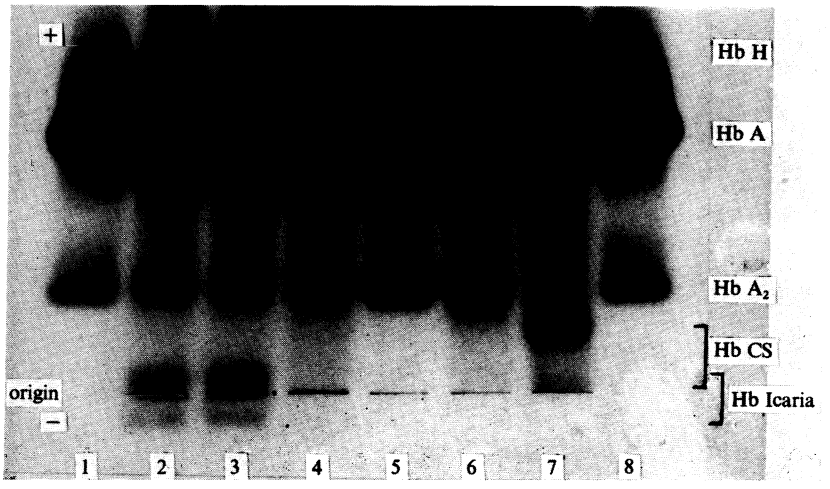


FIGURE 16. The electrophoretic characterization of Hb Icaria. The lysates are from the following sources (left to right): 1, normal adult; 2 and 3, Hb-Icaria heterozygotes; 4, Hb-CS heterozygote; 5, normal adult; 6, Hb-CS heterozygote; 7, an individual heterozygous for Hb CS and α-thal. 1; 8, a normal adult. Vertical descending starch-gel electrophoresis, tris–EDTA–borate system, benzidine stain.

(*Facing p.* 432

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

that there is no gross discrepancy in the time in which the αCS chain takes to become uniformly labelled as compared with the αA chains and thus rule out a reduced rate of assembly or release as the basis for the reduced rate of production of αCS chains. If a reduced rate of assembly were responsible for a 20-fold reduction in the rate of αCS-chain production, compared with that of the αA chain, it seems unlikely that it would have been uniformly labelled after 15 min since it takes 9 min at 37 °C to uniformly label αA chains (Clegg *et al.* 1968). It was noticed also that the specific activity of the αCS peptides was approximately 1/6th of that of the αA peptides, further confirmation that αCS synthesis declines more rapidly than that of αA during erythroid maturation.

TABLE 9. THE ASSEMBLY OF THE αA COMPARED WITH THE αCS CHAIN

Reticulocytes from an individual heterozygous for Hb CS and α-thalassaemia 1 were incubated for 15 min with [^{3}H]arginine. The cells were immediately washed at 4 °C and frozen. After thawing, Hbs, A, H and CS were purified on Amberlite-IRC 50 columns as described in the text. The α and β chains were then separated and fingerprints made of the αA and αCS chains. After staining with 0.02% ninhydrin the peptides indicated were eluted in 4 ml of 10^{-3} M HCl. Samples were then taken for amino-acid analysis and counting.

peptide	molar activity [^{3}H]Arg count per min / μmol [^{12}C]Arg
αA 4	1810
αA 10	2020
αA 14	1880
αCS_1 4	304
αCS_1 10	366
αCS_1 14	360
αCS_1 X	283
αCS_1 Y	279

(*e*) *The slow-moving haemoglobins of Hb H disease from other racial groups*

Following the publication of a preliminary account of the genetics and structure of Hb CS (Milner *et al.* 1971) several workers told us that they had observed slowly-migrating haemoglobin fractions in patients with Hb H disease. Samples of these were obtained from the following sources; two from Thailand, one from Hong Kong, two from Greece, one from Malaysia and a further sample from Malaysia from the child apparently homozygous for Hb CS mentioned already.

Electrophoretic studies. All samples were examined by starch-gel electrophoresis with the tris–EDTA–borate system, pH 8.5. In each case except one from Greece the slowly-moving components migrated in the position of Hbs CS_1 and CS_2, although their relative quantities varied depending on the age and condition of the samples (figure 15, plate 24).

Structural studies. Sufficient red cells were obtained to prepare approximately 10–20 mg of αCS chains. Again the relative yields of αCS_1, αCS_2 and αCS_3 varied but they were always present, and eluted from the CM-cellulose column in approximately the same proportions as would the corresponding genuine Hb CS fractions. The chains were isolated and fingerprinted as described for Hb CS. In each case but one the structure was identical to that of Hb CS. These identifications were based on the isolation of purified α-chain from the abnormal Hb, followed by peptide mapping of trypsin digests of the α chain and analysis of the critical peptides X, Y and W1. Particular attention was paid to peptide X, since in view of current ideas about the origin of Hb CS (see later) it was conceivable that the N-terminal residue

might be different from glutamine in some cases. But for the one exception described below this proved not to be the case.

An account of the isolation and characterization of the one anomolous haemoglobin, now called Hb Icaria, follows.

(*f*) *The isolation and characterization of Hb Icaria*

A blood sample was sent to us by Professor Tsevrenis of Athens from a 52-year-old Greek woman who had a mild thalassaemia-like disorder associated with the presence of trace amounts of a slowly-migrating haemoglobin variant on electrophoretic analysis. It was thought originally that this was a further example of Hb CS but subsequent electrophoretic and structural analysis showed this not to be the case and it has been named Hb Icaria after the name of the Aegian Island where the family live.

Electrophoretic analysis of Hb Icaria. A haemolysate containing the variant was examined by starch-gel electrophoresis and two slowly-migrating haemoglobin components were observed. However when compared with samples of genuine Hb CS it was found that they migrated more cathodally at pH 8.5 (figure 16, plate 24). The two fractions migrated in the general position of Hbs CS_2 and CS_4 and thus appeared to differ by about one charge from the Hb CS_1 and CS_2 components. This finding was observed repeatedly on fresh material and did not appear to be related in any way to the alteration in the mobility of the genuine Hb CS samples which occurs on ageing.

On quantitative cellulose-acetate electrophoresis the two slowly-migrating components did not separate and were eluted as a single band giving values of 0.65 and 0.90 % of the total haemoglobin in duplicate experiments. In addition, Hb A_2 was present as 2.5 % of the total haemoglobin and there was an elevated level of Hb F of 3.9 %. On starch-gel electrophoresis, pH 7.0, trace amounts of a component which migrates in the position of Hb Bart's were observed and subsequent chromatographic studies confirmed the presence of approximately 0.5 % of this variant.

Isolation. The procedure for the isolation of Hb Icaria was exactly the same as that used for the isolation of Hb CS. During Amberlite chromatography it was noted that the material that was retained on the column at room temperature was always tightly bound to the top of the column as a dense band and never moved down the column as a more diffuse zone, as occurred often with Hb CS. This material was eluted with developer 2 containing 0.5 M NaCl, dialysed, concentrated *in vacuo* and examined by starch-gel electrophoresis. The presence of two major bands migrating just before and just after the origin on alkaline starch-gel electrophoresis was confirmed. These fractions were converted to globin and the chains separated on CM-cellulose chromatography in the 8 M urea mercaptoethanol phosphate buffer, pH 6.8. A chromatogram is illustrated in figure 17. The abnormal α-chain components eluted from the CM-cellulose column after the position of the Hb CS α chains, further evidence that this variant is different from Hb CS. The chains were dialysed and freeze-dried for further analysis.

Structural analysis. Peptide maps of α-Icaria were similar but not identical with those of αCS (figure 18). Although peptides Y and W1 were present and had their expected amino-acid compositions, peptide X had a slightly different mobility from αCS X and it lacked the characteristic N-terminal glutamine residue. Further studies revealed that in the α-Icaria chain the glutamine 142-residue had been replaced by lysine (Clegg *et al.* 1974).

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

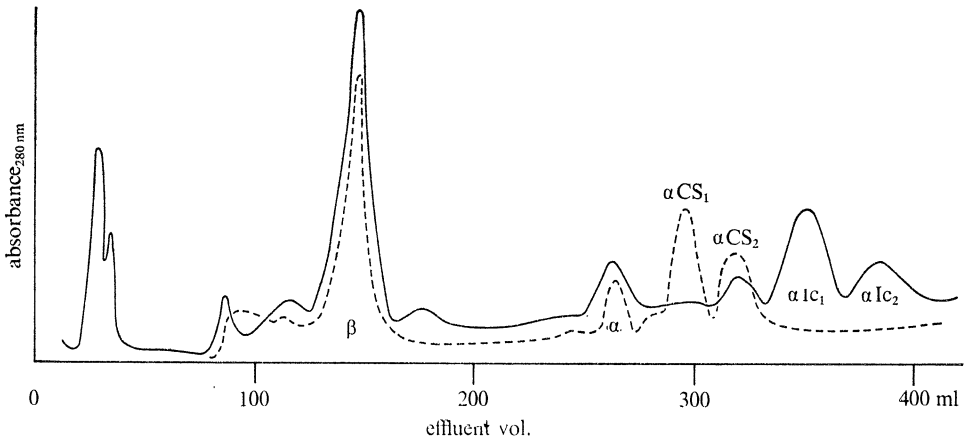


FIGURE 17. The separation of the globin chains of Hb Icaria by CM-cellulose chromatography. A sample of Hb CS globin was run as a control.

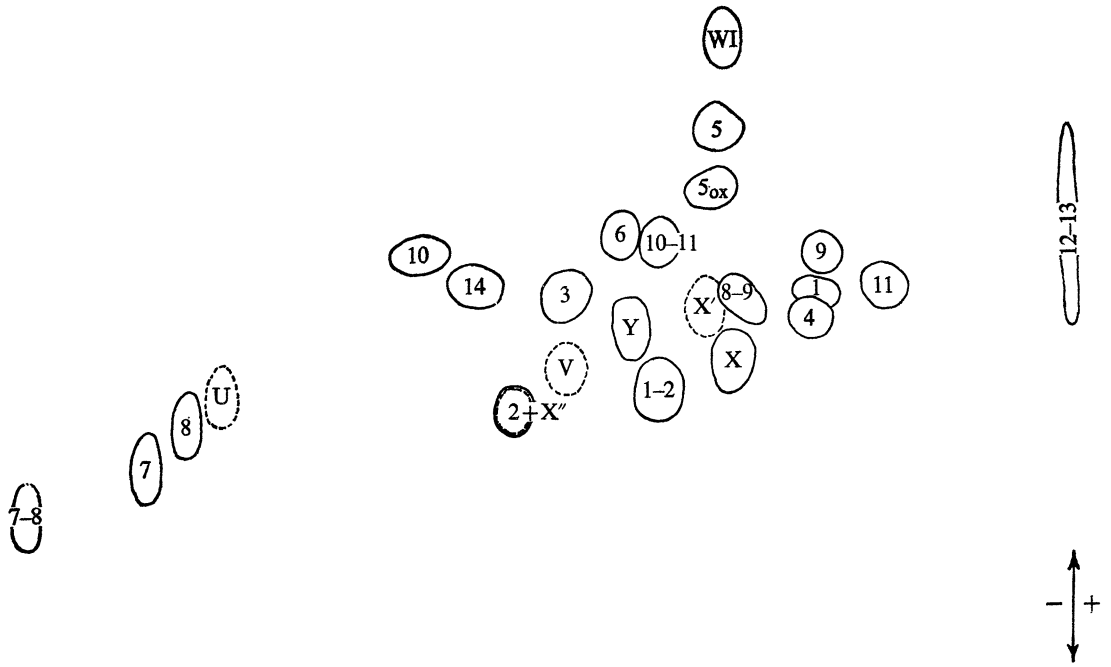


FIGURE 18. Superimposed tracings of fingerprints of αCS and α-Icaria chains. Peptides common to both α chains are drawn in solid lines, those unique to α-Icaria are dotted.

(*g*) *The inheritance of haemoglobins CS and Icaria*

In the families studied here Hb CS appears to be inherited in a simple Mendelian co-dominant fashion. Individuals heterozygous for Hb CS carry between 0.5 and 1 % of the variant and genetic data derived from the Malaysian family (figure 4 and table 10) indicate that the homozygous state is associated with the presence of 5–6 % Hb CS, the remainder being made up by Hbs A and A_2 with a trace of Hb Bart's.

The data presented for family C (figure 1 and table 1) indicate that when Hb CS is inherited

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

TABLE 10. THE HAEMATOLOGICAL AND ELECTROPHORETIC FINDINGS IN THE PATIENT HOMOZYGOUS FOR Hb CS, AND IN HIS FAMILY MEMBERS

The haematological and electrophoretic data was collected by Dr Lie-Injo Luan Eng except for the quantitation of the αCS components in the propositus which was done in Liverpool.

subject	age year	Hb g/dl	10^6 r.b.cs per μl	p.c.v. (%)	m.c.v. fl	m.c.h. pg	m.c.h.c. (%)	retics (%)	Hb pattern	serum iron μg/dl	Hb CS (%)	Hb A_2 (%)	Hb F (%)	Hb Bart's (%)
propositus	12	11.3	4.82	36.5	75.7	23.4	31.0	7.6	$A+A_2+CS$	71	6.7 CS_1 4.5 CS_2 2.2	2.2	1.3	not detectable
after iron therapy		10.3	4.89	34.5	70.6	21.1	29.9	5.7	$A+A_2+CS$	89	5.0	2.0	2.0	trace
father	32	13.4	5.52	41.5	75.2	24.3	32.3	1.6	$A+A_2+CS$	165	0.5	3.1	1.4	not detectable
mother	29	12.4	4.58	40.0	87.3	27.0	31.0	1.2	$A+A_2+CS$	69	1.2	3.0	1.5	not detectable
sister 1	10	12.7	5.16	41.0	79.5	24.6	30.1	0.8	$A+A_2+CS$	110	1.0	3.1	1.0	not detectable
sister 2	6	12.2	5.16	39.5	76.6	23.6	30.9	2.2	$A+A_2+CS$	133	0.6	2.5	1.2	not detectable
brother 1	13	13.4	5.21	42.0	80.6	25.7	31.9	1.8	$A+A_2$	—	—	3.0	—	not detectable
brother 2	7	12.2	4.77	40.5	84.9	25.6	30.1	0.6	$A+A_2+CS$	—	0.5	2.8	1.4	not detectable
brother 3	4	10.0	5.81	35.5	61.1	17.2	28.2	0.8	$A+A_2+CS$	30	0.6	2.8	1.2	not detectable
brother 4	2	12.9	6.26	42.0	67.1	20.6	30.7	2.4	$A+A_2+CS$	59	1.1	2.4	1.8	not detectable
brother 5	(cord blood)	—	—	—	—	—	—	—	$A+A_2+CS+$Bart's	—	1.3	—	—	2.4
	8 days	16.8	5.38	50.0	92.9	31.2	33.6	1.8	$A+A_2+CS+$Bart's	—	0.8	—	58.3	2.5

TABLE 11. THE HAEMATOLOGICAL FINDINGS AND ELECTROPHORETIC DATA FROM THE FAMILY WITH Hb ICARIA

The haematological studies were carried out by Professor Tsevrenis in Athens and the electrophoretic data obtained both in Liverpool and Athens; both sets of data are included. The propositus is number II.2. Structural studies were performed on Hb Icaria isolated from the blood of individual II.1. +, Present; −, absent.

								Liverpool data			Athens data		
	Hb g/dl	p.c.v. (%)	10^6 r.b.cs per μl	m.c.h. pg	m.c.v. fl	abnormal morphology	serum iron μg/dl	Hb Icaria (%)	Hb F (%)	Hb A_2 (%)	Hb Icaria (%)	Hb F (%)	Hb A_2 (%)
I.1	13	41	4.5	28.8	91	+	67	Absent	3.88	2.49	−	< 0.9	2.5
II.1	12.5	39.5	4.6	27.1	86	++	90 130	0.89 0.65	3.88 trace Bart's	2.88 2.70	+	1.6 + Bart's	2.5
II.2	11	36.5	5.5	20.0	66	+++	84	+	1.6	2.9	+	1.6 + Bart's	2.9
II.3	13	41	4.3	30	95	++	93	+	2.2	2.5	+	2.2 + Bart's	2.5

together with the α-thal. 1 gene a condition identical to Hb H disease occurs; two siblings and a maternal cousin appeared to have inherited this combination of genes (family C, figure 1). Our family data does not include the critical matings required to see whether the Hb CS gene segregates dependently or independently of the α-thal. 1 gene. The interaction between the α-thal. 1 and Hb CS genes has been confirmed subsequently in several populations and will be discussed at greater length later together with data on the gene frequencies of Hb CS.

The family data on the patient with Hb Icaria is incomplete (table 11). The variant has been found in both sisters of the propositus but not in their 75 year old mother; the father is dead.

(*h*) *The clinical and haematological findings associated with Hbs CS and Icaria*

In the present studies Hb CS has been observed in the heterozygous state, homozygous state and doubly heterozygous state with the α-thal. 1 gene. Hb Icaria has been found only in the presumed heterozygous state.

Heterozygous state for Hb CS. The haematological findings in a series of Hb CS heterozygotes are summarized in table 1. There is no anaemia, the red-cell indices are normal and stained peripheral-blood films show no specific abnormalities. It appears that the heterozygous state for Hb CS is associated with no haematological abnormalities although to be absolutely certain it would be necessary to compare these parameters in a large number of affected persons with a similar group of normal persons to see if there were any statistical differences between the various red-cell indices; this seems unlikely however and this condition behaves very much like the α-thal. 2 carrier state.

Homozygous state for Hb CS. This has been only observed on one occasion so far and therefore it is difficult to be certain about the clinical features. The affected child was a 12-year-old Malay boy who had no symptoms and was ascertained during a population study because a young brother was found to have Hbs CS and Bart's at birth. The child was normally developed but had a palpable spleen. He was studied twice because he was thought to have mild iron-deficiency anaemia when first examined (see Lie-Injo *et al.* 1974).

The haematological findings (summarized in table 10) showed a normal haemoglobin level for the child's age but the red cells were relatively small and poorly haemoglobinized as shown by the low m.c.v. and m.c.h. values respectively. There was a slight persistent reticulocytosis and the stained peripheral-blood smear showed variation in the shape and size of the red cells with a few target and irregularly contracted forms; Hb-H bodies could not be demonstrated. The cells showed slightly increased osmotic resistance, and the absence of haptoglobins, as demonstrated by starch-gel electrophoresis, provided further evidence of a very mild haemolytic process. This childs' red-cell enzymes were estimated by Dr Lie-Injo Luan Eng and found to be normal, suggesting that there was no other genetic abnormality to account for the low-grade haemolysis. It appears then that the homozygous state for Hb CS is associated with a very mild thalassaemia-like disorder, certainly much milder than Hb H disease, and similar to the heterozygous state for the α-thal. 1 gene.

The heterozygous state for Hb CS and α-thal. 1 (*Hb H disease with Hb CS*). The haematological findings in three siblings with this condition are summarized in table 1. The clinical picture is indistinguishable from that of Hb H disease. These patients have a lifelong history of anaemia, haemolytic or aplastic crises, progressive enlargement of the spleen and an increased propensity to infection. A β-thalassaemia gene and one for a bizzare form of skeletal deformity also were present in this remarkable family but did not appear to be related to the Hb CS gene.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

The propositus (individual III.11, figure 1), was a 19-year-old Chinese girl who presented at the age of 5 to the University Hospital, Kingston, Jamaica, with jaundice, anaemia and a grossly enlarged spleen. Over the next three years she required repeated hospital admissions and her haemoglobin level fell as low as 4 g/dl on one occasion. Splenectomy was performed at the age of 8 years after which time she required no further blood transfusion and was able to lead a normal life. The younger sister of the propositus (III.19) presented at the age of 7 years with jaundice, anaemia and an enlarged spleen and her clinical course has been very much like that of the propositus. A similar clinical picture was also observed in a cousin of the propositus, III.8.

Thus all these patients have typical Hb H disease of the fairly severe type; whether Hb H disease with Hb CS is more severe than Hb H disease alone is not possible to say although in a small series Todd (1971) noted that the presence of Hb CS seems to be associated with a greater reduction of haemoglobin levels than Hb H disease without Hb CS.

5. Discussion

Although at first it seemed as if the minor haemoglobin variants associated with Hb H disease represented simply another type of human haemoglobin mutation it is now clear that they have much broader implications for our understanding of human haemoglobin genetics and in particular of the molecular basis for the α-thalassaemia disorders. Indeed it is evident that, when taken together with recent discoveries about the molecular defect in the Hb Bart's hydrops syndrome, it is now possible to build up a relatively complete picture of the genetics of the α-thalassaemias.

Since the experiments described in the previous sections cover a fairly broad area encompassing protein structure and synthesis, and haematological and genetic data, it will be convenient to discuss each of these topics separately and then finally to summarize what is now known about the molecular pathology of the α-thalassaemias.

(*a*) *The structure of the elongated α-chain variants*

Peptide mapping of tryptic digests of αCS chains established conclusively that Hb CS is an α-chain variant which has an additional 31 amino-acid residues over and above those that constitute the 141 of the normal αA chain. Fragmentation of the αCS_1 chain with CNBr showed that these 31 additional amino acids are attached to the C-terminal end of the αA-chain sequence.

Three major peptides unique to the αCS chain are produced on digestion with trypsin. Of these the sequences of X and Y were established without difficulty. Edman degradation of W1 proceeded satisfactorily as far as the 4th residue; beyond this point the results were ambiguous and the identity of the next residue could not be established conclusively. However, chymotrypsin degradation of W1 did produce a fragment of composition and partial sequence Ala.Leu.Leu.Pro.(Ser, Leu, His) and carboxypeptidase digestion of this peptide released leucine and histidine in good yields as the only products. (No prediction of the C-terminal residue based on the known specificity of chymotrypsin seemed justifiable here. Although leucine is a common substrate, splits at histidine can be obtained in high yield, as for example when α12b.13 is digested.) Only on one occasion was a tentative identification of the 5th residue of W1 made (as serine). However, repeated degradations identified Leu as the 6th

residue. It is not clear why DNS-Ser could not be identified, although the fact that DNS-Leu was subsequently detected suggests that complete cyclization of the serine occurred normally.

While this sequence work was in progress, Efremov *et al.* (1971) reported some preliminary sequence studies on an α-chain variant which was obviously the same as Hb CS. They sequenced two peptides, one of which had an identical sequence to Y, and the other of 8 residues which had an identical sequence to that proposed here for the first eight residues of W1 or the complete sequence of peptide Z. These workers gave no indication of the relative yields of their peptide (Z). Possibly under their conditions of digestion greater yields were obtained than here, although it is quite unusual to get substantial cleavage of Arg-Pro bonds under any circumstances (Milstein, Clegg & Jarvis 1968). There seems little doubt however from the total amino-acid analysis of αCS_1 and the identity of the sequences of Z and the first eight residues of W1 that Z is a partial digestion product of W1. The yields were always very low and it was not found in some digests at all.

The ordering of W1, X and Y was also relatively straightforward. The two critical peptides αCSchyA and αCSchyB lead unequivocally to a sequence X–Y–W1. The loss of the free arginine normally obtained in chymotrypsin digests of αA and the appearance of a chymotrypsin peptide with a sequence Arg.X.Trp. confirmed that X is attached to the C-terminal Arg 141 residue of the normal αA chain.

Structures of the αCS_2 *and* αCS_3 *chains.* All the available evidence indicates that the αCS_2 and αCS_3 chains are shortened versions of αCS_1. The proportions, particularly of αCS_3, varied considerably from haemolysate to haemolysate, and even during the purification of individual samples. The most likely explanation is proteolytic degradation by red-cell enzymes during storage and manipulation. Since the C-terminal residues of the CS_2 and CS_3 chains are leucine and tryptophan respectively it seems likely that some sort of chymotrypsin-like activity is responsible. The fact that the C-terminal residue of the αCS_1 chain is glutamic acid would also suggest that it has not been produced by a similar proteolytic process from a larger unidentified precursor. In any case we have never had any evidence for a minor αCS component which was larger than αCS_1. On many occasions other minor Hb-CS components have been observed on starch-gel electrophoresis, the most prominent one being Hb CS_4. So far we have been unable to isolate this for structural studies so that its identity remains unknown. Its electrophoretic mobility suggests that it should have an α chain which is more basic than αCS_2, but it is difficult to see how such a derivative could arise. It is possible of course that these very basic components represent haemoglobins with the molecular formulae $\alpha_2CS\gamma_2$ or $\alpha_2CS\delta_2$ since such abnormal fetal or minor adult haemoglobins are known to occur in individuals heterozygous for α-chain variants (Weatherall & Boyer 1962).

Structures of the other Hb CS-like haemoglobins. All the minor slow-moving haemoglobins with electrophoretic properties identical with Hb CS that we have examined so far have had structures identical to Hb CS. Hb Icaria, a variant in which lysine replaces glutamine at position 142, behaves quite differently from Hb CS on electrophoresis and chromatography.

However, as will become evident from the discussion of how Hb CS has arisen, other Hb CS-like variants may exist in which residue 142 is not glutamine but which have identical electrophoretic properties to Hb CS. Indeed one such haemoglobin has been reported recently by De Jong, Kahn & Bernini (1975). This variant, called Hb Koya Dora, has been found in a relatively high frequency in Andhra Pradesh. Although it has similar properties to Hb CS and has an elongated α chain, position 142 is occupied by serine rather than glutamine.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Perusal of the genetic code (figure 19) suggests that two more Hb CS-like variants can theoretically exist (142 Leu and 142 Tyr). A third possibility (142 Glu) would have different electrophoretic properties from Hbs CS and Icaria (see below).

It cannot therefore be assumed that Hb CS-like variants are in fact Hb CS, unless the identity of the residue at position 142 can be established. This is most easily done by determining the N-terminal residue and amino-acid composition of peptide X, a task which can be achieved satisfactorily enough if 2 mg or more of purified αX chains are available. Using this approach we have confirmed that the Hb-CS-like variants found in Thailand, Hong Kong, Greece (in one family) and Malaysia are identical to genuine Hb CS.

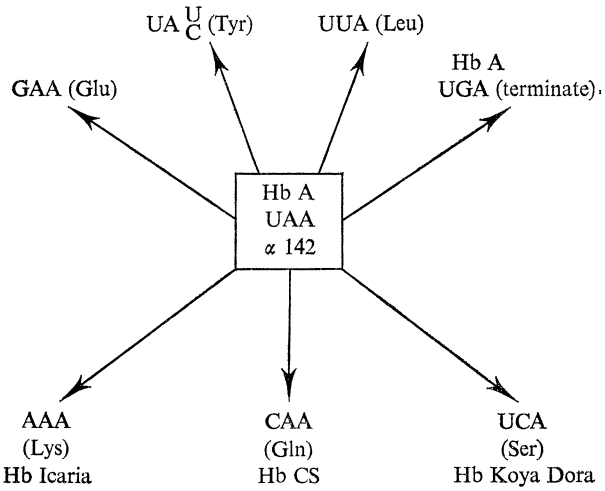


FIGURE 19. The potential and actual α-chain-termination mutant haemoglobins which could be produced by a single base change in the chain-terminating codon UAA.

(*b*) *The molecular basis for the formation of Hb CS*

In our original description of the structure of Hb CS we considered in a general way what sort of genetic processes could result in a variant with such unique properties, and proposed four alternative mechanisms which went some way towards explaining the data. These were that (*a*) Hb CS was a chain-termination mutant in which the normal chain-terminating codon (UAA or UAG) had mutated to a glutamine codon (CAA or CAG) by a single base substitution, (*b*) Hb CS was produced by a supressor mutation. In this case it was assumed that a minor glutamine transfer RNA had undergone mutation in its anti-codon so that it now inserted glutamine in response to a UAA or UAG codon, (*c*) Hb CS was the expression of a minor α-chain locus which is usually repressed in normal individuals and (*d*) Hb CS was the result of an abnormal crossover so that extra genetic material had been attached to a normal α-chain gene.

On the basis of the data available at the time (*a*) seemed the most likely explanation although (*b*) could not be conclusively ruled out. (*c*) and (*d*) were not considered to be strong possibilities and in the light of more recent evidence discussed below can now be ruled out entirely.

The ideas that Hb CS resulted from the expression of a minor α-chain locus or an abnormal cross-over gene both implied that the αCS chains were derived from gene loci not usually expressed or found in humans. Recent work on the structure of human α-chain mRNA has, however, shown that the normal α-chain gene contains all the information required for

the expression of an Hb-CS-like α chain. There is therefore no need to invoke the presence of hypothetical or abnormal genes to explain the formation of what is potentially the mutant product of a normal α-chain gene.

Suppressor mutations have never been described in mammals. In many respects though the Hb-CS mutation could be explained by a suppressor mutation of the sort found in bacteria. Most bacterial suppressors, particularly those for the 'ochre' codon UAA which is considered to be a common natural terminator in microorganisms, operate at low efficiency, typically of the order of 5–15 % (Stretton 1965). Many suppressors are known to be mutants of minor tRNA species, and it is conceivable that a mutant minor glutamine transfer RNA, assuming one exists in erythroblasts, could be acting as an 'ochre' suppressor to produce Hb CS. The evidence, albeit circumstantial, which makes this explanation difficult to accept is the 'interaction' between the α-thal. 1 and Hb CS genes. Why is a single Hb CS gene expressed at three or four times the rate when an α-thal. 1 gene is present in the same cell? Put another way, why should the efficiency of suppression depend on the amount of α mRNA present in the cell? Furthermore, it is difficult to see why suppression should also lead to a deficit in total α-chain production of the order of 15 % (figure 13) in the cell, i.e. produce α-thalassaemia.

Neither of these objections is really conclusive proof against the suppressor hypothesis but make it seem unlikely. An experiment which would decide the issue, however, would be to see whether αCS chains are synthesized in a heterologous cell-free system using globin mRNA isolated from bone marrow of persons with Hb CS. A positive answer would disprove the suppressor hypothesis.

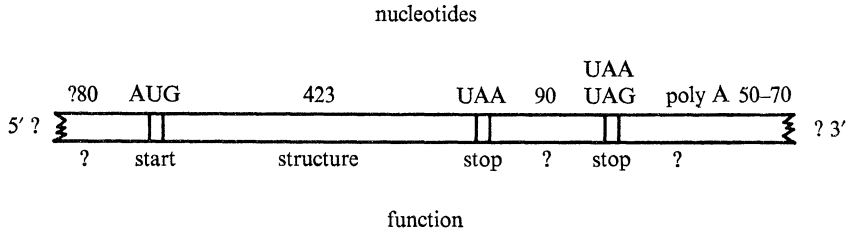


FIGURE 20. A model for human α-chain mRNA based on evidence which is described in detail in the text.

The most plausible explanation for the origin of Hb CS, then, is that it has occurred as a result of a mutation in the normal chain-terminating codon of α-chain mRNA, allowing read-through into extra-structural codons until the next in-phase termination codon is reached. It is known that globin mRNA has a molecular mass of about 220 000 and contains approximately 650 nucleotides (Blobel 1971; Gaskill & Kabat 1971). Since only 423 nucleotides are needed to code for the α-chain sequence it is evident that α mRNA also contains approximately 230 non-coding nucleotides. Of these approximately 70 comprise the poly-A region at the 3′-end (Lim & Canellakis 1970; Burr & Lingrel 1971), leaving another 160 to be accounted for. Approximately 90 of these must code for the 31-residue segment of αCS chain and the remaining 70 are presumably to be found at the 5′-end and at the 3′-end between the 'CS-region' and the poly A (figure 20).

During the last two years considerable advances have been made in elucidating the structure of mRNA. Forget and his colleagues have made substantial progress in the determination of

the nucleotide sequence of human α mRNA, and have identified a number of fragments with sequences coding for known parts of the αA chain, and also portions of the 31-residue segment unique to αCS (Marotta *et al.* 1974). The results establish conclusively that normal α mRNA contains nucleotide sequences coding for the amino-acid sequence unique to αCS. The discovery of the frame-shift mutant Hb Wayne has also augmented these studies. This variant has an abnormal α chain of 146 residues, which differs in sequence from αA from residue 139 onwards (Seid-Akavan, Winter & Rucknagel 1972). By comparing the amino-acid sequences of Hbs Wayne and CS a 26-residue nucleotide sequence can be deduced for α mRNA for codons 139–146, and Marotta *et al.* (1974) have extended this sequence for another 10 nucleotides to the left of codon 139. The chain-terminating codon at position 142 in this sequence is UAA (figure 21), which can mutate to a glutamine codon (CAA) by a single base substitution.

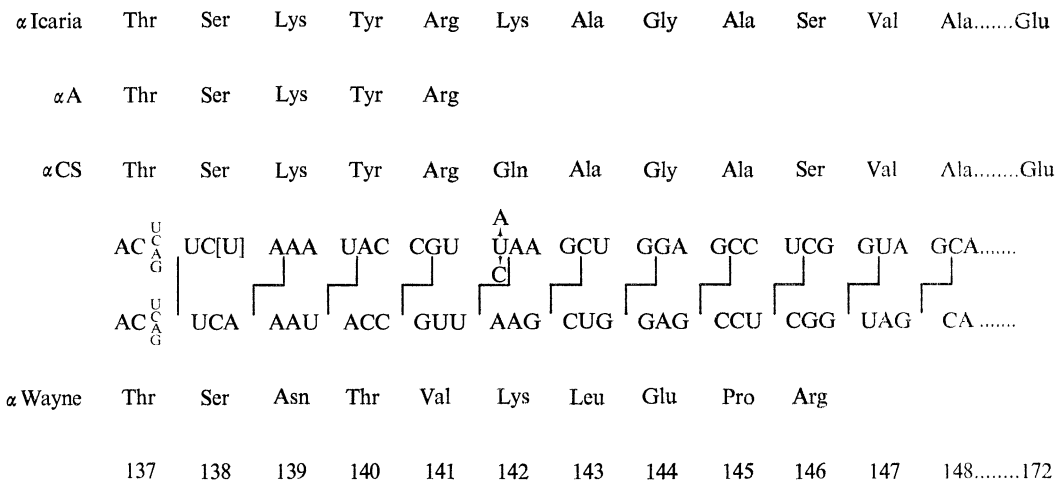


FIGURE 21. Partial amino-acid and nucleotide sequences for the α chains of Hbs A, CS, Icaria and Wayne. The base sequences for codons 138 and 139 have been chosen arbitrarily. Deletion of any one of the four bases before codon 140 produces the nucleotide sequence which defines the amino-acid sequence at the C-terminal end of the α chain of Hb Wayne. In this diagram the third base of the codon for Ser 138 has been deleted; the base sequence of the α mRNA is thus uniquely defined from codon 140 to the second base of codon 148.

If it is accepted that Hb CS is a chain-termination variant it follows that, theoretically at least, there may be a family of variants produced by different mutations of the UAA codon at position 142, but which have identical sequences for residues 143–172. Two of these, Hb Icaria (142 Lys) and Hb Koya Dora (142 Ser) have been reported. Of the three remaining possibilities (figure 19) two (142 Tyr and 142 Leu) would be expected to have electrophoretic properties similar to Hb CS, while the third (142 Glu) would be a more acidic variant with a mobility similar to that of Hb A_2.

(*c*) *Biosynthesis*

The chain-termination hypothesis does not by itself offer a convincing explanation for the low rates of synthesis of Hb CS, which is found only in very low quantities in the red cells of both heterozygotes and homozygotes. This might be so either because it is being synthesized more inefficiently than normal or because its rate of destruction and turnover is unusually rapid. The results of our experiments provide clear evidence that Hb CS is not unstable and that it does not turn over more rapidly than Hb A. In experiments in which the red cells

were incubated with radioactive isotopes and whole-cell globin was fractionated by urea chromatography the specific activity of the αCS chain never exceeded that of the αA chain at incubation times ranging from 5 min to 2 h. Likewise the specific activities of the individual peptides of the αCS chain were considerably lower than those of the corresponding ones for the αA chain. These observations show that Hb CS is being synthesized in reticulocytes at a rate considerably less than is commensurate with its proportions in peripheral blood; this must mean that its rate of synthesis declines during erythroid maturation.† The overall rate of α-chain synthesis in Hb-CS carriers is slightly but significantly less than that of β-chain synthesis, with an average α/β-production ratio of 0.85. This is in the same range as that observed for α-thal. 2 heterozygotes (figure 13).

Why then is the elongated α chain of Hb CS produced less efficiently than the α chain of Hb A? A reduced rate of globin-chain production could occur if there is a reduced amount of α mRNA transcribed, if the mRNA is unstable, or if there is inefficient translation (initiation, elongation or termination) of normal amounts of α-chain mRNA. We have measured the assembly time for the αCS chain and it does not appear to be slower than normal; certainly the rate of translation is not sufficiently retarded to account for the very low level of Hb CS found in heterozygotes. These studies also rule out a marked delay in chain termination as might occur if the normal release mechanism was not operative for the elongated α chain and if it was simply released from the ribosomes by an unusual process, for example by proteolytic cleavage. Since the N-terminal end of the αCS chain is the same as the αA chain there is no reason to believe that the process of initiation is abnormal although this is theoretically possible. However, the fact that the synthesis of Hb CS declines more rapidly than that of Hb A during erythroid maturation, and the finding of a normal translation time, points to a reduced stability of mRNA for the αCS chain as a more likely mechanism for its low level in the erythrocyte.

The pattern of synthesis of the αCS chain during erythroid maturation bears a strong resemblance to that of the δ-chain of the normal minor adult haemoglobin, Hb A_2. This is synthesized only in nucleated red-cell precursors and its production has virtually ceased by the reticulocyte stage of maturation (Rieder & Weatherall 1965; Roberts, Weatherall & Clegg 1972; Lang, White & Lehmann 1972). There are a number of haemoglobin variants called the Lepore haemoglobins which have non-α chains which consist of the N-terminal ends of the δ chains fused to the C-terminal ends of the β chains. It is thought that these variants have arisen by unequal crossing over at meiosis following chromosomal misalignment between the δ- and β-chain loci with the production of δβ-fusion genes (Baglioni 1962). Studies on the synthesis of the Lepore (δβ) haemoglobins and their anti-Lepore (βδ) counterparts have shown that these are also produced only in nucleated red-cell precursors and that their synthesis declines during erythroid maturation (Roberts *et al.* 1972; Lang *et al.* 1972). Furthermore this is not a property inherent in fusion haemoglobins since Hb Kenya, a γβ-fusion variant, is actively synthesized in reticulocytes (Clegg & Weatherall 1974*a*). Rather it seems that the δ chain, and all the variants which have δ-chain residues at one or other end, have an unstable mRNA. Clegg & Weatherall (1974*a*) have recently suggested that the instability of δ-chain mRNA is due to lack of stabilizing regions at the 5′- and 3′-ends of the δ mRNA, which make it more susceptible to nuclease attack, and hence ensure the rapid decline of

† Identical conclusions have been reached recently by Kan and his colleagues (Kan *et al.* 1974).

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

δ-mRNA, and thus δ-chain, synthesis after RNA synthesis stops at the intermediate normoblast stage of red-cell development.

There is no reason to suppose that αCS mRNA has a basically different structure from normal α mRNA and it is unlikely therefore that its instability is due to the same cause as that which results in the degradation of δ-chain mRNA. However, it is interesting that Hb Wayne, like Hb CS, is also produced in very low amounts and that both these variants have in common the property that they arise through the translation of regions of α mRNA which normally remain untranslated. Furthermore, White, Laux & Dennis (1972) have shown that there is an extensive base-paired loop around the chain-termination codon in α mRNA. Possibly this part of the mRNA is normally involved in maintaining its stability, perhaps by preventing degradation by nucleases which attack single-stranded RNA, and that if this region is translated by ribosomes the stabilizing areas are broken down and exposed to nuclease attack.

(*d*) *The genetic significance of the chain-termination mutants; the number of* α*-chain loci in Man*

It has been suspected for some time that the α-chain locus may have duplicated in at least some human populations. The first suggestion that this might be the case came from the observation that in heterozygotes for structural haemoglobin variants levels of α-chain variants are usually lower than those of β-chain variants. Thus it is possible that the α-chain variants are the gene products of only one of four loci whereas β-chain variants are the products of one of only two (Lehmann & Carrell 1968). Certainly there is evidence for multiple α-chain loci in a variety of animal species (reviewed by Clegg 1971) and there is no reason why this should not be so in humans.

Family studies have provided evidence both for and against the idea of multiple α-chain loci in Man. Abramson, Rucknagel, Schreffler & Saave (1970) described two individuals apparently homozygous for the α-chain variant, Hb J Tongariki, in whom no Hb A could be demonstrated. Unfortunately several critical members of this pedigree were not available and it is possible that these individuals had α-thalassaemia and Hb J, although on haematological grounds this seems unlikely. In another family of Hungarian origin two individuals were heterozygous for two α-chain variants called Hbs Buda and Pest. In addition to the abnormal haemoglobins both these persons had Hb A which could only occur if there are at least two α-chain genes on each chromosome (Brimhall *et al.* 1970). Although preliminary genetic data indicated that Hbs Buda and Pest segregate independently, subsequent studies (Hollan *et al.* 1972) have suggested that the validity of this result is clouded by non-paternity. The data obtained from these families have thus given conflicting results regarding the number of human α-chain loci.

However the genetic studies of families carrying the chain-termination mutants have provided a considerable amount of evidence on this point. The findings in the individual homozygous for Hb CS described in the present study show clearly that, assuming that Hb CS results from a mutation of the chain-terminating codon, there must be a minimum of two α-chain loci on each chromosome, at least in the Malaysian population. Similarly, since it is now apparent that the α-thal. 1 gene results from a deletion of all the α-chain loci (see below), the presence of both Hbs A and CS in individuals heterozygous for α-thal. 1 and Hb CS who have Hb H disease provides further evidence that there are two α-chain loci on the chromosome carrying the Hb-CS gene. Similar conclusions can be drawn from genetic studies of Indian families

who carry the chain-termination mutant Hb Koya Dora. Thus one family was found in which several members carried both Hb Koya Dora and another α-chain variant, Hb Rampa, and these individuals also had Hb A, indicating that there must be at least two α-chain loci per chromosome (De Jong *et al.* 1975).

How then are we to reconcile the discrepancy of the findings in Tongariki with those in other world populations?

It is possible of course that the elongated α-chain variants described in this paper result from suppressor mutations and if this were the case the argument outlined above becomes invalid. The findings in the family with Hbs Buda and Pest provide evidence that there are multiple α-chain loci and similar data has been obtained from one of the Hb-Koya Dora families. Perhaps the most likely explanation is that all this data is correct and that there are racial differences in the number of α-chain genes. If the α-chain locus has reduplicated recently there is no reason to suppose that all human populations have the same number of α loci. Indeed it is possible that some populations have even more than two per chromosome. Certainly a variability in the number of α loci could account for a lot of discrepancies in human haemoglobin genetics. For example it would explain some of the differences in α-thalassaemia which have been observed in the Negro and Saudi Arabia populations as compared with the disease in Southeast Asia (Pembrey, Weatherall, Clegg & Perrine 1975). It would also account for the marked variability in the levels of α-chain mutants which occur in heterozygotes in different populations. This problem has been discussed also in relation to the observation that Hb G Philadelphia is often present in relatively large amounts and with the clinical picture of thalassaemia in heterozygotes while Hb Q (α 74 Asp → His) is also associated with α-thalassaemia and the relatively large amount of Hb Q in the carrier state (French & Lehmann 1971; Lehmann & Lang 1974). Until recently the idea of differing numbers of α-chain genes has had to be entirely speculative but with the availability of recently-developed methods for DNA/DNA hybridization, and the production of highly specific cDNA probes for globin genes, it should soon be possible to provide quantitative answers to these questions.

(*e*) *The incidence of the chain-termination mutants in the world populations*

At the time of writing there is relatively little population-survey data relating to the α-chain termination-mutants. Most of the electrophoretic screening procedures which have been used in population surveys would have been too insensitive to pick up the heterozygous state for Hb CS and it may well have been overlooked even in surveys by starch-gel electrophoresis. We have only two sources of information at the moment – the number of cases of Hb H disease with Hb CS and some preliminary population surveys conducted by Lie-Injo and her colleagues in Malaysia and Pootrakul and his colleagues in Thailand (tables 12A, B).

Haemoglobin H disease is extremely common in Thailand, and Wasi, Na-Nakorn, Pootrakul & Panich (1972) have examined haemolysates from 194 affected families by horizontal starch-gel electrophoresis under conditions which would detect small amounts of Hb CS. They found that 85 of these families had Hb H disease associated with Hb CS. This observation suggests that approximately 45% of cases of Hb H disease in Thailand result from the interaction of Hb CS with α-thal. 1. Thus if we know the approximate incidence of the α-thal. 1 and α-thal. 2 genes as derived from estimations of the level of Hb Bart's in Thai infants, we should be able to obtain an approximate estimate of the incidence of Hb CS in Thailand.

Pootrakul, Wasi, Pornpatkul & Na-Nakorn (1970) examined the cord bloods of 1408 Thai newborns in Bangkok and found that 288 of them had increased levels of Hb Bart's, an incidence of 20.4 %. They found that the levels of Hb Bart's fell into three separate groups corresponding to 1–2, 5–6 and approximately 25 % Hb Bart's respectively. These are thought to represent α-thal. 2, α-thal. 1 and Hb H disease respectively. The incidence of the α-thal. 2 genotype in this survey was 10.2 % while those of the α-thal. 1 and Hb-H genotypes were 9.9 and 0.5 % respectively. Assuming that just less than half of the infants with the α-thal. 2 trait carried Hb CS this would give an overall population incidence in the Bangkok region of approximately 4 %. A lower incidence has been found in surveys currently being carried out in Thailand (Wasi, Na-Nakorn & Pootrakul 1974; Pootrakul *et al.* 1974). In a recent cord-blood survey (table 12B) Pootrakul *et al.* (1974) have found that 371 out of 2404 samples contained Hb Bart's, an overall incidence of 15.4 %. The incidence of Hb CS was 1.9 % (46 cases) and no samples containing Hb CS without Hb Bart's were found. In 40 cases, thought to be heterozygous for Hb CS, the levels of Hb Bart's ranged from 1–4.5 % (mean 2.8 %) while 4 had 8.8–11.4 % Hb Bart's and 2 had 32.8 and 29.9 % Hb Bart's. The latter two groups were thought to represent the Hb CS–α-thal. 2 and Hb CS–α-thal. 1 combinations respectively.

TABLE 12A. INCIDENCE OF Hb CS IN DIFFERENT POPULATION GROUPS

(1) *from analysis of Hb-H disease*	
(*a*) Thais: 85 out of 194 families	Wasi *et al.* (1972)
(*b*) Chinese (Hong Kong): 5 out of 43 cases	Todd (1971)
(*c*) Chinese (Singapore): 9 out of 40 cases	Wong (1974)
(*d*) Malays (Singapore): 2 out of 4 cases	Wong (1974)
(*e*) Chinese (Malaysia): 5 out of 8 cases	Lie-Injo *et al.* (1971)
(*f*) Malays (Malaysia): 7 out of 9 cases	Lie-Injo *et al.* (1971)
(2) *adult population surveys*	
(*a*) Malays (Malaysia): 2.24 % out of 536	Lie-Injo & Duraisamy (1972)
(*b*) Chinese (Malaysia): 0.66 % out of 607	
(*c*) Indians (Malaysia): 0.61 % out of 642	
(3) *cord blood surveys*	
(*a*) Malays (Malaysia): 6 out of 492	Lie-Injo (1973)
(*b*) Chinese (Malaysia): 1 out of 501	
(*c*) Indians (Malaysia): 0 out of 438	

TABLE 12B. THE INCIDENCE OF Hb BART'S IN 2404 THAI CORD-BLOOD SAMPLES (POOTRAKUL *ET AL.* 1974)

Hb Bart's range	Hb Bart's mean ± s.d.	genotypes	no.	(%)
with Hb Constant Spring			46	1.9
1.0–4.5	2.82 ± 0.78	Hb CS trait	40	1.7
8.8–11.4	10.0 ± 1.14	α-thal. 2–Hb CS	4	0.2
29.9, 32.6	31.2	α-thal. 1–Hb CS (Hb H disease with Hb CS)	2	0.1
without Hb Constant Spring			325	13.5
0.5–3	1.84 ± 0.59	α-thal. 2 trait	201	8.3
3–4	3.45 ± 0.24	?Hb CS trait†	26	1.1
4–8	5.70 ± 1.10	α-thal. 1 trait	88	3.7
9–13	10.34 ± 1.34	α-thal. 2 homozygosity	5	0.2
19–27	22.88 ± 3.45	α-thal. 1–α-thal. 2 (Hb H disease)	5	0.2
		total	371	15.4

† Hb CS was not detectable.

Pootrakul *et al.* are convinced that they have underestimated the incidence of Hb CS because of the technical difficulties in detecting it in cord-blood samples. An even higher incidence of α-thalassaemia has been found in Northern Thailand (Na-Nakorn & Wasi 1970) but nothing is known as yet about the incidence of Hb CS in this population.

Data on the gene frequency of Hb CS in Malaysia has come from a starch-gel electrophoretic survey of the haemoglobin types of 1785 persons of different racial groups (Lie Injo & Duraisamy 1972). The gene frequencies are summarized in table 12A. It appears that Hb CS occurs in about 2.2 % of Malays, 0.7 % of Chinese and 0.1 % of Indians. Again these must be minimal estimates because of the technical difficulties of demonstrating Hb CS. Preliminary data regarding the frequency of Hb CS in Hb H disease in the Chinese of Hong Kong and in the Chinese and Malay populations of Singapore are given in table 12A.

As shown in this paper Hb CS occurs in the Chinese in Jamaica but gene frequencies are unknown. It occurs sporadically also in Greeks and it appears that there is a relatively high incidence of the related chain-termination mutant, Hb Koya Dora, in parts of India where in a population survey of two tribes from Andhra Pradesh, the Koya Dora and the Konda Reddi, De Jong *et al.* (1975) have found an incidence of 10 % for Hb Koya Dora in the Koya Dora tribe.

The reasons for the remarkably high gene frequencies of the α-chain termination haemoglobin variants is unknown. Presumably since they produce the phenotype of α-thalassaemia they have been subjected to the same selective forces as the α-thalassaemia genes, whatever these may be!

(*f*) *The relation of the chain-termination mutants to thalassaemia*

The genetics and clinical features of the different forms of α-thalassaemia have been reviewed in detail recently (Weatherall & Clegg 1972*a*; Wasi 1973; Wasi *et al.* 1974). Although there is still considerable doubt about the mode of transmission of the α-thalassaemias in some racial groups a fairly straightforward picture of the inheritance of these disorders in Oriental and Mediterranean populations has emerged.

It is clear that the α-thalassaemias result from the interaction of two main genetic variants, the severe, or α-thal. 1, and less severe or α-thal. 2, genes (Wasi 1973). The α-thal. 1 and α-thal. 2 genes can be recognized in the heterozygous state by the finding of approximately 5 and 2 % respectively of Hb Bart's in the neonatal period. In the homozygous state for α-thal. 1 there are no α-chains produced and the clinical picture is that of Hb Bart's hydrops syndrome. In the heterozygous state for both α-thal. 1 and α-thal. 2 the clinical picture is that of Hb H disease.

The findings outlined in this paper provide good evidence that Hb CS can behave in an identical way to the α-thal. 2 gene in that it produces Hb H disease when it interacts with α-thal. 1, and is characterized by a mild thalassaemia-like disorder in its homozygous state. The inheritance pattern like that in family C in the present study (figure 1), in which one parent of a child with Hb H disease has Hb CS and the other α-thal. 1, has been amply confirmed by the data obtained from Malaysia. Lie Injo *et al.* (1971) studied 23 patients with Hb H disease, 12 of whom had Hb CS. In each case one parent had Hb CS and the other had a haematological picture compatible with α-thal. 1. On the other hand, in the remainder of the patients with Hb H disease who did not have Hb CS, the latter was not observed in either parent. Similar findings have been observed in Thailand recently (Wasi *et al.* 1974). Thus it

appears that the Hb-CS gene can, like the α-thal. 2 gene, interact with α-thal. 1 to produce Hb H disease.

Additional evidence that Hb CS can produce an α-thalassaemia-like disorder comes from observations of the Hb CS homozygote. Thus the affected child had slight splenomegaly, a thalassaemic blood picture and traces of Hb Bart's at the age of 12. The associations of Hb Icaria with a thalassaemic blood picture and of Hb Koya Dora with Hb H disease in India (De Jong *et al.* 1975), where it is rare, substantiate the link between the chain-termination mutations and α thalassaemia.

Furthermore, biosynthetic studies showed an α/β-chain production ratio in Hb-CS heterozygotes similar to that in α-thal. 2 heterozygotes. Recently workers in Thailand (S. Pootrakul, private communication) have obtained slightly different results suggesting that there is an overall imbalance of chain production but that it is due to the instability of the αCS chain. We have not observed this instability and are currently carrying out experiments in collaboration with the Thai group to try and sort out this discrepancy. We agree with the Thai group however that there is an overall chain imbalance in the Hb-CS heterozygotes.

Further evidence regarding the similarity between α-thal. 2 and Hb CS might be derived from studies of newborn infants carrying Hb CS. It was suggested by Weatherall (1963) that the presence of Hb Bart's in the neonatal period might be the most sensitive indicator of the presence of a mild α-thalassaemia gene. Thus at the time of the switch from γ- to β-chain synthesis there is normally a very slight degree of chain imbalance with the production of trace amounts of Hb Bart's. In the presence of a mild α-chain deficiency this process is exaggerated with the production of increased amounts of Hb Bart's at this particular phase of development. When adult haemoglobin synthesis is fully established the mild degree of chain imbalance is insufficient to give rise to detectable amounts of Hb H. This hypothesis has received considerable support from the studies in Thailand where it is clear that the α-thal. 1 and α-thal. 2 genes are associated with the presence of approximately 5 and 2% Hb Bart's respectively in the neonatal period, but when affected infants are followed into childhood they do not produce Hb H (Wasi *et al.* 1974).

It follows from the arguments outlined above that if the heterozygous state for Hb CS produces a similar degree of chain imbalance to the α-thal. 2 gene Hb-CS heterozygotes should have increased amounts of Hb Bart's in the neonatal period; preliminary data from Southeast Asia indicate that this is indeed the case. In a survey of the haemoglobin patterns of 1431 newborns in Malaysia (492 Malays, 506 Chinese and 438 Indians) 98 infants were found to have Hb Bart's in their cord blood (Lie-Injo 1973). Of these 39 had Hb Bart's levels above 5% and 38 had Hb Bart's levels below 3.8%; there was one infant with an intermediate level. It was found that none of the babies with more than 5% Hb Bart's in the cord blood had Hb CS. Among the babies with less than 3.8% Hb Bart's 7 had Hb CS. Family studies showed that none of the babies with high levels of Hb Bart's had parents with Hb CS whereas each of the 7 babies with CS had affected family members. Follow-up studies showed that the Hb Bart's disappeared as occurs in α-thal. 2 heterozygotes, while the Hb CS persisted. These observations indicate that the heterozygous state for Hb CS produces a haemoglobin pattern in the neonatal period identical to that of the α-thal. 2 gene. Lie-Injo also makes the interesting point that the incidence of Hb CS in Malaysia, as derived from cord-blood studies, is lower than that observed in the adult population survey data (see previous section). She believes that it is very easy to overlook trace amounts of Hb CS in newborn infants, particularly as

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

the samples tend to be variable age and concentration. It seems likely therefore that cord-blood surveys will underestimate the incidence of Hb CS.

A more recent cord-blood survey carried out in Thailand (Pootrakul *et al.* 1974) has confirmed and extended the findings reported from Malaysia. Thus 46 infants with Hb CS all had Hb Bart's and it was possible to identify the heterozygous state for Hb CS and the Hb CS/α-thal. 1 and the Hb CS/α-thal. 2 combinations by virtue of the levels of Hb Bart's (table 12B). Further there was a significant reduction of the mean cell haemoglobins (m.c.h.) and mean cell volumes (m.c.v.) in the Hb-CS group as compared with normal control cord bloods; the red-cell indices in the Hb-CS group were similar to those of the α-thal. 2 heterozygotes found in the same survey.

Thus the genetic evidence from studies of Hb H disease and the pattern of haemoglobin production of Hb-CS heterozygotes in the neonatal period provides very strong evidence that the heterozygous state for Hb CS produces an identical clinical and haematological picture to the α-thal. 2 mutation. This observation raises the intriguing question of whether all cases of α-thal. 2 are in fact due to the Hb-CS mutation; is it possible that very low amounts of Hb CS may be either overlooked or just not demonstrable by current electrophoretic techniques? We have examined concentrated red-cell lysates prepared from the blood of Hb H disease patients from a variety of racial groups and have convinced ourselves that in many of them, even using starch-gel electrophoretic techniques which would have been capable of picking up as little as 0.2 % of a slowly-migrating haemoglobin variant, there are no Hb-CS components. However, since Hb-CS synthesis declines very early during erythroid maturation, it is possible that in some cases trace amounts of Hb CS are synthesized in young red-cell precursors but not in sufficient quantities to be demonstrable in the peripheral blood. It would be well worth while looking for traces of Hb CS in early red-cell precursors obtained from bone-marrow samples from patients with Hb H disease who do not have Hb CS in their peripheral blood.

(*g*) *A summary of the molecular basis for the α-thalassaemias*

From the data which has been presented in this paper together with that derived from recent studies on the molecular basis of the Hb Bart's hydrops syndrome it is now possible to construct a fairly complete picture of the molecular basis for the common forms of α-thalassaemia, at least as they occur in the Oriental and Mediterranean populations. The α-thalassaemia disorders of clinical importance are the Hb Bart's hydrops syndrome and Hb H disease. There is good genetic evidence that the former condition results from the homozygous state for the α-thal. 1 gene while the latter can result from either the heterozygous states for α-thal. 1 and α-thal. 2, or for α-thal. 1 and Hb CS or related haemoglobins. Any model which attempts to correlate all these observations must take into consideration the following: (1) the strong likelihood that there are at least two α-chain loci on each chromosome in many human populations; (2) that although there is strong evidence for the presence of α-thalassaemia in other populations such as Africa and Saudi Arabia, the Hb Bart's hydrops syndrome and Hb H disease are rare outside Southeast Asia and parts of the Mediterranean, and (3) that a structural variant, Hb Q (α 74 Asp → His), can apparently interact with the α-thal. 1 gene to produce a condition similar to Hb H disease called Hb Q α-thalassaemia (Vella, Wells, Agar & Lehmann 1958; Lorkin *et al.* 1970). In this condition the haemoglobins consist of Hbs Q ($\alpha_2^Q\beta_2$), Q_2 ($\alpha_2^Q\delta_2$), H and Bart's; Hbs A and A_2 are absent indicating that no normal α chains are produced in persons heterozygous for both Hb Q and α-thal. 1.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

The existence of at least two α-chain loci on each chromosome provides an opportunity for a series of α-thalassaemia disorders. Lehmann (1970) suggested that an α-thalassaemia mutation might involve from one to four of these loci and, furthermore, if only two loci were involved these might be on the same or opposite pairs of homologous chromosomes. The α-thal. 1 gene might result from α-thalassaemia mutations at both linked α loci on the same chromosome and thus in the homozygous state all four loci would be involved, resulting in the Hb Bart's hydrops syndrome. The α-thal. 2 gene would result from an α-thalassaemia mutation at only one of a pair of linked α-chain loci. In some populations, such as the African Negro, an α-thalassaemia mutation may never have involved two α-chain loci on the same chromosome so that the Hb Bart's hydrops syndrome cannot occur. These ideas have been expanded by Weatherall & Clegg (1972*a*) and Wasi (1973) and extended to include populations who might have more than two α-chain genes on each chromosome (Pembrey *et al.* 1975). Can they now be applied to the data found in the present study and what can we deduce about the molecular basis of these α-thalassaemia mutations?

Recently Ottolenghi *et al.* (1974) and Taylor *et al.* (1974) have used DNA-hybridization techniques conclusively to demonstrate that the Hb Bart's hydrops syndrome results from a complete deletion of the α-chain loci of affected infants. This observation is entirely consistent with the previous findings that affected babies make no α chains or fragments of α chains (Weatherall *et al.* 1970) and that they have a deficiency of α mRNA (Kan, Todd & Dozy 1974). This suggests that the α-thal. 1 mutation results from a complete loss of the α-chain genes from one chromosome and hence that these loci must be closely linked. Studies of haemoglobin synthesis in Hb H disease have indicated that the reduced rate of α-chain production in this disorder is associated with a reduced amount of α-chain mRNA (Grossbard, Terada, Dow & Bank 1973; Benz, Swerdlow & Forget 1973; Pritchard, Clegg, Weatherall & Longley 1974; Clegg & Weatherall 1974*b*). When taken with the recent studies of Ottolenghi *et al.* and Taylor *et al.* it seems very likely that the α-thal. 2 mutation results from a deletion of one of a pair of α-chain loci, and thus it follows that patients with Hb H disease have a single functioning α-chain locus. It should be possible to test this hypothesis in the near future as purer globin-cDNA probes become available.†

If the α-thal. 1 mutation results from the loss of both linked α-chain loci from a chromosome and α-thal. 2 from the loss of one of the pair, how do the chain-termination mutants fit into the picture? Since individuals who have inherited α-thal. 1 and Hb CS have Hb H disease in which both normal α chains and αCS chains are produced, it seems likely that the chromosome carrying the α locus with the UAA → CAA change in its chain-termination codon also carries a normal α locus. However since Hb CS is produced so inefficiently the overall output from this chromosome will be the same as if it had only a single α-chain locus so that it produces an identical effect to an α-thal. 2 mutation. This interpretation is fully consistent with all the findings outlined in this paper. Thus it appears that all the common α-thalassaemias result from either deletions or chain-termination mutations and that the clinical syndromes of α-thalassaemia result from a series of combinations of these molecular abnormalities all of which result in an overall deficit of α-chain production (table 13). Lehmann & Lang (1974) have suggested that the loss of the α-chain loci has resulted from unequal crossing over of two chromosomes with two linked α-chain genes, a mechanism similar to that suggested for the

† Recent studies by Kan *et al.* (*Nature, Lond.* **255**, 255, 1975) indicate that there is a single α-chain locus in Hb H disease.

production of the different Lepore haemoglobins. Certainly this seems a likely mechanism for the loss of different numbers of α loci in the various types of α-thalassaemia.

Finally what can we deduce about Hb Q α-thalassaemia and α-thalassaemia as it occurs in populations other than those of Southeast Asia and the Mediterranean. The most likely explanation for the findings in Hb Q α-thalassaemia is that the original mutation which caused Hb Q occurred on a chromosome carrying a single α-chain locus, or that if it occurred on a chromosome carrying two loci, one was then lost by an event similar to that which has produced the α-thal. 2 gene, so that when it is inherited together with an α-thal. 1 gene it will produce a condition identical to Hb H disease, i.e. Hb Q α-thalassaemia.

TABLE 13. α-THALASSAEMIA

(*a*) *molecular basis*

α-thalassaemia mutation	molecular mechanism
α-thal. 1	deletion of both α-chain loci
α-thal. 2	deletion of one α-chain locus
Hb CS and related variants	chain-termination mutations
Hb Q	point mutation in an α-chain locus on a chromosome missing one α-chain locus
α-thal. in Africa and Saudi Arabia	? loss of one α-chain locus in populations with variable numbers of α-chain loci, possibly more than two per chromosome

(*b*) *clinical findings related to genotype*

Some of the Hb Bart's levels obtained from Pootrakul *et al.* (1974). α^0, deleted α-chain locus.

suggested α-thal. genotype	clinical or haematological findings
$\alpha^0\alpha^0/\alpha\alpha$	mild thalassaemia disorder; 5% Hb Bart's at birth
$\alpha^0\alpha/\alpha\alpha$	normal haematology; 2–3% Hb Bart's at birth
$\alpha\alpha^{CS}/\alpha\alpha$	normal haematology; 1% Hb CS, 2–3% Hb Bart's at birth
$\alpha^0\alpha^0/\alpha^0\alpha^0$	Hb Bart's hydrops
$\alpha^0\alpha^0/\alpha^0\alpha$	Hb H disease
$\alpha^0\alpha^0/\alpha\alpha^{CS}$	Hb H disease
$\alpha^0\alpha/\alpha^0/\alpha$	mild thalassaemia; 5% Hb Bart's at birth
$\alpha\alpha^{CS}/\alpha\alpha^{CS}$	mild thalassaemia; 5–6% Hb CS; trace of Hb Bart's
$\alpha^0\alpha/\alpha\alpha^{CS}$	mild thalassaemia; 5% Hb Bart's at birth; trace of Hb CS

However the α-thalassaemia disorders as they appear in Negroes and in other high frequency areas such as Saudi Arabia (reviewed by Weatherall & Clegg 1972*a*) are not so easily explained by the model outlined above. There is a relatively high incidence of Hb Bart's in infancy and evidence of α-thalassaemia-like disorders in adults in these populations, but Hb H disease is rare and the Hb Bart's hydrops condition has not been reported. Maybe this paradox can be explained at least in part by the hypothesis proposed by Lehmann (1970). If there are populations in which there is an α-thal.-2-like gene, i.e. one α locus is inactive, but in which there has not been loss of both α-chain loci on homologous chromosomes, the Hb Bart's hydrops syndrome cannot occur. It is interesting to note that in the only case of the presence of an α-chain variant, Hb I, and α-thalassaemia in one individual to be reported in an African (Atwater *et al.* 1960), there was about 30% Hb A present, indicating that some α-chain synthesis must have been directed by the chromosome carrying the α-thalassaemia gene. If it is

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

assumed that in these populations there is only an α-thal. 2-like gene this should be reflected in the levels of Hb Bart's in the newborn infants. These would be expected to fall into two groups, one of about 2–3 % representing the heterozygous state and the other at about 5–6 % representing the homozygous state. The reported levels of Hb Bart's in Negro infants have varied remarkably and the only worker who has found values which fall into these modes is Harris (1973) while others have found levels ranging from 2 to 10 % with no evidence of bimodality. Certainly there is no evidence for such a distribution of Hb Bart's levels in Saudi Arabia (Pembrey *et al.* 1975).

The clinical and haemoglobin-analytical data of α-thalassaemia in Africans and Saudi Arabians has been reviewed in detail by Weatherall & Clegg (1972*a*) and Pembrey *et al.* (1975) and one must conclude from all the published data that the findings are not compatible with the more clear-cut genetic picture which appears from Southeast Asia and the Mediterranean. One possible explanation for the remarkable variability in the α-thalassaemia syndromes is that different numbers of α-chain loci occur in different populations. Indeed it may be that even within the same population individuals may have different numbers of α-chain genes. Certainly at least three have been found in certain primates (Boyer *et al.* 1971) and there is good evidence for the existence of multiple α-chain loci in other animals. A series of deletions or chain-termination mutations involving different numbers of these loci would explain much of the heterogeneity of α-thalassaemias within different populations and indeed may go some way to explain the heterogeneity of the β-thalassaemias since presumably where there is defective β-chain synthesis the degree of α-chain excess may depend at least in part on the number of α-chain loci present. It is interesting that the mild forms of β-thalassaemia are found commonly among the African Negro population (reviewed by Wilcox, Weatherall & Clegg 1975) and even within this population there is a remarkable heterogeneity of the disorder.

These studies would not have been possible without the help and collaboration of our many overseas colleagues. In particular we wish to thank Dr Paul Milner of Jamaica for sending us the material from family C. The haematological studies on this family were carried out almost entirely by Dr Milner and his colleagues in Jamaica. Further samples containing Hb CS were obtained from Dr Lie-Injo Luan Eng of Kuala Lumpur, Malaysia, Professor Phaedon Fessas of Athens, Professor David Todd of Hong Kong and Dr Supa Na-Nakorn and Dr Prawase Wasi of Bangkok, Thailand. The family in which Hb Icaria was discovered was ascertained by Professor Tsevrenis and his colleagues from Athens who also carried out the haematological studies on this family. Dr Lie-Injo Luan Eng also sent us samples from the family in which there was an individual homozygous for Hb CS and she was responsible for all the haematological studies on this family. We are also indebted to Professor Wong Hock Boon of Singapore and Dr Sanga Pootrakul of Bangkok, Thailand, for allowing us to quote unpublished data regarding the gene frequency of Hb CS. Throughout these studies D.J.W. and J.B.C. received financial support from the Medical Research Council and the Wellcome Trust. In addition the Wellcome Trust supplied a special grant for a visit to Jamaica to carry out further studies on family C.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

References

Abramson, R. K., Rucknagel, D. L., Shreffler, D. C. & Saave, J. J. 1970 Homozygous Hb J Tongariki: evidence for only one alpha chain structural locus in Melanesians. *Science, N.Y.* **169**, 194–196.

Allen, D. W., Schroeder, W. A. & Balog, J. 1958 Observations on chromatographic heterogeneity of normal adult and fetal human hemoglobin. *J. Am. chem. Soc.* **80**, 1628–1634.

Ambler, R. P. 1963 The purification and amino acid composition of pseudomonas cytochrome C-551. *Biochem. J.* **89**, 341–349.

Ambler, R. P. & Medway, R. J. 1968 The use of thermolysin in amino acid sequence determination. *Biochem. J.* **108**, 893–895.

Atwater, J., Schwartz, I. R., Erslev, A. J., Montgomery, T. D. & Tocantins, L. M. 1960 Sickling of erythrocytes in a patient with thalassemia-hemoglobin-I disease. *New Engl. J. Med.* **263**, 1215–1223.

Baglioni, C. 1962 The fusion of two peptide chains in hemoglobin Lepore and its interpretation as a genetic deletion. *Proc. natn. Acad. Sci.* (*Wash.*) **48**, 1880–1886.

Benz, E. J., Swerdlow, P. S. & Forget, B. G. 1973 Globin messenger RNA in hemoglobin H disease. *Blood* **42**, 825–833.

Blobel, G. 1971 Release, identification and isolation of messenger RNA from mammalian ribosomes. *Proc. natn. Acad. Sci.* (*Wash.*) **68**, 832–835.

Boyer, S. H., Noyes, A. N., Vrablik, G. R., Donaldson, L. J. & Schaefer, E. W. 1971 Silent hemoglobin alpha genes in apes: potential source of thalassemia. *Science, N.Y.* **171**, 182–185.

Bray, G. A. 1960 A simple efficient liquid scintillator for counting solutions in a liquid scintillation counter. *Annal. Biochem.* **1**, 279–285.

Brimhall, B., Hollan, S., Jones, R. T., Koler, R. D., Stocklen, Z. & Szelenyi, J. G. 1970 Multiple alpha-chain loci for human hemoglobin. *Clin. Res.* **18**, 184.

Burr, H. & Lingrel, J. B. 1971 Poly A sequences at the 3′ termini of rabbit globin mRNAs. *Nature, New Biol.* **233**, 41–43.

Capp, G. L., Rigas, D. A. & Jones, R. T. 1970 Evidence for a new haemoglobin chain (ʒ-chain). *Nature, Lond.* **228**, 278–280.

Clegg, J. B. 1971 Gene duplication and haemoglobin polymorphism. In *Ecological genetics and evolution* (ed. R. Creed), 298–308. Oxford: Blackwell Scientific Publications.

Clegg, J. B., Naughton, M. A. & Weatherall, D. J. 1966 Abnormal human haemoglobins: separation and characterization of the α- and β-chains by chromatography, and the determination of two new variants Hb Chesapeake and Hb J (Bangkok). *J. molec. Biol.* **19**, 91–108.

Clegg, J. B. & Weatherall, D. J. 1967 Haemoglobin synthesis in alpha-thalassaemia (Haemoglobin H disease). *Nature, Lond.* **215**, 1241–1243.

Clegg, J. B. & Weatherall, D. J. 1974*a* β°-thalassaemia – time for a reappraisal? *Lancet* iii, 133–135.

Clegg, J. B. & Weatherall, D. J. 1974*b* Hemoglobin Constant Spring, an unusual α-chain variant involved in the etiology of hemoglobin H disease. *Ann. N.Y. Acad. Sci.* **232**, 168–178.

Clegg, J. B., Weatherall, D. J., Contopolou-Griva, I., Caroutsos, K., Poungouras, P. & Tsevrenis, H. 1974 Haemoglobin Icaria, a new chain termination mutant which causes α thalassaemia. *Nature, Lond.* **251**, 245–247.

Clegg, J. B., Weatherall, D. J. & Milner, P. F. 1971 Haemoglobin Constant Spring – A chain termination mutant? *Nature, Lond.* **234**, 337–340.

Clegg, J. B., Weatherall, D. J., Na-Nakorn, S. & Wasi, P. 1968 Haemoglobin synthesis in β thalassaemia. *Nature, Lond.* **220**, 664–668.

Cooley, T. B. & Lee, T. 1925 A series of cases of splenomegaly in children with anemia and peculiar bone changes. *Trans. Am. pediat. Soc.* **37**, 29.

Dacie, J. V. & Lewis, S. M. 1970 *Practical haematology*, 4th ed. London: J. & A. Churchill.

De Jong, W. W., Khan, P. M. & Bernini, L. F. 1975 Hemoglobin Koya Dora; high frequency of a chain termination mutant *Am. J. Hum. Genet.* **27**, 81–89.

Efremov, G. D., Wrightstone, R. N., Huisman, T. H. J., Schroeder, W. A., Hyman, C., Ortega, J. & Williams, K. 1971 An unusual hemoglobin anomaly and its relation to α-thalassemia and hemoglobin H disease *J. clin. Invest.* **50**, 1628–1636.

Fessas, P. 1963 Inclusions of hemoglobin in erythroblasts and erythrocytes of thalassemia. *Blood* **21**, 21–32.

Fessas, P., Lie Injo Luan Eng, Na-Nakorn, S., Todd, D., Clegg, J. B. & Weatherall, D. J. 1972 Identification of slow-moving haemoglobins in haemoglobin H disease from different racial groups. *Lancet* **i**, 1308–1310.

French, E. A. & Lehmann, H. 1971 Is haemoglobin Gα Philadelphia linked to α-thalassaemia? *Acta haemat.* **46**, 149.

Gaskill, P. & Kabat, D. 1971 Unexpectedly large size of globin messenger ribonucleic acid. *Proc. natn. Acad. Sci.* (*Wash.*) **68**, 72–75.

Gerald, P. S. & Diamond, L. K. 1958 The diagnosis of thalassemia trait by starch block electrophoresis of the hemoglobin. *Blood* **13**, 61–69.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Gouttas, A., Fessas, P., Tsevrenis, H. & Xefteri, E. 1955 Description d'une nouvelle variété d'anémie hémolytique congenitale. *Sang* **26**, 911–919.

Gray, W. R. 1967 Sequential degradation plus dansylation. In *Methods of enzymology*, vol. 11, 469–475 (eds. S. P. Colowick & N. O. Kaplan). New York: Academic Press.

Grossbard, E., Terada, M., Dow, L. W. & Bank, A. 1973 Decreased α-globin messenger RNA activity associated with polyribosomes in α thalassaemia *Nature, New Biol.* **241**, 209–211.

Harris, D. J. 1973 Levels of hemoglobin Bart's in a sample of American black infants. *Abstract Am. J. Hum. Genet.* **25**, 32 A.

Hollan, S. R., Szelenyi, J. G., Brimhall, B., Duerst, M., Jones, R. T., Koler, R. D. & Stocklen, Z. 1972 Multiple alpha chain loci for human haemoglobins: Hb J-Buda and Hb G-Pest. *Nature, Lond.* **235**, 47–50.

Hunt, J. A. & Lehmann, H. 1959 Haemoglobin 'Bart's': a foetal haemoglobin without α-chains. *Nature, Lond.* **184**, 872–873.

Ingram, V. M. & Stetton, A. O. W. 1959 Genetic basis of the thalassaemia diseases *Nature, Lond.* **184**, 1903–1909.

Kan, Y. W., Todd, D. & Dozy, A. M. 1974 Haemoglobin Constant Spring synthesis in red cell precursors. *Brit. J. Haemat.* **28**, 103–106.

Lang, A., Lehmann, H. & King-Lewis, P. A. 1974 Hb K Woolwich the cause of a thalassaemia. *Nature, Lond.* **249**, 467–468.

Lang, A., White, J. M. & Lehmann, H. 1972 Synthesis of Hb Lepore ($\alpha_2 \delta\beta_2$): influence of δ and β nucleotide sequence on synthesis of δβ chain. *Nature, New Biol.* **240**, 268–271.

Lehmann, H. 1970 Different types of alpha-thalassaemia and significance of haemoglobin Bart's in neonates. *Lancet* ii, 73–75.

Lehmann, H. & Carrell, R. W. 1968 Difference between α and β-chain mutants of human haemoglobin and between α and β-thalassaemia. Possible duplication of the α-chain gene. *Br. med. J.* **4**, 748–750.

Lehmann, H. & Lang, A. 1974 Various aspects of α-thalassaemia. *Ann. N.Y. Acad. Sci.* **232**, 152–158.

Lie-Injo Luan Eng 1973 Haemoglobin Bart's and slow-moving haemoglobin X components in newborns. *Acta Haemat.* **49**, 25–35.

Lie-Injo Luan Eng & Duraisamy, G. 1972 The slow moving haemoglobin X components in Malaysians. *Human Hered.* **22**, 118–123.

Lie-Injo Luan Eng, Ganesan, J., Clegg, J. B. & Weatherall, D. J. 1974 Homozygous state for Hb Constant Spring (slow moving Hb X components). *Blood* **43**, 251–259.

Lie-Injo Luan Eng & Jo, Bwan Hie 1960 A fast-moving haemoglobin in hydrops foetalis. *Nature, Lond.* **185**, 698.

Lie-Injo Luan Eng, Lopez, C. G. & Lopez, M. 1971 Inheritance of haemoglobin H disease. *Acta haemat.* **46**, 106–120.

Lim, L. & Canellakis, E. S. 1970 Adenine-rich polymer associated with rabbit reticulocyte messenger RNA. *Nature, Lond.* **227**, 710–712.

Lingrel, J. B. & Borsook, H. 1963 A comparison of amino acid incorporated into the hemoglobin and ribosomes of marrow erythroid cells and circulating reticulocytes of several anaemic rabbits. *Biochemistry* **2**, 309–320.

Lorkin, P. A., Charlesworth, D., Lehmann, H., Rahbar, S., Tuchinda, S. & Lie-Injo, Luan Eng 1970 Two haemoglobins Q α74 (EF3) and α75 (EF4) Aspartic Acid → Histidine. *Br. J. Haemat.* **19**, 117–125.

Marotta, C. A., Forget, B. G., Weissman, S. M., Verma, I. M., McCaffrey, R. P. & Baltimore, D. 1974 Nucleotide sequences of human globin messenger RNA. *Proc. natn. Acad. Sci.* (*Wash.*) **71**, 2300–2304.

Milner, P. F., Clegg, J. B. & Weatherall, D. J. 1971 Haemoglobin H disease due to a unique haemoglobin variant with an elongated α-chain. *Lancet* i, 729–732.

Milstein, C. 1966 The disulphide bridges of immunoglobin x-chains. *Biochem. J.* **101**, 338–351.

Milstein, C., Clegg, J. B. & Jarvis, J. M. 1968 Immunoglobin λ-chains – The complete amino acid sequence of a Bence-Jones protein. *Biochem. J.* **110**, 631–654.

Na-Nakorn, S. & Wasi, P. 1970 Alpha thalassemia in Northern Thailand. *Am. J. Hum. Genet.* **22**, 645–651.

Ottolenghi, S., Lanyon, W. G., Paul, J., Williamson, R., Weatherall, D. J., Clegg, J. B., Pritchard, J., Pootrakul, S. & Wong, Hock Boon 1974 The severe form of α-thalassaemia is caused by a haemoglobin gene deletion. *Nature, Lond.* **251**, 389–392.

Pembrey, M. E., Weatherall, D. J., Clegg, J. B. & Perrine, R. P. 1975 Haemoglobin Bart's in Saudi Arabia. *Br. J. Haemat.* **29**, 221.

Pootrakul, S., Wasi, P., Pornpatkul, M. & Na-Nakorn, S. 1970 Incidence of alpha thalassemia in Bangkok. *J. med. Assn Thailand.* **53**, 250.

Pootrakul, S., Wasi, P. & Na-Nakorn, S. 1974 Private communication.

Pritchard, J., Clegg, J. B., Weatherall, D. J. & Longley, J. 1974 Isolation of human messenger RNA. *Br. J. Haemat.* **28**, 141.

Rieder, R. F. & Weatherall, D. J. 1965 Studies on hemoglobin biosynthesis; Asynchronous synthesis of hemoglobin A and hemoglobin A_2 by human erythrocyte precursors. *J. clin. Invest.* **44**, 42–50.

Rigas, D. A., Koler, R. D. & Osgood, E. E. 1955 New hemoglobin possessing a higher electrophoretic mobility than normal adult hemoglobin. *Science, N.Y.* **121**, 372.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Roberts, A. V., Weatherall, D. J. & Clegg, J. B. 1972 The synthesis of human haemoglobin A_2 during erythroid maturation. *Biochem. biophys. Res. Commun.* **47**, 81.

Sanger, F. & Thompson, E. O. P. 1963 Halogenation of tyrosine during acid hydrolysis. *Biochim. biophys. Acta* **71**, 468–471.

Schwartz, E. 1969 The silent carrier of beta thalassaemia. *N. Engl. J. Med.* **281**, 1327–1333.

Seid-Akhaven, M., Winter, W. P. & Rucknagel, D. 1972 Haemoglobin Wayne. A frame shift variant occurring in two distinct forms. *Blood* **40**, 927.

Sofraniadou, K., Kaltsoya, A., Loukopoulos, D. & Fessas, P. 1968 Hemoglobin 'Athens', an alpha-chain variant with unusual properties. *Abstracts of simultaneous sessions. 12th Int. cong. Haemat*, p. 56. New York: Grune and Stratton.

Stretton, A. O. W. 1965 The genetic code. *Br. med. Bull.* **21**, 229–235.

Taylor, J. M., Dozy, A., Kan, Y. W., Varmus, H. E., Lie Injo Luan Eng, Ganeson, J. & Todd, D. 1974 Genetic lesion in homozygous α thalassamemia (hydrops fetalis). *Nature, Lond.* **251**, 392–394.

Todd, D. 1971 Slow-moving haemoglobin bands in haemoglobin H disease. *Lancet*, ii, 439.

Vella, F., Wells, R. H. C., Ager, J. A. M. & Lehmann, H. 1958 A haemoglobinopathy involving haemoglobin H and a new (Q) haemoglobin. *Br. med. J.* **1**, 752–755.

Wasi, P. 1973 Is the human globin α-chain locus duplicated? *Br. J. Haemat.* **24**, 267–273.

Wasi, P., Na-Nakorn, S. & Pootrakul, S. 1974 The α-thalassaemias. In *Clinics in Haematology* **3**, 383 (eds. D. J. Weatherall & W. B. Saunders). London:

Wasi, P., Na-Nakorn, S., Pootrakul, P. & Panich, V. 1972 Incidence of haemoglobin Thai: a re-examination of the genetics of α-thalassaemic disease. *Ann. Hum. Genet. Lond.* **35**, 467–470.

Wasi, P., Na-Nakorn, S., Pootrakul, S., Sookanek, M., Disthasongchan, P., Pornpatkul, M. & Panich, V. 1969 Alpha and beta-thalassemia in Thailand. *Ann. N.Y. Acad. Sci.* **165**, 60–82.

Weatherall, D. J. 1963 Abnormal haemoglobins in the neonatal period and their relationship to thalassaemia. *Br. J. Haemat.* **9**, 265–277.

Weatherall, D. J. & Boyer, S. H. 1962 Evidence for the genetic identity of alpha chain determinants in hemoglobins A, A_2 and F. *Bull. Johns Hopkins Hosp.* **110**, 8–22.

Weatherall, D. J. & Clegg, J. B. 1972*a* *The thalassaemia syndromes*, 2nd ed. Oxford: Blackwell Scientific Publications.

Weatherall, D. J. & Clegg, J. B. 1972*b* α-thalassaemia due to a structural haemoglobin variant. *Biochimie* **54**, 665.

Weatherall, D. J., Clegg, J. B., Na-Nakorn, S. & Wasi, P. 1969 The pattern of disordered haemoglobin synthesis in homozygous and heterozygous β-thalassaemia. *Br. J. Haemat.* **16**, 251–267.

Weatherall, D. J., Clegg, J. B. & Naughton, M. A. 1965 Globin synthesis in thalassaemia: an *in vitro* study. *Nature, Lond.* **208**, 1061–1065.

Weatherall, D. J., Clegg, J. B. & Wong, H. B. 1970 The haemoglobin constitution of infants with the haemoglobin Bart's hydrops foetalis syndrome. *Br. J. Haemat.* **18**, 357–367.

Weatherall, D. J., Gilles, H. M., Clegg, J. B., Blankson, J. A., Mustafa, D., Boi-Doku, F. S. & Chaudhury, D. S. 1971 Preliminary surveys for the incidence of the thalassaemia genes in some African populations *Ann. trop. Med. Parasit.* **65**, 253–265.

White, H. B., Laux, B. E. & Dennis, D. 1972 Messenger RNA structure, compatibility of hairpin loops with protein sequence. *Science, N.Y.* **175**, 1264.

Willcox, M., Weatherall, D. J. & Clegg, J. B. 1975 Homozygous β-thalassaemia in Liberia. *J. med. Genet.* (in the Press).

Wong, H. B. 1974 Private communication.

Woods, K. R. & Wang, K.-T. 1967 Separation of dansyl-amino acids by polyamide layer chromatography. *Biochim. biophys. Acta* **133**, 369–370.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

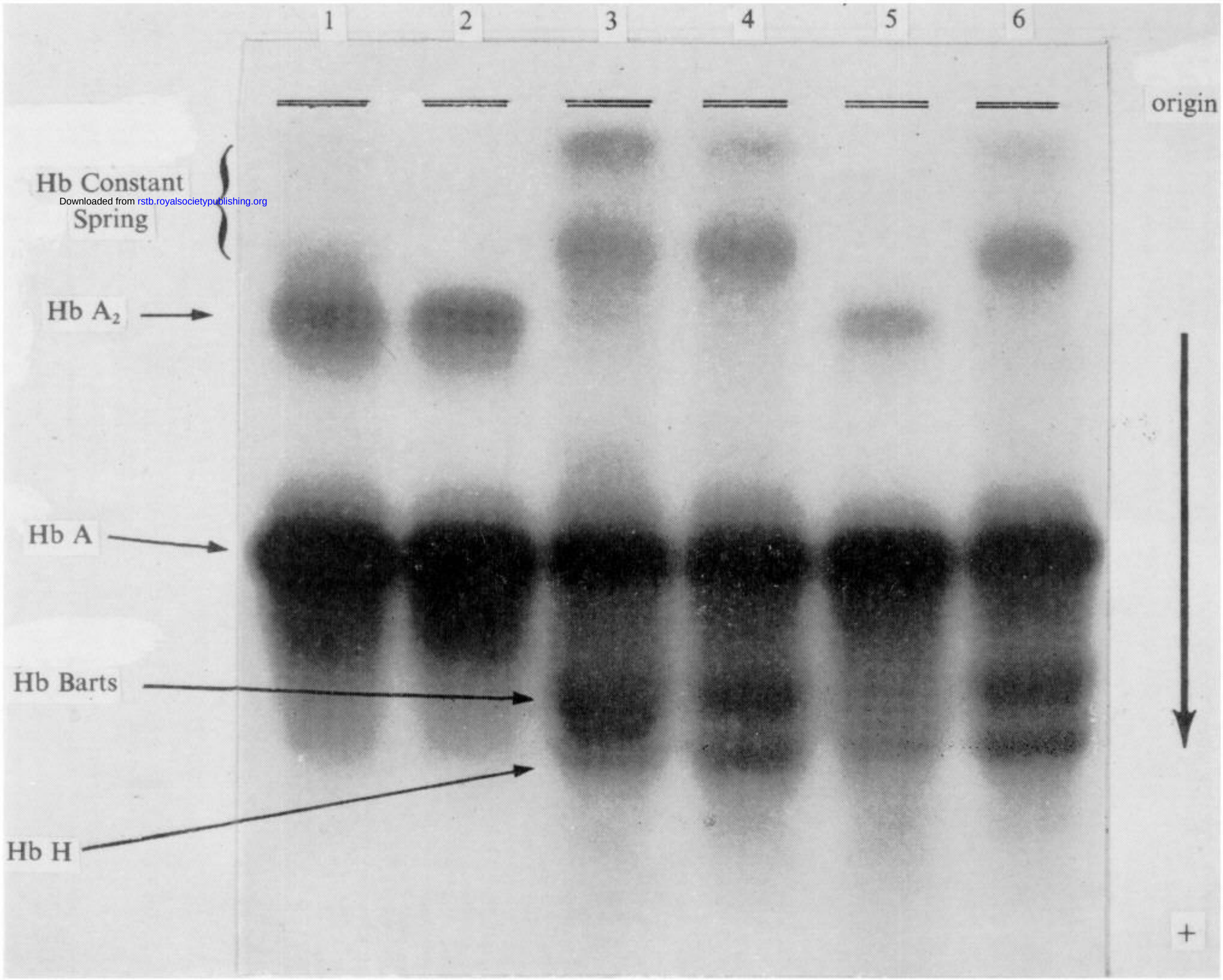


FIGURE 2. Haemoglobin analysis of members of family C. The lysates shown are from the following (left to right): 1, an Hb-CS heterozygote; 2, a normal individual; 3 and 4, siblings heterozygous for α-thal. 1 and Hb CS; 5, an individual with Hb H disease without Hb CS (control not from family C); 6, the third sibling heterozygous for Hb CS and α-thal. 1. Vertical descending starch-gel electrophoresis, tris–EDTA–borate buffer system, pH 8.5, benzidine stain.

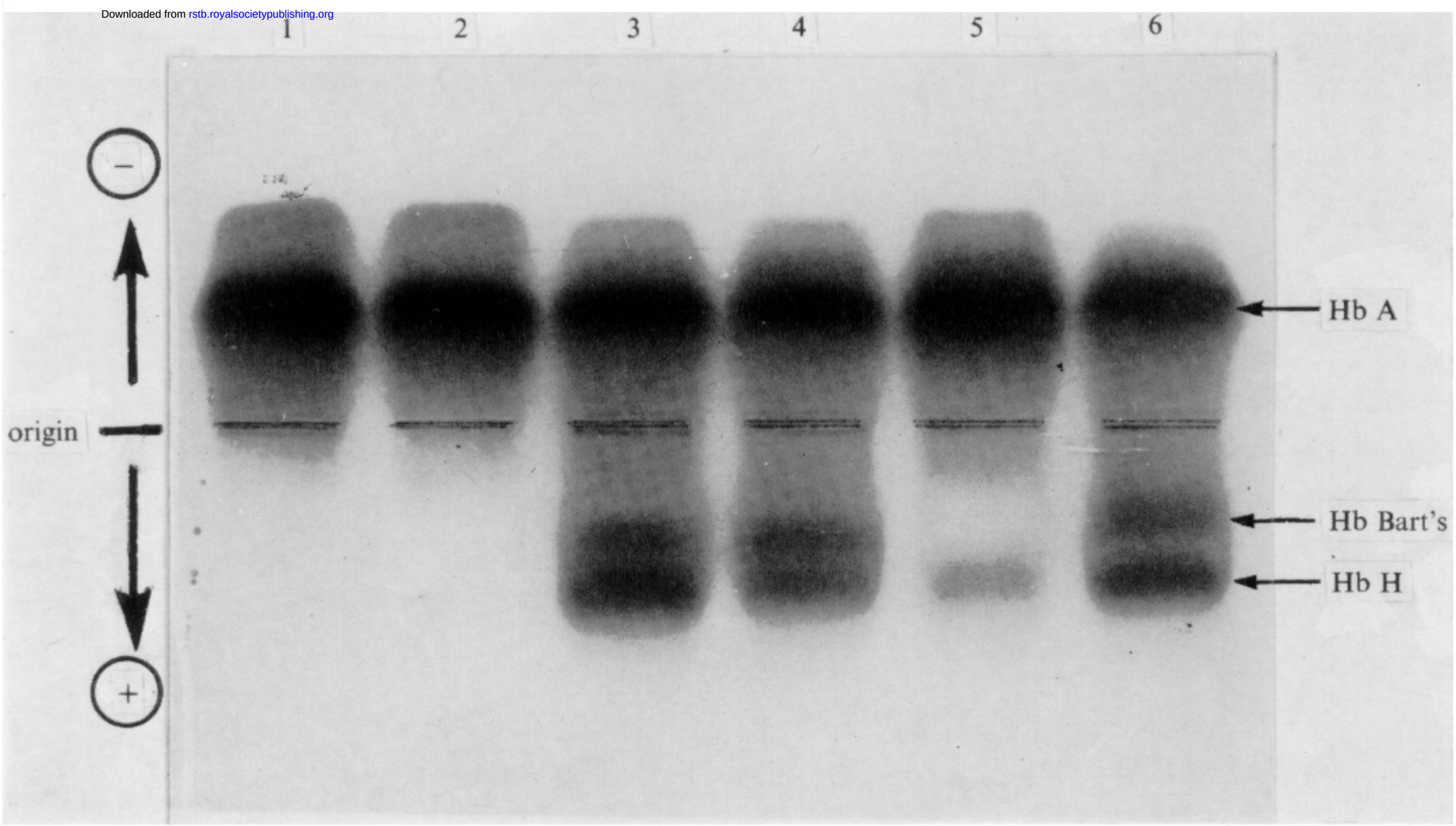


FIGURE 3. Haemoglobin analysis of members of family C. The lysates are from the following (left to right): 1 and 2, normal individuals; 3 and 4, individuals heterozygous for α-thal. 1 and Hb CS; 5, an individual with Hb H disease without Hb CS (control); 6, an individual heterozygous for α-thal. 1 and Hb CS. The anodal migration at pH 7.0 is characteristic of Hbs H and Bart's and Hb CS did not separate under these conditions. Vertical descending starch-gel electrophoresis, phosphate buffer system, pH 7.0, benzidine stain.

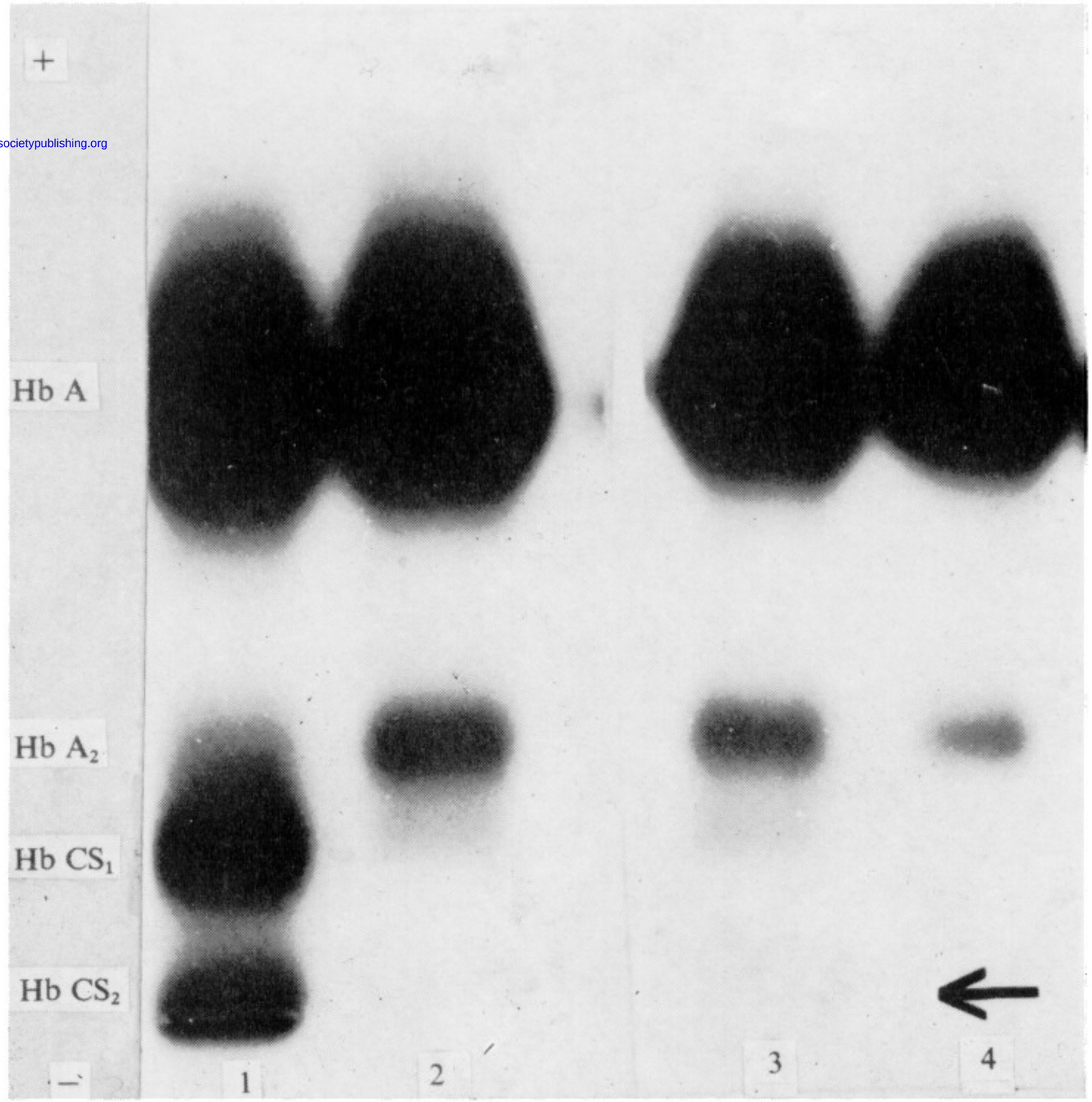


FIGURE 4. Haemoglobin analysis of lysates from members of the family with an Hb-CS homozygote. The lysates are from the following (left to right): 1, the child homozygous for Hb CS; 2 and 3, the heterozygous parents; 4, a normal individual. In addition to Hbs CS_1 and CS_2 there were some faint bands visible behind the origin which were not strong enough to photograph. Vertical descending starch-gel electrophoresis, tris–EDTA–borate buffer system, pH 8.5, benzidine stain.

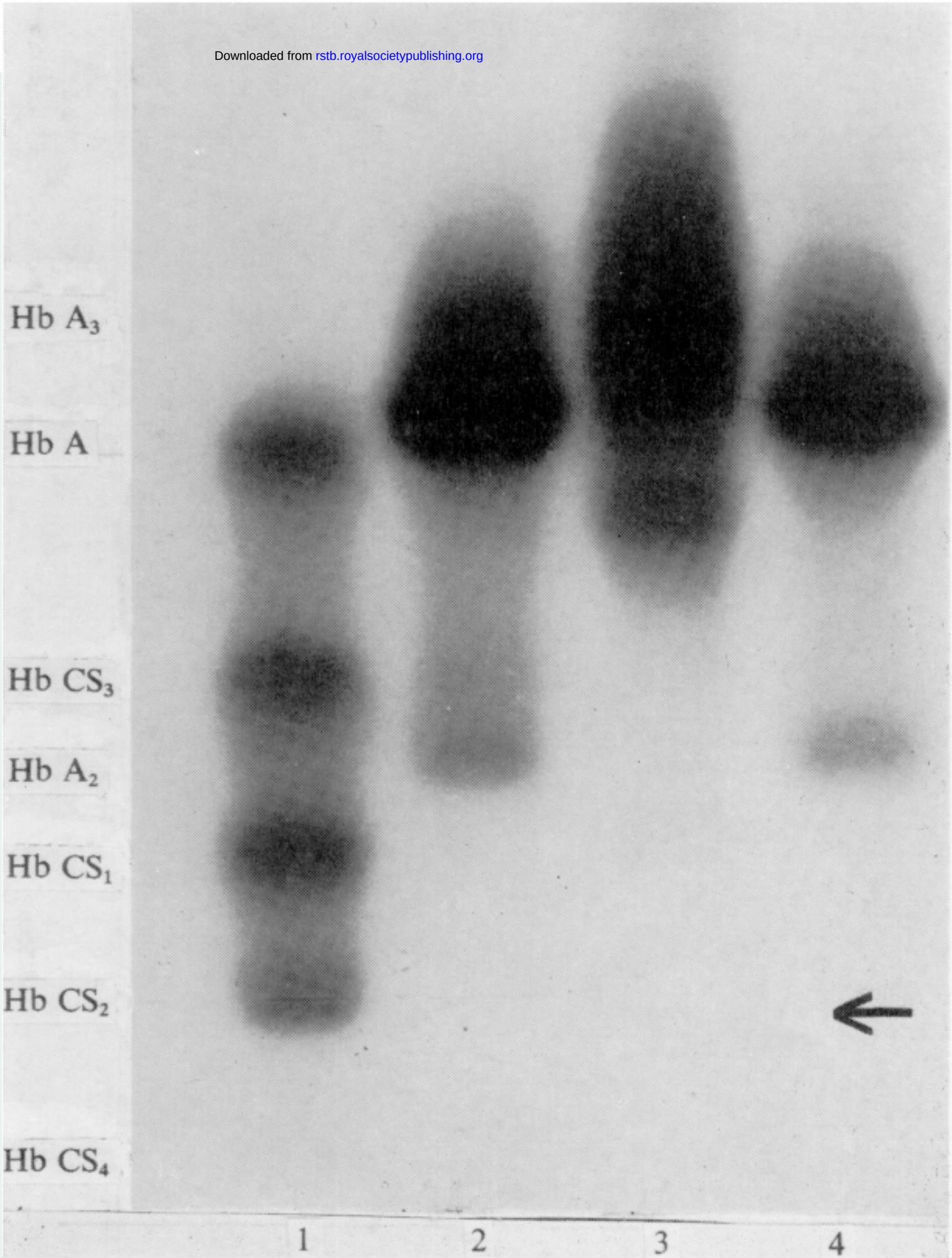


FIGURE 5. For description see opposite.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

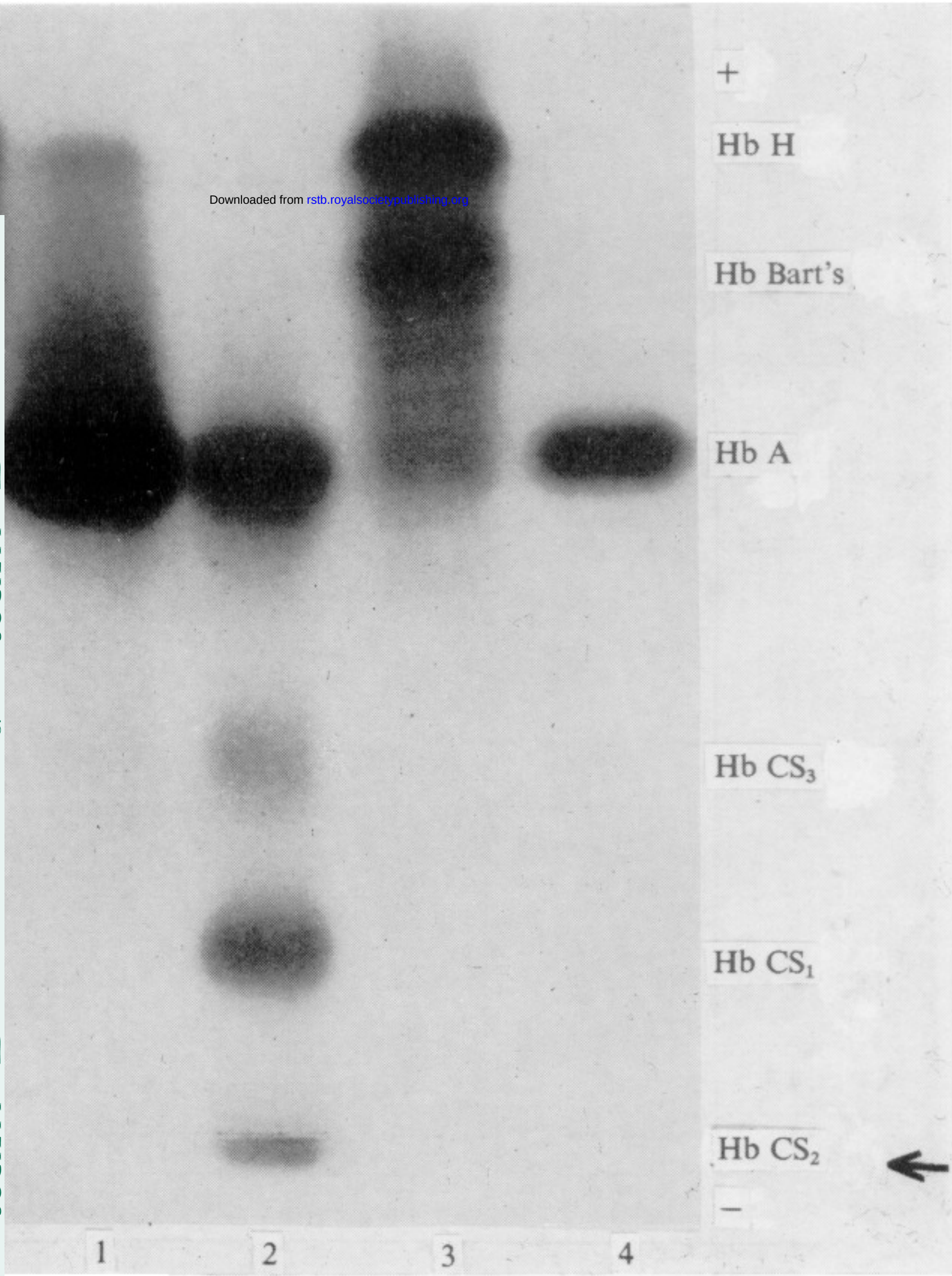


FIGURE 6. For description see opposite.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

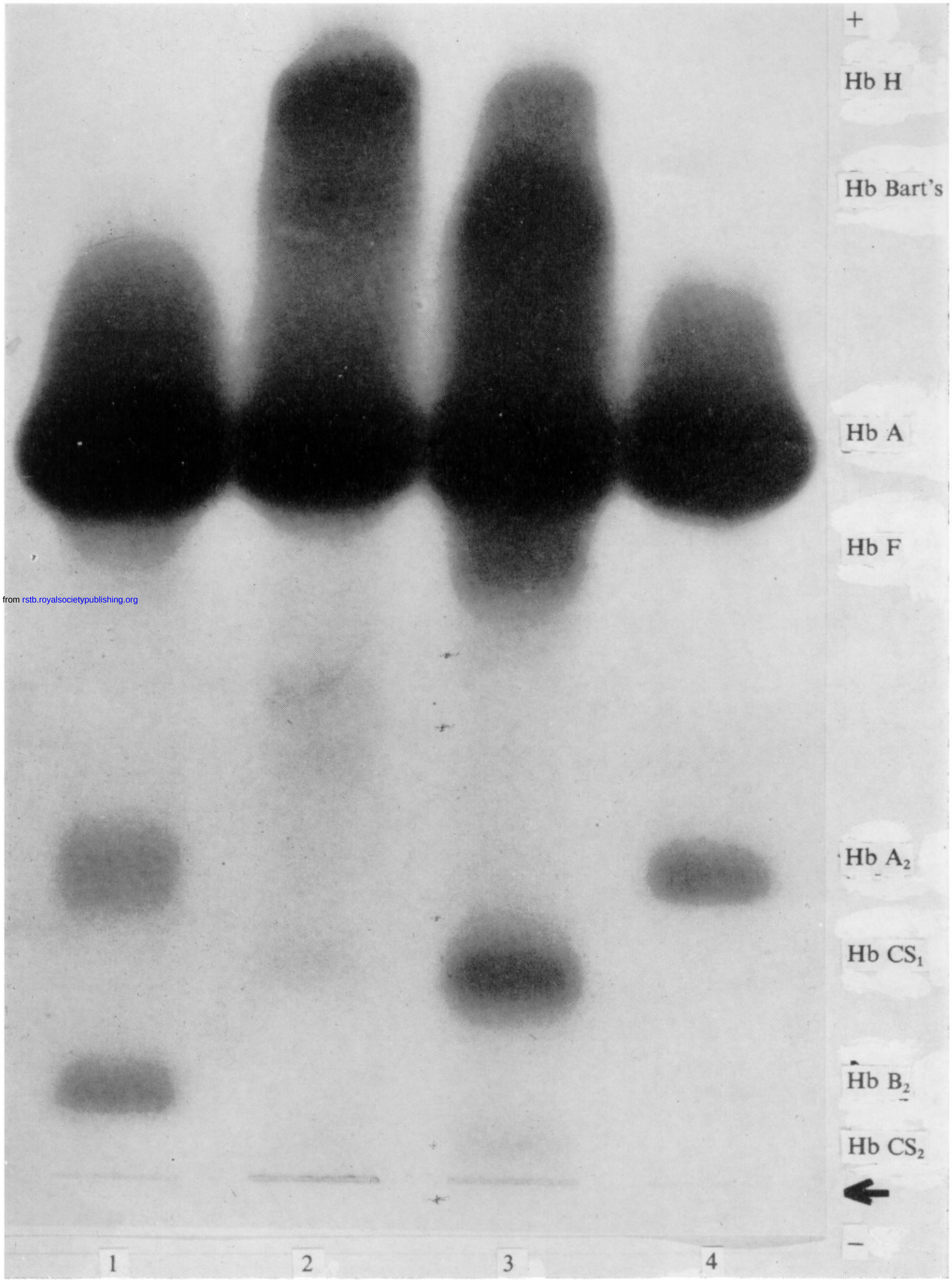


FIGURE 15. The electrophoretic pattern of Hb-CS-like fractions from different racial groups. Left to right: 1, a lysate from an individual heterozygous for Hb B_2, a δ-chain mutant, shown as a control; 2, a lysate from an individual heterozygous for α-thal. 1 and Hb CS from Thailand; 3, a lysate from an individual heterozygous for Hb CS and α-thal. 1 from Malaysia; 4, a normal adult lysate as a control. Vertical descending starch-gel electrophoresis, tris–EDTA–borate system, pH 8.5, benzidine stain.

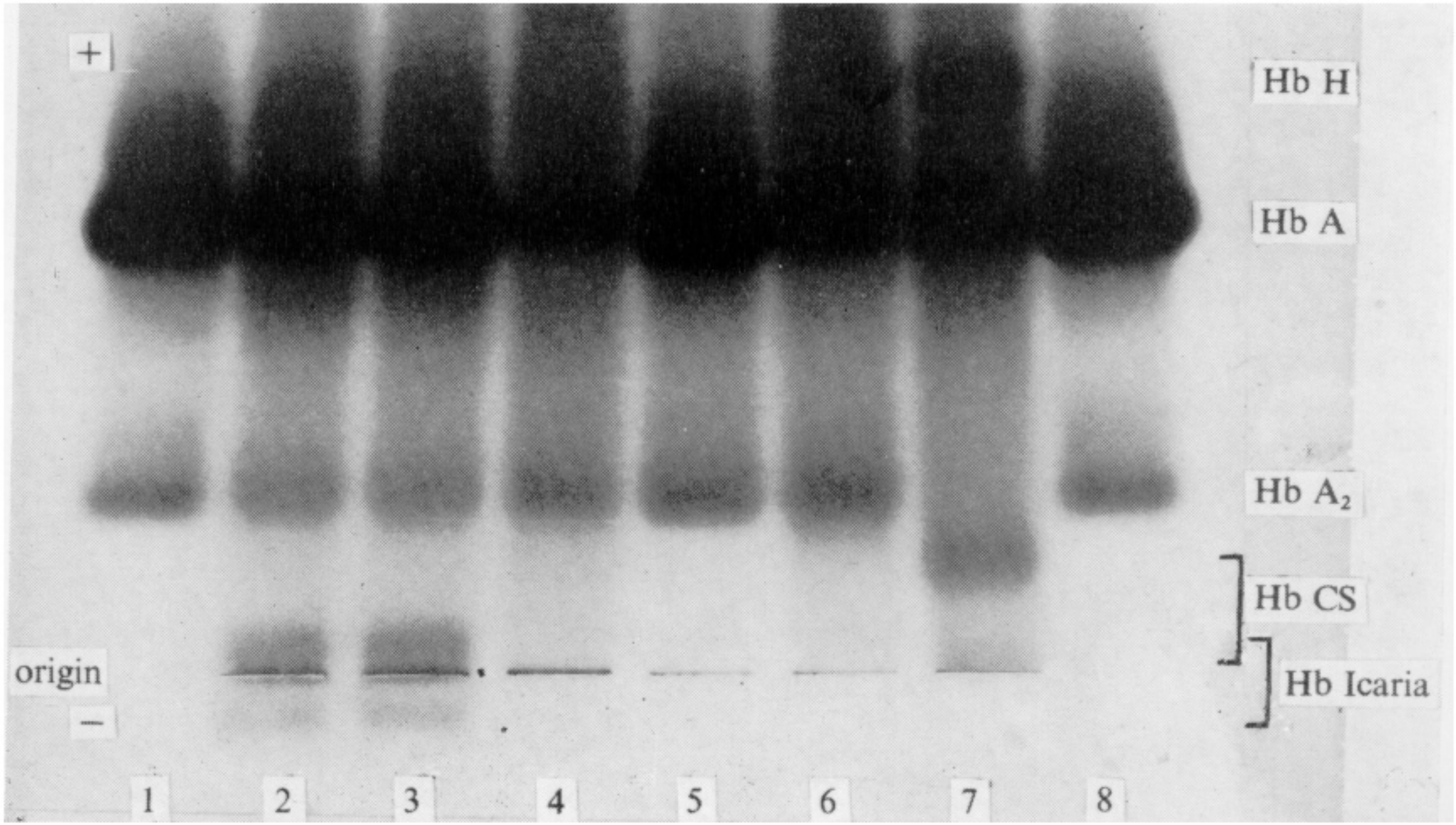


FIGURE 16. The electrophoretic characterization of Hb Icaria. The lysates are from the following sources (left to right): 1, normal adult; 2 and 3, Hb-Icaria heterozygotes; 4, Hb-CS heterozygote; 5, normal adult; 6, Hb-CS heterozygote; 7, an individual heterozygous for Hb CS and α-thal. 1; 8, a normal adult. Vertical descending starch-gel electrophoresis, tris–EDTA–borate system, benzidine stain.

(*Facing p.* 432