Basic Biology Course

9 Protein Synthesis

BASIC BIOLOGY COURSE
UNIT 3
REGULATION WITHIN CELLS

BOOK 9

Protein Synthesis

MICHAEL A. TRIBE, IRWIN TALLAN, MICHAEL R. ERAUT &
ROGER K. SNOOK
University of Sussex

CAMBRIDGE UNIVERSITY PRESS

CAMBRIDGE
LONDON · NEW YORK · MELBOURNE

Foreword

This book is part of a Basic Biology Course for undergraduates written by the Inter University Biology Teaching Project team at Sussex. The trial version of the book has been extensively re-written in collaboration with Dr I. Tallan (on leave from the University of Toronto, 1974–5).

The 'central dogma', a term proposed by Crick, is fundamental to all aspects of biology. The central dogma emphasizes the basic relationship between DNA, RNA and protein; with DNA serving as a template for both its own duplication and the synthesis of RNA; and RNA, in turn, the template for protein synthesis. An understanding of these basic relationships is the main aim of this book.

We have not attempted to provide a comprehensive survey of the research work going on in this field, because that has been the task of other books and is not the purpose of this. Instead, we have tried to concentrate here on specific examples to provide evidence for the theme of our story. It is for this reason we have looked particularly at haemoglobin synthesis, because firstly, haemoglobin is a protein which is familiar to many students; secondly, its ready availability from red blood cells and its uniqueness of function make it very suitable for illustrating the principles of protein synthesis; and thirdly, and very importantly, it is also a protein on which much biophysical and biochemical research has been done.

Brighton, Sussex, 1975

M.A. Tribe
I. Tallan
M.R. Eraut
R.K. Snook

Acknowledgements

This book was developed under the auspices of the Inter University Biology Teaching Project and is the responsibility of the Sussex University Project team. However, it owes a great deal to the students who studied and criticized our earlier versions and to many colleagues, both at Sussex and elsewhere, who made constructive suggestions for its improvement.

In particular we would like to thank:

Dr B. Hogan, Dr D. Peacock, Dr J. Sampson and the late Professor A. Korner for help and encouragement in the early stages;

the Nuffield Foundation for financially supporting the project from 1969–72;

Cambridge University Press for their continued interest and support in publishing the materials;

and Mrs S.A. Collier, project secretary, for typing the manuscript.

We are extremely grateful to the following for allowing us to use their photographs:

M. Ashburner, Dept of Genetics, University of Cambridge (fig. 36)

W. Beerman, Max-Planck Institut für Biologie, Tübingen (figs. 34, 35)

D.W. Fawcett, Dept of Anatomy, Harvard Medical School (figs. 26, 33)

C. Pelling, Max-Planck Institut für Biologie, Tübingen (fig. 10)

A. Rich, Dept of Biophysics, Massachusetts Institute of Technology (figs. 9, 22)

J. Sampson, Imperial Cancer Research Fund, Mill Hill, London (figs. 24, 25)

H.S. Slayter, Dept of Biophysics, Massachusetts Institute of Technology (figs. 9, 22)

Contents

Introduction

Discussion

Previous books (5, 6, 7 and 8) in this Unit have stressed the importance of enzymes in regulating metabolic activities within cells. But little has been said about how or where the enzymes themselves are synthesized; or about the way or ways in which the cell is able to regulate the synthesis of enzymes.

This aspect of regulation is especially important because hundreds of reactions go on within each living cell and most of these reactions are enzyme-catalysed. Not only do all these enzymes have to be synthesized, but their synthesis has to be controlled. There is considerable variation from moment to moment in a given cell type and it would be extremely inefficient if the cell continuously made all the enzymes which it was capable of synthesizing, whether or not such enzymes were required. Moreover, all cells of an organism arise from pre-existing cells by cell division, starting with the fertilized egg. So the fertilized egg has to have the potential for making all the enzymes which characterize a particular organism. Some of these enzymes will be common to all cells (they all respire and excrete) but others will be specific to a particular organ. Thus the fertilized egg has to have the potential for producing both the specific set of enzymes which directs the development and function of, say, a liver cell and the rather different set of enzymes which directs the development and function of a brain cell or a skin cell, etc. How then can it possibly control its 'output' of genetic information in order to regulate subsequent development?

We know that cells are extremely efficient and do possess the ability to regulate their enzyme production, but precisely how this regulation is achieved is far from clear. Although geneticists, biochemists and cell biologists have all made valuable contributions to these fundamental problems of molecular biology, convincing experimental evidence is still relatively scant. There is sufficient evidence, however, to encourage a large number of ideas, hypotheses and models and it is our purpose to examine some of these here.

Preknowledge requirements

This book assumes a knowledge of protein structure, atomic structure and valency of common elements, bond formation, enzymic activity and standard free energy change (ΔG°), as given in Book 7 (*Enzymes*). It also assumes a knowledge of mitosis, meiosis, fertilization, mutation and the genetic consequences of these events – as given in Book 3 (*Dynamic Aspects of Cells*).

Many of the experiments quoted in the book make reference to globin, the protein constituent of haemoglobin, which you should know is responsible for carrying oxygen in the red blood cells (erythrocytes) of mammals and many other animals.

Objectives

At the end of this book you should be able to:

(1) Explain in outline how the three-dimensional molecular structure of a protein is synthesized from twenty different amino acids.

(2) Describe in outline the molecular structure of DNA and RNA.

(3) Demonstrate an understanding of the genetic code by correctly solving problems related to base pairing, and by using the table of codons, amino acid sequencing.
(4) Specify where and how protein synthesis takes place in the cell.
(5) Recall the similarities and differences between the types of RNA that are involved in translation.
(6) Recall experimental evidence supporting:
  (i) the nature of the message;
  (ii) transcription of the message;
  (iii) identification of the messenger;
  (iv) translation of the message, particularly for one specific protein, haemoglobin.
(7) Explain, by giving specific examples, how mutations in genetic material can induce changes in protein structure and function.
(8) Analyse and interpret correctly experimental data concerned with protein synthesis from, or similar to that found in, this book and presented in tabulated, graphical or micrograph form.
(9) Demonstrate an understanding of the operon hypothesis to show how the synthesis of proteins might be regulated.
(10) Define and use all terms in the Glossary, except specific chemical formulae.

Instructions on working through programmed text

All this book is programmed, with questions and answers arranged sequentially down the page. You are provided with a masking card and a student response booklet. Cover each page in turn, and move the masking card down to reveal two thin lines

---

This marks the end of the first question on that page. Record your answer to the question under the appropriate section heading in the response booklet provided. Then *check* your answer with the answer given. If your answer is correct, move the masking card down the page to the next double line and so on. If any of your answers are incorrect retrace your steps and try to find out why you answered incorrectly. If you are still unable to understand the point of a given question, make a note of it and consult your tutor. The single thick line

---

is a demarcation between one frame and the next. '*Intermissions*' in this book are convenient stopping and starting points in the programme, since it is unlikely that you will have time to read through the whole book in one session. Always read the appropriate intermission again before going on to a new section. Additional stopping points are marked by thick double lines.

Protein synthesis

An examination of protein structure

1 Enzymes like all proteins, are large molecules made up of a specific number of 'building blocks'.
What are these 'building blocks'?

Amino acids

2 How do amino acids get inside (*a*) animal cells and (*b*) plant cells?

(*a*) Mainly across the cell membrane (i.e. animal cells rely on an exogenous source); even though most animal cells are capable of synthesizing some amino acids.
(*b*) Higher plant cells and many micro-organisms (but not all) are capable of synthesizing all twenty amino acids required for protein synthesis. Amino acids can also gain access into these cells via the cell membrane.

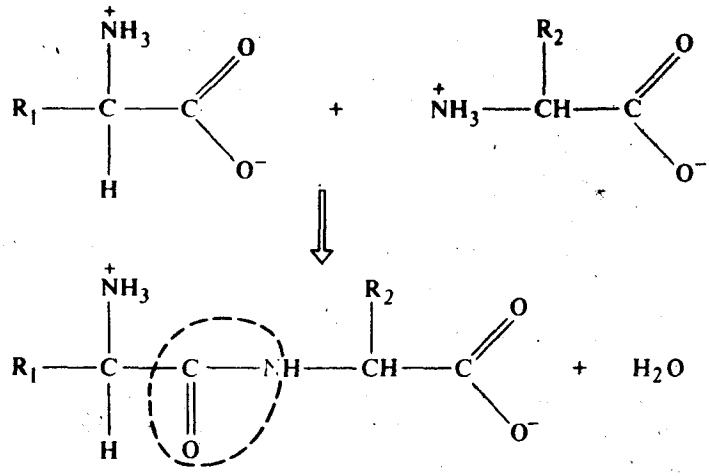
3 In order to synthesize a particular enzyme, what must a cell be able to do?

Assemble together the various amino acids in a specific order which is characteristic of that enzyme.

4 The primary structure of all proteins is determined by the specific order of the constituent amino acids joined together *into a chain*. Although only twenty different kinds of amino acids are used, their proportional representation varies greatly from protein to protein. Moreover, since protein chains typically contain several hundreds of amino acids, the number of different amino acid sequences possible is astronomically large. This permits a high degree of specificity among proteins.
How many different sequences are still possible in a protein chain of 100 amino acids?

20^{100}

5 Obviously the amino acids cannot be linked together randomly if we want multiple copies of a specific protein. But, before considering how the specific sequence is determined, let us consider further the nature of the chemical bond linking the amino acids together. The linkage is always between the carboxyl group of one amino acid and the *amino group* of the adjacent amino acid (see fig. 1).



A dipeptide

Fig. 1

The linkage shown above is referred to as a _____?_____ bond; and a protein is often referred to as a _____?_____ chain of amino acids.

peptide bond; *polypeptide* chain

6 An important feature of the C–N peptide bond is that it has a 40% double bond character. This is compensated for by the C=O bond having a 40% single bond character; and, in effect, the electrons of the p orbital of the N atom as well as those of the O atom are shared between the two bonds. This has two important consequences:

(1) The bond is stabilized, i.e. ΔG for its formation is lower.*

(2) All the atoms attached to the C and N atoms have to be co-planar, and this prevents the free rotation of the peptide bond.

The formation of a dipeptide and a molecule of water from two amino acids (see previous page) requires a ΔG° of +4.2 to +16.8 kJ/mole, depending on which amino acids are being bound. Furthermore, water has a very much higher concentration in a cell (≙ 100 times greater) than any other molecule.

In the light of this information would it be possible for proteins to form spontaneously from amino acids? Explain your answer.

* An explanation of the terms ΔG and ΔG° is given in Book 7 in this series; and, more briefly, in the Glossary on p. 101 under 'Standard free energy change'.

No, since $\Delta G°$ for peptide bond formation is always positive (i.e. energy-requiring). Also the water concentration in the cell will tend to 'push' the reaction in the direction in which water molecules are consumed (i.e. hydrolysis of peptide bonds).

---

7 What substances are available in the cell which might make good this deficiency in free energy?

Compounds like ATP in which the $\Delta G°$ of hydrolysis has a high negative value at neutral pH

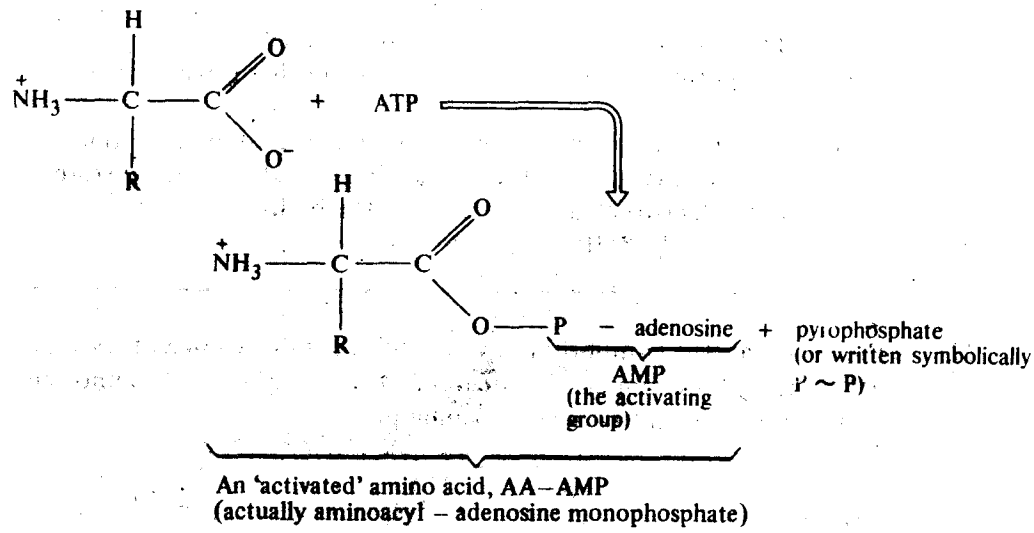
---

8 How can the energy of these compounds become available to the amino acids for biosynthetic work?

By coupling the biosynthesis of peptides (where $\Delta G°$ is not greater than +16.8 kJ/mole) with the breakdown of ATP. In fact in this case ATP is broken down to AMP and pyrophosphate with a $\Delta G°$ of −33.6 kJ/mole.

---

9 The coupling reaction is as follows:



Why *cannot* this coupled reaction simply be two independent reactions?

i.e. (i) the hydrolysis of ATP ($\Delta G°$ negative)

$$ATP \longrightarrow AMP + P{\sim}P$$

and (ii) the synthesis of the AA–AMP complex ($\Delta G°$ positive)

$$\text{Amino acid} + AMP \longrightarrow \text{Aminoacyl–AMP} + H_2O$$

Because most of the free energy liberated by hydrolysis of ATP would be lost as heat

---

10 In the coupled reaction shown above some of the energy that was 'stored' in ATP is transferred to the aminoacyl–AMP complex. Thus the amino acid is considered to be in an activated state. However, the coupled reaction has a slightly positive $\Delta G°$. What are the implications of this for the concentration of the activated amino acid complex?

The concentration is likely to be fairly low, because the back reaction is thermodynamically favoured over the forward reaction.

---

11 How could the cell continue to synthesize 'activated amino acids' if the back reaction is thermodynamically favoured?
(*Hint*: there are a number of possibilities.)

Some possibilities are:

(i) Its synthesis might be coupled with the conversion of pyrophosphate (the other product) to two phosphate ions.

$$P{\sim}P \xrightarrow{\text{Pyrophosphatase}} P_i + P_i$$

The reaction has a $\Delta G°$ of -29.4 kJ/mole.

or (ii) Its synthesis might be coupled with another (unspecified) reaction.

or (iii) Its synthesis might be coupled with a later step in peptide bond formation and the release of some of the energy stored in the 'activated amino acid' (i.e. the hydrolysis of the aminoacyl–AMP complex).

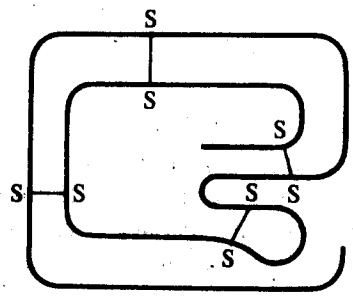
---

12 Is the idea that 'the equilibrium favours polypeptide chain synthesis' in contradiction to the implication of frame 4 that 'amino acids cannot be linked together randomly' when forming proteins?

NO. Frame 4 dealt with specific sequencing, suggesting that there must be some informational source determining the order in which the amino acids are linked. However, the formation of each peptide bond must be thermodynamically feasible if the chain is to be synthesized. In fact, there is a very large negative $\Delta G°$ for the net conversions of amino acids and ATP to polypeptide chains and AMP.

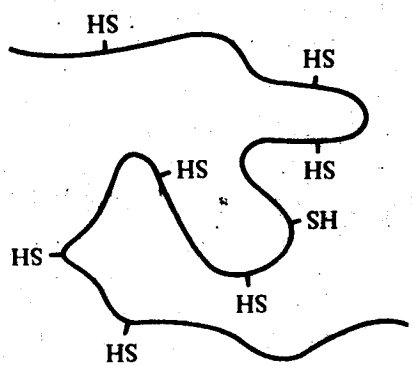
---

13 Discussion so far has centred around the synthesis of polypeptide chains, the *primary structure* of proteins. No reference has been made to the secondary or tertiary structures, which together determine the three-dimensional conformation of the protein molecule. However, we know from Book 7 that the three-dimensional spatial folding of an enzyme into a globular form is essential for its activity. The question we now need to ask is whether such folding can take place spontaneously as the polypeptide chain is formed from amino acids in the aqueous surroundings of the cell, or whether some additional hypothesis is needed to account for the protein's secondary and tertiary structure. Let us, therefore, examine ribonuclease, a fairly simple enzyme which breaks down ribonucleic acid (RNA). Ribonuclease isolated from the pancreas of cattle has a sequence of 124 component amino acids. Examination of its three-dimensional structure shows it to contain four disulphide linkages (i.e. positions where the two S atoms of two cysteine amino acids are covalently linked).



A simplified model of ribonuclease; native or active conformation.

Loss of enzyme activity may be brought about by heating, acidification or exposure to high concentration of salts or urea.



Inactive conformation of ribonuclease after treatment with urea.

When the enzyme is restored to physiological conditions (i.e. those resembling the living cell) and the denaturing agent(s) removed, enzymic activity is gradually restored over a period of time.

(i) Could the enzyme have more than one three-dimensional configuration? Why?

(ii) Is it likely to have more than one enzymatically *active* conformation?

(iii) What light do these experiments shed on the idea that the amino acid sequence of the polypeptide chain is also responsible for the spontaneous formation of the enzyme's tertiary structure?

(i) Looking at the *inactive* conformation of ribonuclease without reference to the native conformation, we cannot be certain which are the proper pairings for the cysteine residues; therefore, the enzyme certainly could have more than one three-dimensional configuration if the pairings for the cysteine residues were to be completely random.

(ii) It is unlikely from the evidence presented in Book 7 that there is more than one enzymatically *active* conformation.

(iii) The experiments indicate that the amino acid sequence of ribonuclease is also responsible for its *secondary* and *tertiary* structure.

---

14 Not all proteins regain their biological activity when they are restored to physiological conditions and the denaturing agents removed. Still assuming that the amino acid sequence plays the major role in determining how the polypeptide chain is folded, can you suggest why some proteins might not 'renature'?

The conditions at the time of synthesis may be unique. For example, folding might begin during synthesis (i.e. before the chain is completed), eliminating some of the possible secondary and tertiary conformations, including the conformation favoured when the completed chain is 'refolded'.

---

15 Is the native (active) conformation of the enzyme in aqueous surroundings necessarily the structure which has the least free energy (i.e. the most stable conformation thermodynamically)?

No, the active conformation of any enzyme is not necessarily the one with least free energy, although in most cases it is.

---

16 The specific three-dimensional conformation of an enzyme within the living cell is determined both by the intrinsic properties of the polypeptide chain itself and the properties of the aqueous medium surrounding it. As each amino acid is added to the growing chain, the tertiary conformation should reflect the structural adjustments that will produce local minimum free energy states.

With these properties in mind, which of the following features will maintain the enzyme's three-dimensional structure?

(i) The rotation of C—C bonds in amino acid side-chains.

(ii) Intramolecular hydrogen bonding between the –NH of one peptide linkage and the $>C=O$ of the fourth amino acid behind it.

(iii) Sections of the polypeptide chain in the form of α-helices.

(iv) Proline residues which distort α-helices and create kinks or bends.

(v) The positioning of polar groups on the outside of the globular mol-

ecule where they can be 'hydrated', while non-polar groups are 'tucked' inside.

(vi) Ionic bonding between polar side-chains.

(vii) The formation of disulphide bridges (S–S) between –SH groups on cysteine residues.

(viii) The vibrational energy of the molecule.

(ix) The lack of free rotation in the C–N peptide bonds.

(*Hint*: you may need to use the knowledge acquired in Book 7 in the consideration of some of these features.)

Most of them contribute to the maintenance of the enzyme's three-dimensional structure, except (i) and (viii).

(i) The rotation of C–C bonds in amino acids certainly contributes to events which lead to the formation of the tertiary structure, but once the tertiary structure is attained, rotation of C–C bonds is minimized. It therefore plays no significant part in maintaining the tertiary structure.

(viii) The vibrational energy of the molecule is related to temperature; and whilst the vibrational energy will have little effect on the tertiary structure at normal or low temperatures, it will have a marked disrupting (denaturing) effect at high temperature; the molecule will 'shake itself apart'.

---

17 Intermission 1

The role of genetic material in determining protein specificity was indicated by the demonstration that genetic defects are associated with defective proteins. Although examples date back to Garrod's study in 1908 of 'inborn errors of metabolism' in man, much of the pioneer work associating the structural determination of a polypeptide with a gene was done by Beadle, Tatum and their co-workers in the early 1940s using the bread mould, *Neurospora crassa*.

The first mutant protein to be chemically compared with its normal counterpart was an abnormal haemoglobin in man found among individuals suffering from *sickle-cell anaemia*. Pedigree analyses have established that sickle-cell anaemia is associated with a single genetic defect; individuals with the conditions being homozygous for a recessive gene. The name of this anaemia derives from the peculiar elongated and sickle shape of the red corpuscles of people suffering from the disease under conditions of low partial pressure of oxygen. Fragility of these red blood cells causes the severe anaemia, as well as numerous secondary clinical effects, arising especially as a consequence of the blockage of small vessels.

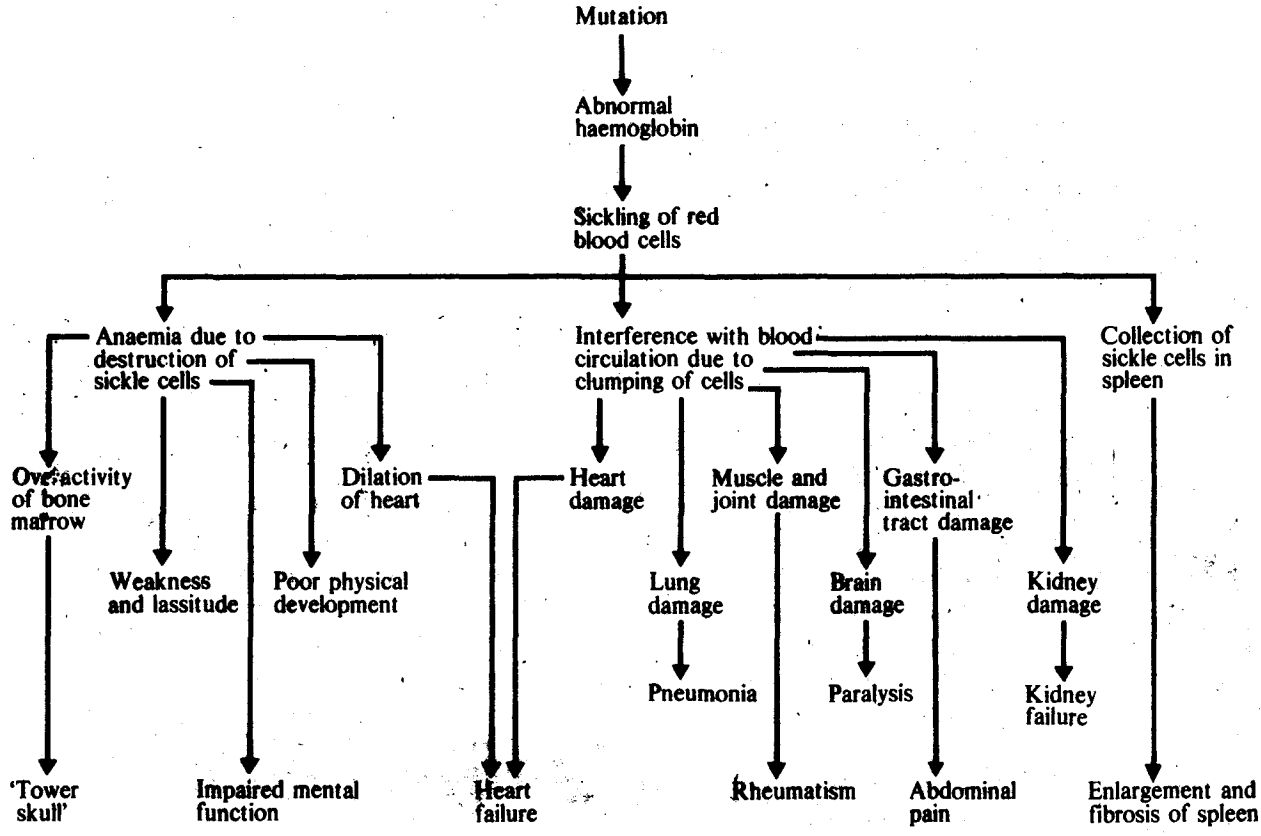


Fig. 2. Table of symptoms associated with the sickle-cell trait.

Fig. 2 shows how this single mutation can induce a multiplicity of effects. The next series of frames will consider the manner in which the protein is altered in the mutant organism. In addition to telling us much about the relationship between the genetic material and the protein, it provides us with further insight into the relationship between the protein's structure and its biological activity.

Sickle-cell anaemia

18 As we showed in Book 7, proteins carry electrical charges because of the ionization of amino, carboxyl and other functional groups in the molecule. These properties can be used to advantage in *electrophoresis*, because when proteins are placed in an electric field, the *net* charge on each molecule will cause migration towards either cathode or anode. Electrophoresis, therefore, provides a means of separating different proteins whose positions in the electrical field may be detected subsequently by special stains (e.g. ninhydrin reagent).
What factors will affect the speed of migration of a given protein in the electric field?

(i) The net charge on the molecule;
(ii) the shape of the molecule;
(iii) the polarizing voltage applied to the electrodes establishing the field;
(iv) the solvent in which the protein is placed and the pH of the medium.

---

19 In 1949 Pauling with his co-workers showed that the haemoglobin from individuals with sickle-cell anaemia (HbS) can be separated from normal adult haemoglobin (HbA) by their differential migration in an electric field (*electrophoresis*). Whereas the normal haemoglobin in reduced state (carbomonoxy-haemoglobin) has a net negative charge at pH 6.9 and is, therefore, attracted to the positive electrode in the electrical field (as in fig. 3*a* below), the reduced state of sickle-cell haemoglobin has a net positive charge at this pH. It, therefore, tends towards the negative electrode in the electrical field as in fig. 3*b* below).

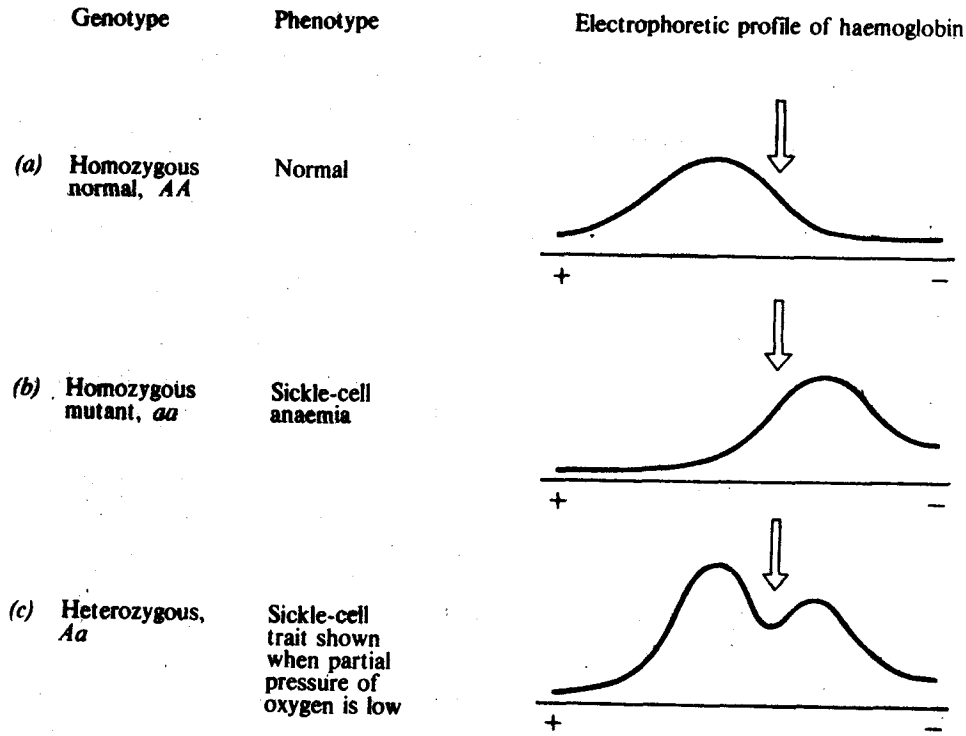


Fig. 3

(i) Why is there no HbA present in individuals with sickle-cell anaemia?
(ii) Why are both haemoglobins present in heterozygous individuals with sickle-cell trait?

(i) HbA is replaced by the variant haemoglobin form HbS. However, this technique does not distinguish between a change in the primary structure of haemoglobin or a simple change in its tertiary structure exposing fewer free carboxyl groups.
(ii) The heterozygote has two alleles of the gene, one determining HbA and the other HbS. In the heterozygote both are expressed, i.e. both proteins are synthesized.

---

20 The normal adult haemoglobin (HbA) is a molecule composed of four polypeptide chains, each associated with a haem (iron-containing) group and functioning in oxygen transport. The polypeptide chains are of two types; there are two α-chains, each with 141 amino acid residues and two β-chains, each with 146 amino acid residues. The problem then was to compare the amino acid sequence of these (HbA) globin chains with those of HbS. Although chemical methods were available for a stepwise breakdown of a chain, i.e. splitting off the residues as free amino acids one at a time, these procedures, unfortunately, are reliable for only short lengths from each end of the chain. Interpretation of the sequence, therefore, soon becomes confused by both asynchronous degradation and the increasing possibility of repeats among the amino acid residues released. Therefore, the protein must be broken into fragments and these smaller peptides separated and isolated so that each type of fragment can be analysed individually for its amino acid sequence.
Can you suggest what type of methodological approach is needed to obtain the peptide fragments?
(*Hint*: peptide fragments created by such non-specific attacks as acid hydrolysis are usually too small and too numerous to be of much value.)

What is needed is a means of breaking the polypeptides into a limited number of unique fragments by preferentially attacking only certain peptide linkages.
One such method uses trypsin, which is known to split the protein chains specifically at the carboxyl side of arginine and lysine residues. Though the number and kind of fragments obtained by trypsin digestion will depend upon the location of arginine and lysine residues, each molecule of the protein will be attacked in the same way, thus giving the desired uniformity of result.

---

21 The next step is to separate the different types of fragments so that their amino acid sequences can be determined.
How would you separate the different types of fragments?
(*Hint*: we have already discussed the technique in this Unit.)

There are several procedures that will distinguish certain physical and/or chemical differences among the fragments. One possibility is electrophoresis (see frames 18 and 19), after which the position of the frag-