Biochimica et Biophysica Acta, 1000 (1989) 147–150
Elsevier

A case of sickle-cell anaemia *: a commentary by

Vernon M. Ingram

Department of Biology, Massachusetts Institute of Technology, Cambridge, MA (U.S.A.)

on 'Abnormal human haemoglobins. I. The comparison of normal human and sickle-cell haemoglobins by 'fingerprinting''
by V.M. Ingram
Biochim. Biophys Acta 28 (1958) 539–545
with 'II. The chymotryptic digestion of the trypsin-resistant 'core' of haemoglobins A and S'
by J.A. Hunt and V.M. Ingram
Biochim. Biophys. Acta 28 (1958) 546–549 (summary)
and 'III. The chemical difference between normal and sickle cell haemoglobins'
by V.M. Ingram
Biochim. Biophys Acta 36 (1959) 402–411 (summary)

The time, 1956, was right for working on the structure of an abnormal human haemoglobin, suspected to be the cause of sickle-cell anaemia. The disease was called a molecular disease by Pauling in 1949, the first time this notion had been applied, implying that an abnormal molecule was responsible for the symptoms. In that year, Linus Pauling, Harvey Itano, John Singer and Ibert Wells had shown that the haemoglobin in sickle-cell anaemia carries about three more positive charges per molecule than the normal adult protein, but it was not known what caused this difference. Also in 1949, James Neel had proved that sickle cell anaemia was due to homozygosity in a recessive autosomal gene. It remained to put these ideas together by discovering the precise nature of the chemical difference between wild-type and sickle-cell protein. This might prove that a gene controls the structure of a protein and might show how it does so. The surprise was in the fact that only a single amino acid was substituted in the protein by the sickle-cell allele, showing amazingly detailed control of the gene over the structure of its protein. It was also unexpected that so small a change could cause so serious a disease.

The time was right in 1956, because new tools for examining protein structure had just been devised. The methods used by Walter Schroeder and by Herbert Scheinberg in earlier attempts at quantitative amino acid analysis of the whole protein with its more than 100 amino acids were not sufficiently sensitive to show why the abnormal protein had a different electric charge near neutral pH. Even the new column chromatographic methods for amino-acid analysis developed by Stanford Moore and Bill Stein did not have the resolution necessary for what eventually turned out to be a single amino-acid difference. Clearly, it was necessary

Vernon Ingram

* The British spelling is employed in deference to the original publication.

0006-3002/89/$03.50

to cut the protein into smaller, more manageable pieces and examine these by other methods.

It was known that proteolytic enzymes, particularly trypsin, will cleave a protein cleanly into manageable peptides. In the case of haemoglobin some 30 peptides of about 10 amino acids each were expected, since it was shown that haemoglobin consists of two identical halves. Quite recently, a neighbour in Cambridge, Fred Sanger, had developed paper chromatographic methods and also paper electrophoretic techniques that could separate and characterize small peptides. It was a relatively simple step to combine the two techniques at right-angles and thereby separate a mixture of peptides from a protein on a single sheet of paper to produce a two-dimensional map. We called this a 'fingerprint', since the chemical structure of each peptide determined its location on the map and the collection of peptides produced a map characteristic of the structure of the protein. If two proteins differed in their structure in only a small segment, affecting the behaviour of the protein only a little, that same difference could make a large difference to the behaviour of the small peptide containing the difference. This proved to be the case in sickle-cell anaemia haemoglobin, where an octapeptide showed remarkably different electrophoretic and chromatographic behaviour, as well as the loss of negative charge, characteristic of the whole protein. Eventually it was shown that this peptide happened to be the N-terminal peptide, a fact of little significance.

The time was right also, because a few years earlier Pehr Edman had developed a chemical method for degrading a peptide one amino acid at a time in stepwise fashion, also identifying the particular amino acid at each step. John Sjöquist had made the method easier to use. The important feature of this procedure is that it can be used with very small quantities of peptide, such as became available from paper chromatography or paper electrophoresis. Nowadays, of course, HPLC-column methods are widely used to separate and prepare the peptide fragments and the Edman procedure has been completely automated so that its laborious steps can be performed while we sleep!

Haemoglobin was a favourite protein for structural biochemists, physical chemists and, of course, X-ray crystallographers. J.D. Bernal had shown in the 1930s that it was possible to get X-ray diffraction photographs from single protein crystals. The reasons for haemoglobin's popularity were not only its attractive colour, which is indicative of its important physiological role, but also the simple fact that blood was one of the easiest tissues to obtain from animals, even from humans. Red blood cells are the easiest cells to isolate and haemoglobin is by far the major constituent of these cells. It was easier to prepare in pure form than any other protein. These facts were not lost on Max Perutz, the X-ray crystallographer and director at the M.R.C. Unit in Cambridge, in fact my boss. (Mentor would be a better word.) By 1956 Perutz had worked on the structure of crystalline horse oxyhaemoglobin for a long time. Eventually he would solve its complete three-dimensional structure, as did John Kendrew for sperm-whale myoglobin. What strange creatures we turn to in search of the protein or the molecule that allows us to answer the questions of the moment with certainty and with least difficulty! Perutz was also very interested in sickle-cell haemoglobin, although I do not know the origin of his interest. By the time I arrived at the M.R.C. Unit, he had already published a paper in 'Nature' showing that deoxygenated sickle-cell haemoglobin was very insoluble and would form paracrystalline aggregates in solution. This was a hint towards explaining the bizarre 'sickle-shape' of red blood cells in the circulation of sickle-cell anaemics. Not only was my attention drawn towards haemoglobin in general and sickle cell haemoglobin in particular, Perutz was the reason why Tony Allison *, a virologist, visited the M.R.C. Unit during my stay there, bringing with him specimens of sickle cell haemoglobin. He wanted to take X-ray photographs of deoxygenated (aggregated) sickle-cell haemoglobin and thereby prove and characterize the aggregated state of the protein. However, Allison's experiments were unsuccessful for technical reasons, but he left behind samples of sickle-cell haemoglobin. Max and Francis Crick suggested to me that, as a protein chemist, I might want to apply some of the techniques of peptide fingerprinting that I had developed for another purpose to the problem of the difference between normal adult haemoglobin and sickle-cell haemoglobin. And so I did, with the help of my assistants, first Leslie Barnett and later Rita Fishpool. I remember that our first fingerprints looked like a modern watercolour left out in the rain! Then they improved and eventually gave the characteristic fingerprints of sickle-cell haemoglobin and normal haemoglobin. We felt that we had to do five different, unrelated samples of normal and five different, unrelated samples of sickle-cell haemoglobin to prove our point. It was fortunate that Herman Lehmann at St. Bartholomew's Hospital in London was willing to search his capacious freezer for the appropriate samples. That freezer had samples of haemoglobin from all over the world.

---

* Allison was a proponent of J.B.S. Haldane's notion that sickle-cell anaemia, although fatal when homozygous, was surprisingly common in many areas at a gene frequency of around 10%. Haldane sought to explain this fact by postulating balanced polymorphism as the cause, the heterozygote having an advantage over the normal homozygote. Allison thought that this might be due to partial resistance of the Hb^S heterozygote to the malaria that was or is endemic where sickle cell anaemia is common. Thanks to the work of Allison and many others, this is now the accepted explanation.

The fact that the DNA structure had just been solved by Francis Crick and Jim Watson in the room next door was certainly a powerful stimulus to the work and to the interpretation of the single amino-acid difference. The coding DNA sequence for haemoglobin was still a long way in the future, as was the genetic code. Nevertheless, the notion of a mutation replacing only one amino acid in a long peptide chain fitted well with the concept of a change in only a few bases in a long DNA base-pair sequence of the haemoglobin gene. In the event it would turn out to be a single base-pair change, which was as neat as could be.

Then followed a heady period in which my graduate students, John Hunt and Tony Stretton, and colleagues Makio Murayama and Seymour Benzer, and I applied our methods to other abnormal haemoglobins, C, E, D and I, as well as to the normal minor adult haemoglobin A_2. Other laboratories also began to apply the technique. Quite quickly the amino-acid substitutions were seen to be examples of allelism, hetero- and homozygosity, they illustrated co-dominance and evolutionary changes after gene duplication and unequal crossing over *, and eventually deletions and read-through mutations. The amino-acid substitutions in the abnormal human haemoglobins multiplied rapidly and played a part in the unravelling of the parameters of the genetic code. During this time I moved to the Massachusetts Institute of Technology in the other Cambridge.

The struggle towards the genetic code was at one stage concerned with the possibility of 'overlapping coding triplets of bases in DNA'. This was eventually ruled out in part by noting that certain amino-acid substitutions caused by single mutations would be forbidden. The human haemoglobin substitutions of valine in Hb^C and lysine in Hb^S, at the same position, were two of the instances then known that contradicted the overlapping code.

In the next few years it became apparent from our work and from work in other laboratories that other frequent and inherited haemoglobinopathies (HbC, HbE) showed also single amino-acid substitutions and a similar pattern of inheritance. Although much less serious diseases, their rather frequent occurrence still has no convincing explanation. Next, infrequent haemoglobinopathies were investigated in several laboratories by essentially similar methods and found to be single amino-acid substitutions **. It was only later that deletions and read-through mutations were found in some rare abnormal haemoglobins.

Only in the case of haemoglobin C could the demographic findings be used to postulate a single region in Ghana and Upper Volta as the origin of this mutation. Hence, a single mutational event was postulated. This had not been possible in the case of sickle-cell anaemia, because presumed population migrations seemed to have spread the mutated gene over very wide areas of Africa ***.

All the known abnormal haemoglobins could be shown to be inherited in a simple Mendelian manner. Sometimes two genes for different abnormal haemoglobins affecting both α- and β-chains were present in the same individual. Abnormal haemoglobins at that time were usually identified and characterized by their abnormal mobility on electrophoresis. They were named originally by capital letters, not logically according to mobility, but rather by the order in which they were discovered. As their numbers grew, it became necessary to add the place name of their first occurrence to the official name of the haemoglobin. The ability to pinpoint amino-acid substitutions in the various abnormal haemoglobins soon led to a classification of these mutations where the actual amino-acid substitution had occurred, e.g., $Hb\beta^{6Val}$.

The finding of the sickle-cell substitution made it important to define which of the two different peptide chains in the haemoglobin molecule was involved or whether it was perhaps both, since at that time the amino-acid sequence of the two chains was not known. It turned out that only the β-chain was affected. This was a relief for those of us who had adapted Beadle and Tatum's 'one gene – one enzyme' hypothesis to state 'one gene – one peptide chain'. The finding that a single mutation in a particular gene changes one amino acid in a peptide chain strengthened the concept of a direct relationship between the linear genetic material and the linear peptide chain. It certainly brought protein biochemists and geneticists closer together.

The existence of the abnormal human haemoglobins acted as a stimulus to protein chemists to define the peptide chains of all the human haemoglobins – the adult α-, β- and δ-chains, the foetal γ-chain, the embryonic ϵ- and ζ-chains. The work was spearheaded by Walter Schroeder at Cal. Tech. and by Lyman Craig, Bob Hill and Bill Konigsberg at the Rockefeller In-

* Corrado Baglioni in my laboratory showed unequal crossing over in the abnormal Hb^{Lepore}, which contains a non-alpha peptide chain that was part β- and part δ-chain.

** The rare abnormal haemoglobin $D^{Los\ Angeles}$ and the rare haemoglobin D^{Punjab} were found to have the identical amino-acid substitution. Was this a case of two independent mutations or was it due to a traveller?

*** At one time it seemed that the mutation rate for the sickle mutation was much higher than expected. A study showed two children among some 300 with sickle-cell anaemia whose mothers were not sickle-cell trait carriers, i.e., were not heterozygous for the mutation. This would indicate an extraordinarily high mutation rate. It later turned out that child exchanges were quite frequent in that particular population and were not necessarily officially documented.

stitute, but eventually it was taken up by others. Many years later the knowledge of these peptide chains was instrumental in unravelling the structure of the human (and animal) globin gene clusters, the 'α-chain' gene cluster, containing γ, α- and ζ-genes, and the 'β-chain' gene cluster, containing β-, δ- and ϵ-genes. They and certain animal globin gene clusters are still favourites for studying gene expression and for examining the factors that activate one or other globin gene to be expressed at different stages of development.

It was apparent to us quite early on that a mutation in an α-chain would affect haemoglobins in mid and late foetal as well as adult stages of human development, since the α-chain is common to all the haemoglobins in the late foetus and the adult. Tony Stretton and I sought to bring some order to the confusion that existed amongst the descriptions of the various types of thalassaemia. This is a widespread group of inherited haemoglobinopathies in which there are only normal haemoglobin peptide chains, but a gross deficiency in certain haemoglobins, or more precisely in one or other haemoglobin peptide chain. We postulated that the thalassaemias could be divided into α- or β-thalassaemias *. Since α-chains are common to all haemoglobins, a thalassaemic mutation affecting the production of α-chains would cause an anaemia at all stages of development; a β-thalassaemia would affect only adults, etc. Our understanding of the structural and genetic basis for these molecular diseases became gradually better. In the last 10 years, through the work of Bernard Forget, Arthur Bank, Stuart Levy, David Weatherall and others and with the advent of powerful methods of DNA cloning and DNA sequencing, the molecular biology of the thalassaemias has become clear. This understanding contributed greatly to our understanding of molecular genetics and of protein synthesis. Certain thalassaemias illustrate the effects of large and small deletions. Others affect the processing of immature messenger RNA for globins and thereby shed light on some steps in the synthetic process. At first sight the human organism might seem to be a poor prospect for the study of fine or gross molecular genetics, since planned crosses are hardly possible. In addition, the generation time is very long! Yet it was in the human system that the 'one gene – one peptide chain' notion and the idea of a mutation so small as to affect only a single amino acid was first seen. Charles Yanofsky's elegant and elaborate studies in the tryptophan synthetase gene of *E. coli* followed soon. His work demonstrated even more thoroughly these same principles in a bacterial system.

What was the effect of the original observation on the treatment of sickle-cell anaemia? Once, when I was very excited about the original finding, I was brought down to earth by the comment: "Very interesting, but what does it do for the patient?" What indeed ! Even today, 30 years later, the answer is – very little. There have been big improvements in the management of the sickle-cell patient to the degree that it is even possible for a woman with sickle-cell anaemia to bear a child, though not perhaps desirable from a population point of view. The improvements in treatment cannot be claimed by the molecular biologists, however; they belong to the clinicians. The molecular biologist can only contend that through their work a prenatal test of the foetus' DNA can reveal the presence of homozygosity for the sickle-cell mutation **. This makes it possible to advise the parents that a therapeutic abortion might be desirable. Perhaps at some distant future date the molecular biologist turned cell biologist will be able to repopulate the homozygote's bone marrow with his/her erythroid precursor cells carrying additional DNA with normal wild-type genes for the β-globin chain. Such a person might make enough of the wild-type β^{6Glu}-chain, i.e., enough normal haemoglobin, to prevent the sickling of these cells when they become deoxygenated in the tissues. The patient would then perhaps be no worse off than the largely asymptomatic heterozygote. This treatment would have the disadvantage that it would keep two Hb^S genes in the child-bearing population.

Molecular population genetics tells us that in the case of sickle-cell anaemia the removal of malaria from the population will lead to the eventual loss of the Hb^S gene, a process that takes many generations ***.

It is often difficult to go from the genetics of a mutation to finding the particular DNA sequences involved. The next step, to understand the role of these sequences in the etiology of the disease, is usually even harder, although we were lucky in the case of sickle-cell anaemia. However, to move from that point to proposing and perfecting a method of treatment is more difficult still. Yet these are the challenges facing molecular biologists in the realm of genetically determined diseases.

* The concept of β/δ thalassaemias came much later.

** The foetus at the time of amniocentesis does not yet synthesize the adult β^{6Val}-chain. Therefore, its haemoglobin is not yet characteristic of the disease.

*** Some 300 years ago, Dutch settlers brought many slaves from the malarial, sickle-cell anaemia region of the West Coast of Africa to the island of Curaçao and also to the mainland, including Surinam. Curaçao was and is free of malaria, the mainland of Surinam was swampland infested with malaria. In 1965, many generations later, J.H.P. Jonxis reported that the frequency of the Hb^S gene is only 6.5% on Curaçao, but 18.5% in Surinam. On the other hand the frequency of the Hb^C gene is the same, at 5.8 and 6.0%, respectively, indicating that the population in the two locations came from a single pool. The study shows that malaria increases the frequency of sickle-cell disease and that malaria has little if any influence on the fitness of carriers of the Hb^C gene.

Correspondence: V.M. Ingram, Dpartment of Biology, Room 56-601, Massachusetts Institute of Technology, Cambrdge, MA 02139, U.S.A.

ABNORMAL HUMAN HAEMOGLOBINS

I. THE COMPARISON OF NORMAL HUMAN AND SICKLE-CELL HAEMOGLOBINS BY "FINGERPRINTING"

V. M. INGRAM

Medical Research Council Unit for Molecular Biology, Cavendish Laboratory, University of Cambridge (Great Britain)

NEEL[1] demonstrated in 1949 that sickle-cell anaemia is a disease inherited in a strictly Mendelian manner, and PAULING *et al.*[2], showed in the same year that the blood of sickle-cell patients contains a haemoglobin which differs electrophoretically from normal haemoglobin. Since then, several other abnormal haemoglobins have been discovered. They have been studied intensively by many workers for the sake of both their medical and their genetic interest, because they provide one of the clearest cases where mutations in a Mendelian gene are reflected in changes in the chemical structure of a protein. The medical aspect is important, because several of the abnormal haemoglobins give rise to severe haemolytic anaemias, which thus provide one of the rare instances when a disease is directly attributable to an inherited change in a protein molecule. However, in this series of papers it will be the genetic and biochemical aspect with which we shall be primarily concerned. Like other proteins, haemoglobin is thought to be the first protein product of the gene. The sequence and spacing of the mutations on the gene responsible for its synthesis should therefore be reflected in the sequence and spacing of the changes in amino-acid composition which these mutations produce in the haemoglobin.

The first three papers of this series will contain a comparative study of the amino-acid composition of sickle-cell and normal haemoglobin. The results have already been described in two preliminary notes[5,8]. In this fuller account the experimental work will be described in detail.

Chemical studies on haemoglobins A (normal) and S (sickle-cell) by previous authors[5] may be summarised by saying, first, that whatever difference exists between them must be small, and secondly, that the known charge difference of two to three units per molecule[2] is probably due to haemoglobin S having fewer carboxyl groups[6]. When it is remembered that human haemoglobin is made up of two identical halves[7], this difference is probably one of a single carboxyl group in the half molecule. It was not known whether the two haemoglobins differed in amino acid sequence or on the other hand, in the folding of the polypeptide chains leading to masking of some groups. As a result of the present work we can now say for certain that there is a difference in amino acid composition and that this does involve a carboxyl group, and explain the known electrophoretic behaviour. In fact, the polypeptide chains of the two haemoglobins differ in only one amino acid residue; one of the glutamic acid residues of haemoglobin A had changed to valine in haemoglobin S[5,8]. No other

References p. 545.

changes could be detected. The detailed evidence which led to these conclusions will now be described.

Part I will deal with the comparison of tryptic digests of the two haemoglobins and Part II, the paper published simultaneously[13], with the examination of the trypsin-resistant "cores". Part III will describe the chemical studies on the differing peptides derived from the two proteins.

MATERIALS AND METHODS

The blood samples which reached us had usually been collected in citrate–dextrose solution. The red cells had been sedimented by centrifuging at low speed and washed 4–6 times, sedimenting each time from about 4 volumes of 0.9% NaCl. The samples of normal European blood were worked up in the same manner in our laboratory. Sickle-cell red cells from patients homozygous for the sickle-cell gene, which had already been washed before they arrived, were given a few more washes. The cells were lysed by freezing and thawing; the cell ghosts were removed by centrifugation at 20,000 × *g* for 20 min. The clear solution of haemoglobin—usually 10 ml—was dialysed for 3 days against 3 × 1 l of cold distilled water containing 1 ml of 1 *M* PO_4 of pH 7.2 per l. This dialysis was followed by another high-speed centrifugation; the samples were stored frozen. The early experiments were performed on a sample of sickle-cell haemoglobin kindly supplied in pure solution by Dr. A. C. ALLISON of Oxford.

The purity of these haemoglobin solutions was checked by paper electrophoresis at pH 7.2 in 0.1 *M* PO_4 buffer and at pH 8.6 in 0.06 *M* Veronal buffer. Whatman 3MM was used as a free hanging strip and 1–2 mg of haemoglobin were applied along a 1″ line on the paper. All samples gave clear single bands, though a small amount of slightly faster material was visible in haemoglobin S. This was probably foetal haemoglobin, but the quantities present were not likely to be large enough to interfere with the examination of the trypsin digests of these proteins. Experience has borne this out. Paper electrophoresis of haemoglobin samples at pH 6.4 in the pyridine–acetic acid–water buffer of MICHL[9] showed only traces of 3–4 ninhydrin-positive substances which were the same in both haemoglobin A and S.

In order to render haemoglobin susceptible to trypsin attack, it first had to be denatured. A 2% haemoglobin solution was brought to pH 8.0 with dilute NaOH and heated in a narrow test tube in a water bath at 90° for 4 min. The sample was then immediately cooled. A fine light brown suspension was obtained; this showed little tendency to settle and trypsin attacked it readily.

The trypsin used for these digestions was the 2 × crystallised, salt-free protein obtained from Worthington Biochemicals, Freehold, New Jersey, U.S.A. It was used without further purification. The first trypsin digestions were carried out in dilute phosphate buffer under the following conditions: Haemoglobin concentration 2% or 4% in 4.0 ml; trypsin to haemoglobin ratio, by weight, 1:50; 0.05 *M* Na_2HPO_4–HCl buffer pH 8.1; 38°. Two drops of toluene were added to prevent contamination. The incubation of the suspension lasted for 43–48 h, with occasional mixing. As the haemoglobin had been denatured in the presence of the buffer, the precipitate was much coarser and settled easily, unlike the preparation described in the previous paragraph. As the digestion proceeded, a lot of the precipitate disappeared, but a dark brown residue remained undissolved, presumably containing all the haeme groups. The supernatant had little colour. Surprisingly this method of digestion gave reproducible peptide mixtures. It was used for the early "fingerprints"[5].

Since then improved methods of trypsin digestion have been developed and all later digests were carried out in a type of "pH-stat". 4-ml samples of a 2.2% haemoglobin solution, denatured by heating in salt-free solution at pH 8, were stirred continuously at 38° in an atmosphere of nitrogen; trypsin 0.10 ml of a 0.5% solution in 0.001 *M* HCl was added at zero time. The vessel contained a glass electrode and an agar–KCl bridge to a calomel electrode. The splitting of peptide bonds liberated hydrogen ions; 0.5 *N* NaOH was added manually by means of an Agla micrometer burette to keep the pH between 7.95 and 8.00 using an EIL direct reading pH meter. It is well known that when alkali consumed is plotted against time the slope of this curve never flattens completely. In preliminary experiments, the end point was determined by adding further aliquots of trypsin solution and observing that after a certain time—about 90 min—there was no additional alkali uptake (Fig. 1). After the digestion, the pH was adjusted to 6.5 with 1 *N* HCl and the precipitate removed by centrifugation. The supernatant was stored frozen; it was used for fingerprinting and for the preparation of various peptides. Thus the improved procedure reduced digestion time from about 40 h to less than 2. The purification and chymotryptic digestion of the trypsin-resistant "core" is described in part II of this series.

References p. 545.

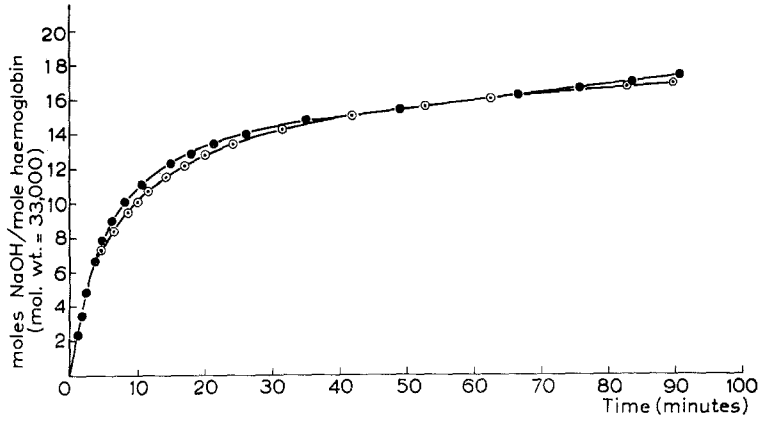

Fig. 1. Course of the trypsin digestion of haemoglobin A and S. ⊙—⊙ Haemoglobin A; ●—● Haemoglobin S.

"Fingerprinting"

The separation of the peptides obtained by tryptic digestion was achieved by a two-dimensional combination of paper electrophoresis and chromatography. Paper electrophoresis was carried out on Whatman No. 3MM paper which had been dipped into MICHL's volatile buffer[9]: pyridine–glacial acetic acid–water, pH 6.4, in the proportions 10: 0.4:90 by volume. AnalaR grade chemicals were used without further purification. Excess liquid was removed carefully by blotting firmly between two sheets of fresh blotting paper. The moist No. 3MM paper was immediately placed onto a horizontal piece of $^1/_4$″ polished plate glass which rested on the two buffer vessels, each 3″ high. The sample of peptides—0.10 ml of the original solution which had been evaporated and dissolved in 0.020 ml of water—was applied with a construction pipette at the point indicated in Fig. 2. Similarly, a second peptide mixture could be applied to the other paper. At once, a second sheet of plate glass was placed on top of the first one, thus enclosing the moist filter paper as a sandwich. Fig. 2 shows a plan of the arrangement and in particular the way the paper is cut.

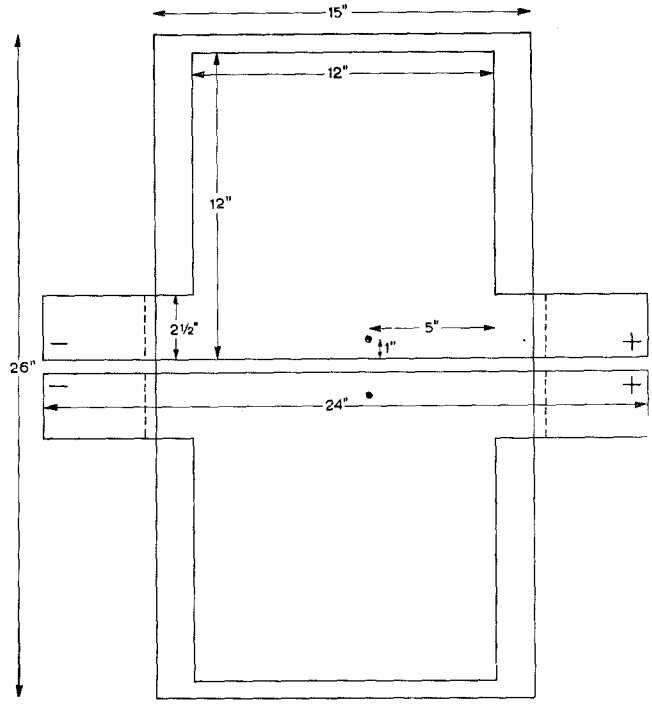

Fig. 2. Dimensions of filter paper used for fingerprinting.

References p. 545.

The tags dip into buffer vessels (not shown). This arrangement allows two separations to be carried out under identical conditions; it is now generally used to compare peptide mixtures from normal haemoglobins with those from the abnormal proteins. Great care is taken to handle pairs of papers in the same way. Preferably they are cut from a single sheet of paper, 24″ × 24″, with the machine direction of the paper always pointing in the same direction.

After the sample had been applied, 15 min were allowed to elapse to give the buffer time to diffuse into the sample, even though the latter was already at pH 6.5. A potential of 19 V/cm was applied, and as the current rose, this fell to about 14 V/cm after 150 min. The glass sandwich got comfortably warm during the electrophoresis. The tags were trimmed off the papers and the buffer was removed by drying at room temperature in a current of air for not less than two hours and preferably longer. The dry papers, now 12″ × 12″, were hung for two hours in the air space above the butanol–acetic acid chromatographic solvent. This was made up from AnalaR grade chemicals in the proportions *n*-butanol–glacial acetic acid–water, 3:1:1 by volume; it was freshly prepared each Saturday for use in the next week beginning on Monday.

Ascending chromatography was carried out overnight, usually for 15 h. The papers were again dried in a current of air at room temperature. The peptide spots were revealed by dipping the papers into 0.2% ninhydrin in acetone and allowing development to take place at room temperature or in a warm place. This usually took 24 h to reach maximum intensity. Since it was difficult to preserve these "finger prints" even with the copper spray[10], reflex prints were taken on document paper or colour transparencies were made.

One-dimensional paper electrophoresis on 3MM paper in the above volatile pH 6.4 buffer was frequently used for examining peptide or amino acid mixtures and for isolating peptides. The arrangement was a glass sandwich similar to the fingerprinting apparatus but measuring only 12″ × 6″. Another volatile buffer[9] at pH 3.6: pyridine–glacial acetic acid–water, 1:10:90 by volume, was useful for one-dimensional separations. For these, the bottom member of the glass sandwich was replaced by a copper plate $^1/_4$″ thick, cooled by copper coils carrying tap water. A thin layer of silicone greases and another of thick polythene film separated the copper plate from the paper. Up to 35 V/cm could be used without noticeable rise in the temperature of the upper glass plate. When the isolation of a peptide was required, the papers were usually first washed by descending irrigation with the pH 6.4 buffer, followed by soaking in water and drying. Such papers were developed by spraying with 0.025% ninhydrin in butanol and gentle heating. Alternatively guide strips were used. The peptides were eluted with water or 20% aqueous acetic acid. Descending chromatography on No. 1 paper *n*-butanol–glacial acetic acid–water, 4:1:5, was used for separating peptide mixtures, especially those which are uncharged at pH 6.4. Amino acids were identified by the two-dimensional system of REDFIELD[11] using Whatman No. 1 paper, 20 × 20 cm. A multiframe holding 12 papers could be processed in about 7 h. After steaming for 10 min and dipping into 0.2% ninhydrin in acetone, the papers were developed at room temperature. Good separations were obtained when the solvents were made up each day and were kept at the fairly constant temperature of a basement room. Occasionally amino acids were separated by ascending chromatography on No. 1 paper butanol–acetic acid–water solvent, 3:1:1, overnight.

RESULTS AND DISCUSSIONS

Comparison of haemoglobin A and S tryptic digests

The most striking feature of the fingerprints of haemoglobin A and S (Fig. 3) is the faithful reproduction of the intricate and highly characteristic pattern of peptide spots in both proteins. All except one; this is the No. 4 peptide which appears amongst the uncharged peptides of a haemoglobin A fingerprint, but is positively charged in haemoglobin S and appears in a new position nearer to the cathode. Four questions arose: is such a fingerprint characteristic of all individuals of a certain genotype; is it a true representation of all the peptides in a trypsin digest; are all the peptides, which occupy similar positions really identical in structure, and finally what is the difference between the haemoglobin A and S No. 4 peptides?

The haemoglobins of five different normal white persons and of five different negro sickle-cell patients have been examined by the method of trypsin digestion followed by separation of the peptides by the "fingerprinting" technique. No differences in behaviour were found within either group. The comparison with normal negro haemoglobin has yet to be done, but it is known that electrophoretically it is like

References p. 545.

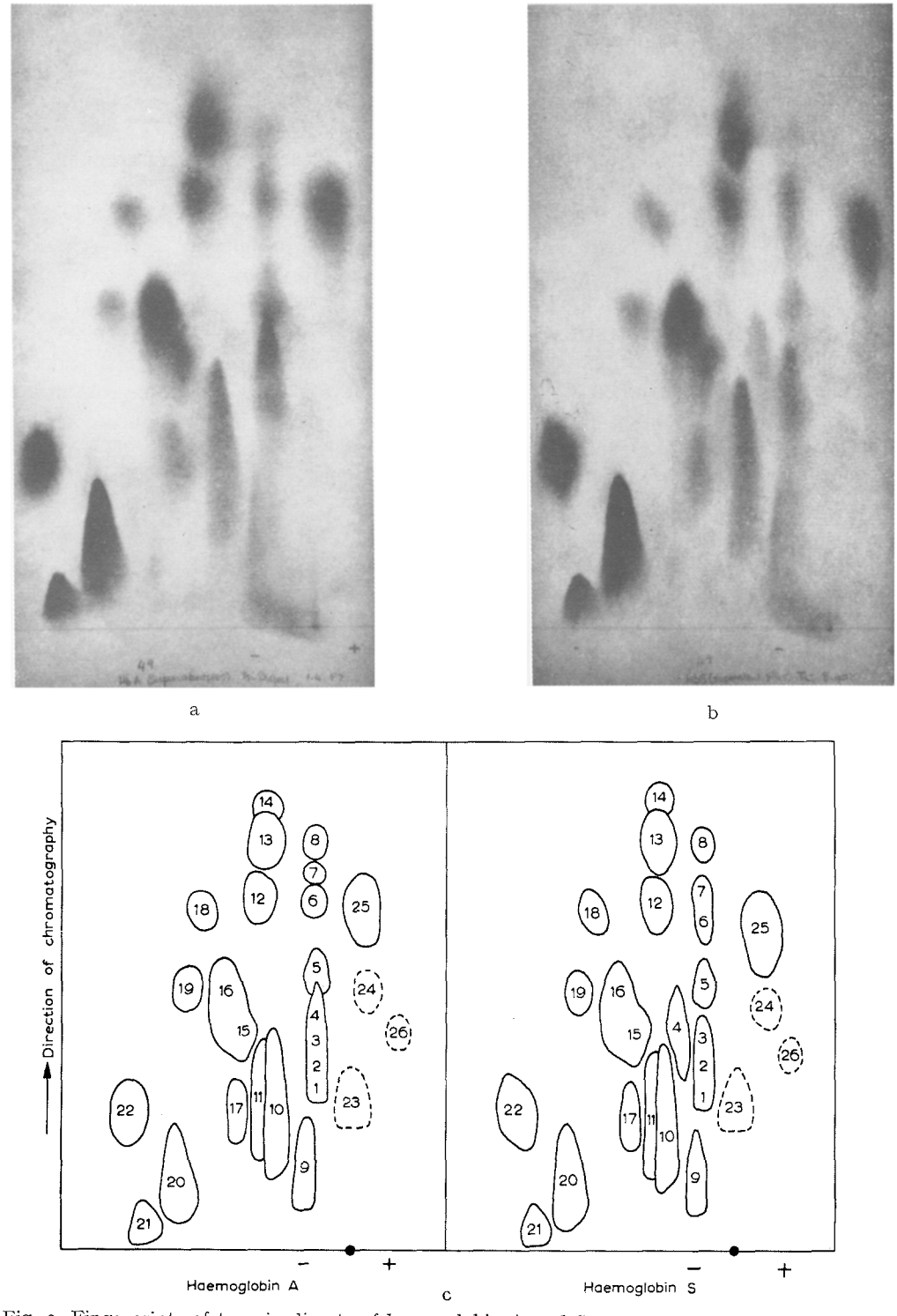

Fig. 3. Fingerprints of trypsin digests of haemoglobin A and S. a, Haemoglobin A; b, Haemoglobin S; c, Tracings, showing the system for numbering peptide pairs. Dotted lines indicate peptides which only became visible after heating the chromatogram.

normal European haemoglobin A. The fingerprint contains altogether some 26 peptides, though, as the numbering in the figure shows, not all of them are separated. Their further separation which will be discussed below, has provided good evidence that the tracing in Fig. 3 shows all the peptides present in the tryptic digest with the exception of the trypsin-resistant core which is discussed in Part II. As judged by the intensity of their ninhydrin colours and those of their constituent amino acids, the majority of peptides are present in good yield. Earlier "fingerprints"[5] had been prepared from 43-h trypsin digests in buffered solutions, as compared with the 90-min incubation periods now used and the reduction in the proportion of trypsin to a quarter of the old value. The earlier fingerprints showed thirty peptide spots which may have corresponded to some thirty-two or -three peptides. The extra peptides had probably arisen through traces of chymotrypsin which were present in the trypsin and which got a chance to act because the period of digestion was too long and the enzyme concentration too high. The present fingerprints show 18–19 lysine-containing and 6 arginine-containing peptides. The trypsin-resistant core probably contains another 3 lysine peptides, but no arginine. These figures compare very well with the latest data given by STEIN *et al.*[12] for normal haemoglobins, namely about 21 lysines and 6 arginines per half molecule of haemoglobin.

It was particularly necessary to examine the poorly resolved "backbone" of neutral peptides. They were prepared by elution of the neutral band after one-dimensional paper electrophoresis at pH 6.4; this fraction was then examined by descending paper chromatography in butanol–acetic acid–water, 4:1:5, and revealed a number of peptides which were identical in both proteins except for peptide No. 4[5]. A better way of examining peptides 1–4 was by paper electrophoresis at pH 3.6 where they separated well. One can therefore be reasonably certain now that the fingerprint shows all the trypsin peptides.

Duplicate fingerprints of haemoglobin A and S when carefully prepared, are sometimes so good that individual peptide spots are superimposable. This is an indication, but no more, that any peptide from haemoglobin A will have the same structure as the corresponding S peptide. Individual peptides from the two haemoglobins have been compared on the same paper, by extended paper chromatography with butanol–acetic acid or by paper electrophoresis at pH 3.6. In all cases they have formed identical patterns and their ninhydrin colours relative to other peptides have been closely similar. In addition, analogous peptides have been eluted, purified further if necessary, hydrolysed completely and examined qualitatively for their amino acid contents. This procedure does not detect tryptophan or cysteine; neither does it distinguish between, say, 3 or 4 amino acids of the same kind in one peptide, though it would detect a difference between one and two. Careful comparison of pairs of these amino acid chromatograms has failed to show any difference between analogous peptides, except for the two No. 4 peptides. Here the loss of one of two glutamic acids was obvious, as was the gain of the extra valine in the haemoglobin S peptide. With the above reservations in mind, it is possible to say that probably all trypsin peptides, except No. 4, have similar chemical structures. Similar arguments hold for all peptides obtained by the chymotryptic digestion of the "core".

The determination of the detailed chemical structure of the No. 4 peptides from haemoglobins A and S will be described in Part III. The results described here and by HUNT AND INGRAM[13], lead one to conclude with reasonable confidence that most of

References p. 545.

the amino acid sequences of the two haemoglobins are very similar and perhaps identical, and that the only difference lies in the No. 4 peptide. It does not follow that there are not additional differences in the folding of the polypeptide chains, but they must be small, since they are not detectable in the X-ray diffraction patterns from single crystals of these proteins[7]. The genetic implications of these findings have already been discussed[5,8].

ACKNOWLEDGEMENTS

The author is greatly indebted to Dr. A. C. ALLISON, Oxford, and to Drs. S. BRENNER AND G. SEAMAN, Cambridge, for supplying blood from patients with sickle-cell anaemia. The great interest shown by Drs. M. F. PERUTZ AND F. H. C. CRICK, is gratefully acknowledged, as are many helpful discussions with Dr. H. LEHMAN, St. Bartholomew's Hospital, London. Grateful thanks are also due to Mrs. RITA FISHPOOL for her valuable assistance.

SUMMARY

Samples of human haemoglobin from normal adults and from sickle-cell anaemic patients have been heat denatured and then digested with trypsin. The resultant peptide mixtures were separated in two dimensions by paper electrophoresis and paper chromatography. Such "fingerprints" show that the two proteins differ in only one peptide.

REFERENCES

1 J. V. NEEL, *Science*, 110 (1949) 64.
2 L. PAULING, H. A. ITANO, S. J. SINGER AND I. C. WELLS, *Science*, 110 (1949) 543.
3 H. A. ITANO, *Ann. Rev. Biochem.*, 25 (1956) 331.
4 W. W. ZUELZER, J. V. NEEL AND A. R. ROBINSON, in *Progress in Hematology*, Vol. I, Grune & Stratton, New York, 1956.
5 V. M. INGRAM, *Nature*, 178 (1956) 792.
6 I. H. SCHEINBERG, R. S. HARRIS AND J. L. SPITZER, *Proc. Natl. Acad. Sci. U.S.*, 40 (1954) 777.
7 M. F. PERUTZ, A. M. LIQUORI AND F. ERICH, *Nature*, 167 (1951) 929.
8 V. M. INGRAM, *Nature*, 180 (1957) 326.
9 H. MICHL, *Monatsh. Chem.*, 82 (1951) 489.
10 E. KAWERAU, *Biochem. J.*, 48 (1951) 281.
11 R. R. REDFIELD, *Biochim. Biophys. Acta*, 10 (1953) 344.
12 W. H. STEIN, H. G. KUNKEL, D. COLE, D. M. SPACKMAN, S. MOORE, *Biochim. Biophys. Acta*, 24 (1957) 640.
13 J. A. HUNT AND V. M. INGRAM, *Biochim. Biophys. Acta*, 28 (1958) 546.

Received November 27th, 1957

ABNORMAL HUMAN HAEMOGLOBINS

II. THE CHYMOTRYPTIC DIGESTION OF THE TRYPSIN-RESISTANT "CORE" OF HAEMOGLOBINS A AND S

J. A. HUNT AND V. M. INGRAM

Medical Research Council Unit for Molecular Biology, Cavendish Laboratory, University of Cambridge (Great Britain)

SUMMARY

No differences were detected between the "trypsin-resistant cores" of haemoglobins A and S. The core accounts for a considerable portion of the molecule, poor in lysine and arginine, but rich in aromatic amino acids and those with the long non-polar side-chains.

ABNORMAL HUMAN HAEMOGLOBINS. III THE CHEMICAL DIFFERENCE BETWEEN NORMAL AND SICKLE CELL HAEMOGLOBINS

V. M. INGRAM*

Molecular Biology Unit of the Medical Research Council, Cavendish Laboratory, University of Cambridge (England)

(Received September 20th, 1958, revised manuscript received July 13th, 1959)

SUMMARY

Tryptic digests of normal human and sickle cell haemoglobins contain a peptide fragment (peptide 4) which apparently alone of all the peptides has a different chemical structure in the two haemoglobins. This peptide has been degraded and its amino acid sequences have been determined in the two cases. The sole alteration is the replacement of a glutamic acid residue of normal haemoglobin by valine in sickle cell haemoglobin.