GENES AND ENZYMES

Patricia H. CLARKE
Department of Biochemistry, University College London, Gower Street, London WC1E 6BT, UK

One gene – one enzyme?

'This . . . . locates the error in the penultimate stage of the catabolism of the aromatic protein fractions which is in accordance with the fact that all the tyrosine and phenylalanine, both exogenous and endogenous, is swept into the net and goes to contribute to the excreted homogentisic acid in alkaptonuria. We may further conceive that the splitting of the benzene ring in normal metabolism is the work of a special enzyme, that in congenital alkaptonuria this enzyme is wanting whilst in disease its working may be partially or even completely inhibited' [1].

This very clear statement by Garrod in his Croonian lectures in 1908 leaves no doubt that he was certain that in his investigations of certain congenital metabolic disorders he was dealing with defects in specific enzymes. He had observed, as had others before him, that some disorders such as alkaptonuria, although rare among the general population, tended to occur in more than one member of an afflicted family. Garrod confirmed earlier findings that the dark colour of the urine of alkaptonuriacs was due to the presence of homogentisic acid, which characteristically darkened on exposure to air especially in alkaline solution. He showed that feeding tryptophan to an alkaptonuriac boy had no effect on the excretion of homogentisic acid, whereas feeding tyrosine resulted in a marked increase. He also showed that the output of homogentisic acid by alkaptonuriacs was not affected by 'intestinal disinfection' so that there could be no question of microbial conversion of tyrosine to homogentisic acid in the gut followed by reabsorption. This was a very clear cut demonstration in alkaptonuriacs of what later came to be called a metabolic block. Garrod discussed his findings with Bateson, who pointed out that the familiar incidence of alkaptonuria 'finds a ready explanation if the anomaly in question be regarded as a rare recessive character in the Mendelian sense' [1]. Since Garrod had good evidence to suggest that 'the anomaly in question' was a defect in the splitting of the benzene ring by a specific enzyme, and since Bateson was willing to accept that the inheritance of this defect was in accordance with the newly rediscovered laws of Mendel, one might have thought that only one third of a century from the appearance of 'the enzyme', the stage was already set for a period of fruitful collaboration between biochemists and geneticists.

However, neither biochemists nor geneticists could develop these ideas at this time. Bateson's own account of Garrod's findings now appears very shadowy. ' . . . as Garrod has shown, alkaptonuria must be regarded as due to the absence of a certain ferment which has the power of decomposing the substance alkapton' [2]. Later, he was to show considerable insight into the relations which might exist between genes and enzymes. During the first decade of this century, the geneticists were still accumulating examples of the segregation of characters in a variety of species. They badly needed more evidence to assess whether Mendel's laws were of general application or whether they had only limited application to special cases. One of the main difficulties which was to persist for many years was that many of the heritable differences which could be observed among groups of plants or animals were rather complex ones and it was often difficult, if not impossible, to sort them out in such a way that clear-cut segregation of characters, comparable to Mendel's tall and dwarf peas, could be seen at all. Another difficulty which was to recur time and time again, and sometimes in suprising ways, was to distinguish the elements of heredity from their manifestations in differences which could be

observed in individuals with respect to colour, height or other attributes. Johannsen in 1909, (cited by Carlson in 'The Gene: A Critical History' [3]) made a valuable suggestion, which might have cleared up this confusion once and for all, in proposing that the postulated units of heredity should be called *genes*; that the total of all the entities determining inheritance in an individual should be called the *genotype*; that the observable characteristics of the individual, the raw data of the geneticist, should be called the *phenotype*. However, the material available to most geneticists was still so intractable that this proposal did not meet with general acceptance for a long time and confusion was rife for many more years to come.

In due course the concept of the gene as the elementary and indivisible unit of heredity became accepted although progress had been painful, slow and often acrimonious. The association of certain characters in genetic crosses had at first been taken as evidence against Mendelian segregation but was now interpreted as gene linkage. The role of the chromosomes in inheritance was also becoming clearer so that it was possible to relate the hypothetical genetic elements to a definite physical entity. Sturtevant [4], following up Morgan's suggestion that the 'strength' of gene linkage might be dependent on the distance apart of the genes, calculated the relative distances between six genes associated with the X chromosome of *Drosophila* from the numbers of crossovers, and constructed the first genetic map. In 1922 Morgan was able to say .... 'The evidence from crossing over has led to the conclusion that the hereditary elements, the genes, are arranged in linear order in the chromosome' [5]. About this time, Muller, who had very excited to hear about the strange bacteriophages isolated by Twort and d'Herelle, thought that if these really were simple arrangements of genes, fundamentally like the genes of the chromosome, 'they would give us an utterly new angle from which to attack the gene problem'. This thought led him to ask the prophetic question: 'Must we geneticists become bacteriologists, physical chemists and physicists, simultaneously with being zoologists and botanists? Let us hope so' [6]. But at that time no geneticists actually did this and the biochemists, chemists and physicists were not thinking about genetics at all. Metabolism, the composition of tissues and the chemistry of biological compounds were the main preoccupations of biochemists during the 1920s.

With the development of new techniques it became possible to carry out more precise studies on enzymes, to purify them and even crystallize them. Metabolic investigations were extended from yeasts and mammals to a variety of plant and animal species and in some laboratories to bacteria. This resulted in the gradual building-up of the concept of the essential unity of biochemical processes in spite of the diversity of biological activities apparent in nature. Among the biochemists of this period was M. Stephenson at Cambridge, whose work on bacterial metabolism and bacterial enzyme synthesis laid some of the foundations for the later use of bacteria as model systems for exploring biochemical genetics. During the 1930s it was found that compounds which were vitamins for man and animals could also be growth factors for bacteria. Knight [7] showed that *Staphylococcus aureus* required both nicotinamide and thiamine for growth and when Lohmann and Schuster [8] identified thiamine pyrophosphate as the coenzyme of yeast carboxylase, the potential of bacteria for fundamental biochemical studies became apparent. It was not until about ten years later, that their potential for genetic investigations became realised.

Is the gene an enzyme?

The geneticists meanwhile were trying to find out more about genes. The fruitfly, *Drosophila*, first introduced as an experimental organism around 1900 was now being widely used. The development of *Drosphila* genetics, which was mainly the work of Morgan and his colleagues, made it possible greatly to extend fundamental studies on the nature of the gene. *Drosophila* has a relatively short mean generation time of a couple of weeks, and fairly large populations can be maintained in simple equipment and fed quite cheaply. It was not difficult to produce mutant forms, thus providing populations carrying more and more characters for genetic analysis. In line with traditional studies the characters first chosen for study were mostly differences in morphology but as time went on, it became possible to examine characters which could be related to

underlying differences in biochemical reactions. The most obvious examples of these were the studies on eye-colour mutants. Among those who worked on this were Beadle and Ephrussi who concluded in 1936 [9] that there were two genes in control of the biochemical reactions leading to the synthesis of the eye pigments of the mutants they had examined. Beadle is now associated by most biochemists and microbiologists with his later work on *Neurospora* and in the Annual Review of Biochemistry 1974 he recalls the chain of reasoning which led him to change over from working with a fly to working with a mould [10]. He explains that during 1940–41 he was sitting-in on some lectures by Tatum on comparative biochemistry and realised that it would be easier to examine the gene–enzyme relationship if he took a known biochemical pathway and looked for mutants, rather than first making mutants and then looking for enzyme changes as he had been doing in *Drosophila*. At that time he was unaware that 40 years earlier Garrod had obtained evidence to suggest that one gene determined one enzyme and the early work with *Neurospora* developed entirely from his own work with *Drosophila* and that of other workers on plant pigments. Within a very few years Beadle and Tatum had shown that the genes for the separate steps in a biochemical pathway could be identified by mutations and that one gene, and one only, was concerned with each step of the pathway [11]. In many cases the details of the biosynthetic pathways with which they were concerned had not been fully established. An indirect effect of their work was to improve this situation. The many mutants with enzyme lesions were valuable by-products of the genetic studies, and the ways in which they were put to work can be traced in texbooks of biochemistry where mention was soon made of the identification of enzymes and intermediates of amino acid biosynthetic pathways from experiments using *Neurospora* mutants (see for example Fruton and Simmonds on histidine biosynthesis [12]).

It was now well-known that chromosomes contained both protein and nucleic acid and Beadle [13] discussed whether these molecules could be equated with genes. The most striking characteristic of a gene was the very specific effect it could produce and it was possible to see this as similar to the high specificity of enzymes with respect to their substrates and the reactions catalysed. Beadle said, in 1945, '. . . . it is probable that gene specificity is determined by the protein component' [13]. It would have been rather difficult at the time to reach any other conclusion. Proteins, and in a different way carbohydrates, were the only macromolecules whose structure and function were understood even superficially, and better the molecules you know than the molecules you don't know. Reductionist arguments for the view that genes were not just proteins but were themselves enzymes, had been put forward much earlier. Troland in 1917 [14] dismissed Bateson's qualified statement that enzymes, and the bodies on which enzymes act, were the consequences of genetic factors and not the genetic factors themselves as 'intellectual blindness'. Troland also said: 'On the supposition that the actual Mendelian factors are enzymes, nearly all these general difficulties instantly vanish, and I am not acquainted with any evidence that is inconsistent with this supposition'.

The problems that were so lightly dismissed by Troland were to do with the twin attributes of the gene. First, it had to have some way to reproduce itself and secondly, it had to have some way to express itself in the phenotype or, to be more precise, there had to be a way for a change in a gene to appear as a change in an enzyme. The simplest way to restate the problem of gene replication was to say that the process was autocatalytic, that the biological catalysts were enzymes and that it was logical to suppose that the genes were special sorts of enzymes which could reproduce themselves (the extreme case was that all enzymes could reproduce themselves). It was easier to hold this view at a time when no-one knew how proteins were synthesised and when few people had faced the inherent difficulties in replicating a molecule containing a unique combination of amino acids. It was the complexity of the whole process of gene reproduction and expression which had led Bateson, quite correctly, to think of enzymes as the consequences of genes and not as the genes themselves. However, such reasoning could not rule out that *some* proteins might be genes, and while the evidence about the nature of genes and enzymes rested mainly on experiments with higher organisms these arguments could continue indefinitely.

The bacteriologists were interested in certain problems of heredity and had been for a long time.

In earlier days, during the latter part of the 19th century, they had needed to establish with great care the conditions needed for isolating and maintaining pure cultures and they found that most of the characteristics of their cultures were maintained indefinitely. However, it was also common knowledge that variants appeared spontaneously which differed from the parent strain with respect to such properties as colonial form or the capacity to produce pigments. Dubos, in: 'The Bacterial Cell' in 1945 [15] pointed out that 'analysis of these phenomena in terms of classical genetics presents many difficulties'. Although it had been reported from time to time that conjugation of bacteria could be seen under the microscope, the reports were not generally accepted nor were the reports that nuclear structures in various stages of division could be seen by appropriate staining methods. Nevertheless, even if bacteria reproduced solely by binary fission, there had to be some way of ensuring genetic continuity and Dubos thought that one could safely assume that bacterial inheritance took place through the agency of genes. Dubos suggested that the simplest possible arrangement for bacteria would be 'a single gene string existing as a rod or granule' [15]. This was a pointer to the future when it would at last be recognised that the genetic structure of procaryotes was far less complex than that of eucaryotes.

One of the most interesting of the morphological variations of bacteria was the spontaneous change from smooth to rough colonial forms. With the pneumococci this change could be related to the loss of capsular polysaccharide together with loss of virulence. It had been known for many years that rough pneumococci could be transformed back to smooth capsulated forms with heat-killed cells or extracts [16] but with the publication in 1944 of the now legendary paper of Avery, McCleod and McCarthy [17], bacterial genetics and the whole story of the biochemistry of the gene entered a new dimension altogether. Avery's experiments showed that the substance eliciting the pneumococcal transformation was DNA and that only 0.003 μg was sufficient to produce transformants from a few millilitres of bacterial suspension. They concluded their paper, 'If the results of the present study on the chemical nature of the transforming principle are confirmed, then nucleic acids must be regarded as possessing biological specificity the chemical basis of which is as yet undetermined . . .' [17]. In 'The Bacterial Cell' [15] the following year Dubos was hopeful that these experiments could provide the pattern for analysing bacterial variability and inheritance.

It was not long after this that Tatum and Lederberg [18] demonstrated that exchange of genetic material could occur by conjugation of bacteria and this began the revolution in genetics that has made *Escherichia coli* strain K12 the most widely cultivated organism in the world. In due course it was shown that the bacterial genes were indeed arranged in a single string and that this string was a circle of DNA [19,20]. The large bacteriophages, whose existence had so enthralled Muller nearly half a century earlier, were also shown to contain DNA and moreover it was found, by 1952, that for replication of the bacteriophage to occur in an infected host, it was essential for the DNA to enter the cell, but the bacteriophage coat, the protein component, could be left outside and abandoned [21].

Thus, within a few years from the publication of Avery's paper on the nature of the transforming principle the idea that proteins (enzymes) were genes had become untenable but, while the evidence was based solely on microbial experiments and until the genetics of bacteria and bacteriophages had become clearer, the claim that genes were really proteins was bound to crop up again from time to time and duly did so [22]. The hope of Avery [17] that the chemical basis of the specificity of DNA as carrier of genetic information would be firmly established was realised in 1953, when, following much detailed chemical and crystallographic work by a number of workers in various laboratories, Watson and Crick proposed the double helix structure [23]. They pointed out that this structure not only satisfied the physicochemical characteristics of the DNA molecule, but could allow for faithful replication by making copies of each of the complementary strands. From that time on the chemical nature of the gene was no longer in doubt and the details of its replication and the events by which enzymes were made could be pursued with confidence.

From gene to enzyme

Some of the most pertinent evidence on how

enzymes are made came from studying variations in rates of bacterial enzyme synthesis in different growth media. Observations going back to the turn of the century indicated that many catabolic enzymes were made only if the potential growth substrate was present in the growth medium. By the 1930s there were many observations of the effect of the environment on bacterial enzyme synthesis and it became accepted that bacteria could 'adapt' to produce catabolic enzymes and could be 'trained' to dispense with growth supplements [24,25]. The use of the terms 'adaptation' and 'training', though justifiable at the time to describe the observations, turned out to be unfortunate and resulted in more than semantic confusion. The question which needed to be asked in these experiments was 'were the variations which could be observed due to mutational changes in some of the bacteria, which would then be selected by the environment and outgrow the original strain, or were they due to temporary non-inherited variations which would be lost when the culture was returned to the original medium?'.

Much of the earlier data came from comparing the enzymic activities of washed suspensions of bacteria grown in different media. This produced a wealth of information on the effects of such factors as the complexity or otherwise of the growth medium, the nature of the major carbon and nitrogen sources, pH, oxygen and whether or not glucose was present in the growth medium. Gale, in 1943, analysed the results of studies of this type and said that the distinction should be drawn between '... the potential enzymic constitution representing the repertoire of an organism, and the actual enzymic constitution which is that part of the potential constitution which is produced in response to a given set of growth conditions' [26]. This was a clear distinction between genotype and phenotype but 'adaptation' continued to be used indiscriminately to describe both phenotypic and genotypic changes in bacterial cultures. When the synthesis of catabolic enzymes following the addition of a new substrate to a growing culture was examined in more detail it could be demonstrated that this was truly a phenotypic event, and to remove all confusion the phenomenon of enzyme synthesis in response to substrate became known as *induced enzyme synthesis.* Paradoxically this became a valuable tool in the hands of the bacterial geneticists.

β-Galactosidase was of singular importance in unravelling the details of enzyme synthesis and Jacob and Monod pointed out in 'The Lactose Operon' in 1970 [27] that one of the main reasons for studying this enzyme was the availability of a large number of lactose analogues. The finding that β-galactosidase could be induced by galactosides which were not substrates of the enzyme, made it clear that inducer specificity was not the same as substrate specificity [28]. This made it possible to use non-substrate inducers to study enzyme synthesis under conditions of gratuity, when the enzyme was not needed for growth, with the inducer concentration kept constant over the experimental period. Further, it meant that there must be another molecule, other than the enzyme, which interacted specifically with β- galactosides. By 1953, it was known that the synthesis of induced enzymes occurred within a very short time from the addition of inducer to exponentially growing cultures and that this was 'de novo' protein synthesis, and that no high molecular weight precursors were involved. A complete pool of amino acids was required indicating that the synthesis of enzyme protein was an all-or-none phenomenon and this led directly to the idea that enzyme synthesis took place by a template mechanism [29–31].

Further studies with β-galactosidase suggested the identity of the template and explained why inducer specificity might differ from enzyme specificity. By this time bacterial genetics had become much less mysterious. Hayes in 1953 [32] found that the donor or male bacteria of *E. coli* K12 possessed a sex factor F which was missing from the recipient or female bacteria; they became known as F^+ and F^- respectively. A little later [33] the Hfr males were discovered, which were more efficient in conjugal transfer, and it was found that in these strains the sex factor F had become integrated in the circular bacterial chromosome. Occasionally F factors arise which have acquired some of the chromosomal genes in addition to their usual genes determining replication and transfer activity [34]. These are the F′ factors and a particularly useful one was F′ *lac* carrying genes for lactose catabolism. The F′ *lac* factors could be retained in bacteria which already carried *lac* genes on the chromosome. Thus it was possible to carry out dominance tests with bacteria which were partial diploids. Bacteria normally carry only a single chromosome and are therefore haploid

organisms, but partial diploids may be obtained experimentally by inserting extrachromosomal elements carrying the appropriate genes.

By this time some mutants had been isolated which were defective with respect to β-galactosidase activity or for uptake of galactosides (*lacZ, lacY*), and others which were constitutive, and able to synthesise enzyme and permease in the absence of inducer (*lacI*). Using the diploid dominance test it was shown that inducibility was dominant to constitutivity, irrespective of whether the *lacI*$^+$ gene was carried on the chromosome or the F′ *lac*, and whether or not it was on the same gene segment as the *lacZY* gene region it controlled. The inference was that the *lacI* gene determined the structure of a molecule, the *lac* repressor, which prevented the expression of the *lac* genes unless an inducer was present [31]. The chemical nature of the *lac* repressor was established following its isolation by Gilbert and Müller-Hill in 1966 [35]. It is now known to be a tetrameric protein which combines specifically, and with very high affinity, with a short gene region (*lacO*) comprising 27 nucleotides very near the *lacZ* gene [36,37]. It was suggested by Jacob and Monod in 1961, that by binding to the operator region, *lacO*, the repressor prevented the transcription of the DNA of the *lac* operon genes into a complementary RNA copy – the *lac* messenger. This was the postulated template for protein synthesis and within a very few years a number of specific messenger RNAs had been identified. At the same time experiments with in vitro systems were under way, and in 1961 Nirenberg and Matthaei [38] announced that an in vitro system consisting of ribosomes, cell extracts containing enzymes, nucleotides and other factors together with phenylalanine could translate a polyuridylic acid RNA messenger into a polypeptide chain consisting entirely of polyphenylalanine. This was the first step in deciphering the code whereby a sequence of three bases in the DNA of the gene defines a single amino acid of the polypeptide chain of an enzyme.

The story of the unravelling of the genetic code, the start and stop signals, the designated reading frame, which could be displaced by the insertion or deletion of bases, and the identification of the exact codons used in vivo for a particular amino acid has been well documented elsewhere and is part of contemporary biochemistry and genetics. The earlier geneticists had to be content with mapping the distance apart of two genes, but the finer resolution of microbial genetics made it possible to detect recombination between mutational sites very short distances apart within a single gene. Benzer, with bacteriophage T_4, was able to detect crossovers between adjacent nucleotides and thus reached the limits of possible genetic recombination [39]. The colinearity of gene and protein is a basic assumption in all this work and was given formal proof by experiments relating the genetic map of mutational sites of structural genes for the head protein of bacteriophage T_4 [40] and the *trpA* gene of *E. coli* [41] to the sequences of their proteins. We are now in no doubt that mutations in the sequence of bases in the DNA of a gene may be reflected in the sequence of amino acids in the polypeptide chain of an enzyme or, in with some mutational changes, in the absence of a completed enzyme molecule.

Control of gene expression

The explanation of the control of β-galactosidase synthesis rested on the hypothesis that there could be two types of gene, the *structural gene* determining the amino acid sequence of the enzyme and a *regulator gene* which determined a specific repressor protein. A similar mechanism could be invoked to explain why biosynthetic enzymes were not synthesised when the end-product was present, if it was assumed that in the latter case the regulator gene produced an inactive form of a repressor which was only activated when combined with the end-product. In both cases the control was negative and involved the prevention of transcription; frequently several genes were found, in *E. coli* at least, to be closely linked and under co-ordinate control thus constituting the functional genetic control element, the *operon*. The operon theory filled so many requirements for a logical theory of gene–enzyme interaction that it met with immediate success and it was overlooked that it would have been equally satisfactory if the regulator gene produced an activator protein which facilitated transcription when combined with the inducer. Examples of this were found later, and it is interesting that the first evidence for positive control was for the enzymes for L-arabinose catabolism, the *ara* operon, and came from genetic tests for domin-

ance [42]. Later it was found that many operons for catabolic genes in *E. coli* were regulated by dual control and required a general activator protein, determined by the *crp* gene, needing cyclic AMP to get it into the active state. The cyclic AMP level is low in the presence of a good carbon source such as glucose so that glucose may prevent transcription even when the inducer is present. These conclusions, based in the first instance on in vivo experiments with actively growing bacteria, have now been fully confirmed with in vitro systems able to transcribe DNA into RNA and to translate RNA into protein [43,44].

It was quite logical to suppose that an enzyme might act as either an activator or repressor for the transcription of its own genes or for the genes determining other enzymes. The most versatile of the candidates which have so far appeared to fill this role is glutamine synthetase. This most important enzyme is at the heart of bacterial nitrogen assimilation and in addition to being absolutely required by many species at low concentrations of ammonia it is able to stimulate the synthesis of other enzyme systems which might provide nitrogen for growth [45]. Fig.1 is based on the work of Magasanik and colleagues and shows the way in which

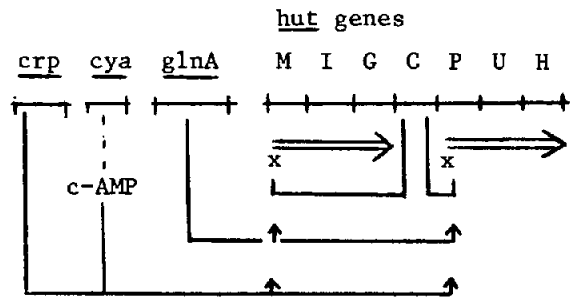

Fig.1. Regulation of the histidine utilization pathway genes (*hut*) of *Salmonella typhimurium*. The *hut* genes are arranged in two operons under the negative control of the regulator gene *hutC*. The left-hand operon is *hutMIGC*; *hutM* is the promoter; *hutI* determines imidazolonepropionate hydrolase; *hutG* determines formiminoglutamate hydrolase. The right-hand operon is *hutPUH*; *hutP* is the promoter; *hutU* determines urocanase; *hutH* determines histidase. Gene *cya* determines adenyl cyclase and *crp* the c-AMP receptor protein. Gene *glnA* is the structural gene for glutamine synthetase. The bold arrows indicate the direction of transcription [45].

glutamine synthetase may activate transcription of the *hut* genes concerned with histidine utilization. The *hut* genes of *Salmonella typhimurium* are arranged in two linked operons both under the negative control of the *hutC* regulator gene. Transcription requires the presence of the inducer urocanate, the product of the first enzyme histidase, and one of two possible activator proteins. The *hut* operon can be activated by the *crp* gene protein and cyclic AMP, but this will not be effective in the presence of a good carbon source although if the ammonia concentration is low the activator function may be taken over by glutamine synthetase. Thus, this complex control system makes it possible to use histidine as carbon and nitrogen source for growth when no alternative carbon or nitrogen sources are available but avoids unnecessary degradation of this valuable amino acid.

Comparative studies have shown that biosynthetic pathways are similar among different bacterial species but that the regulatory patterns vary a great deal, which suggests that they may have appeared at a much later stage of evolution when speciation had already arrived in the procaryotic world. The regulator proteins are required to interact with both low molecular weight effector molecules and specific regions of DNA. In the case of the *lac* repressor the total sequence of the protein is known and also the DNA sequence at which it binds to the *lac* operator [37,46], but so far this has not indicated the way in which regulator genes may have evolved.

One of the pathways which has been explored in great detail is that for the biosynthesis of tryptophan and Crawford [47] has recently reviewed comparative aspects of gene–enzyme relationships. In *E. coli* 5 genes are linked in a cluster and regulated as a single operon by an unlinked gene determining the *trp* repressor. In other species the genes for the tryptophan enzymes are distributed among several gene clusters; in *Pseudomonas* species one group is under negative control by the *trp* repressor, one enzyme is constitutive and the last enzymes of the pathway are induced by one of the intermediates. In this pathway we can see also an example of an enzyme, tryptophan synthetase, with two different subunits which are, of course, determined by two separate structural genes (two genes – one enzyme). These are also in *E. coli* examples of one gene determining

more than one enzyme activity. In some species the two subunits of anthranilate synthase are determined by separate genes but in *E. coli* the functions of the smaller subunit are carried by a protein which also has phosphoribosyl transferase activity and both these functions are determined by a single gene *trpD*(G). The two enzyme activities, phosphoribosyl-anthranilate isomerase and indoleglycerol phosphate synthase, are also carried by a single protein determined by gene *trpC*(F) (one gene – two enzymes). Fig.2 compares the gene–enzyme arrangements of the tryptophan biosynthetic pathway of *E. coli* with that for *P. aeruginosa*. *Salmonella typhimurium* resembles *E. coli* but other species have the same arrangement as *P. aeruginosa* of one gene for each enzyme, which suggests that in the past gene fusion may have occurred in some species but not in others.

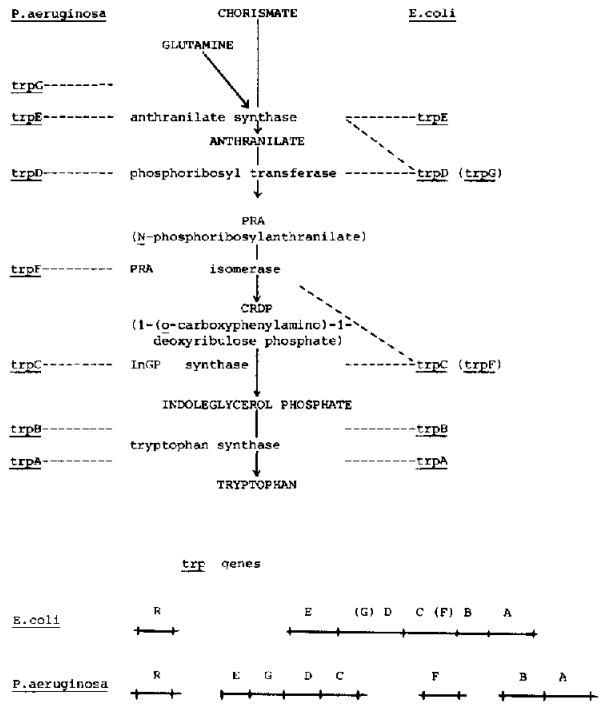

Fig.2. Biosynthetic pathway for tryptophan: Genes and enzymes. In *E. coli, trpE* determines the large subunit of anthranilate synthase; *trpD* determines the glutamine amido-transferase function of anthranilate synthase and phospho-ribosyl transferase; *trpC* determines both PRA isomerase and InGP synthase activities; *trpB* and *trpA* determine the two subunits of tryptophan synthetase. The tryptophan biosynthetic genes of *E. coli* form a single operon under negative control by the regulator gene *trpR*. The three *trp* gene clusters in *P. aeruginosa* are regulated independently and only *trpEGDC* genes are repressed by tryptophan [47].

Another biosynthetic pathway for which there is evidence for gene fusion is the histidine biosynthetic pathway. In *Salmonella typhimurium* a single gene *hisB* determines the structure of a protein which has both phosphatase and dehydratase activity and catalyses steps 7 and 9 of the pathway while in other organisms separate genes and proteins are required [48]. The possibilities of gene fusion appear not to have been exhaust for this operon since it has been possible in laboratory experiments to fuse two other genes *hisC* and *hisD* and to obtain a protein with both histidinol dehydrogenase and amino transferase activity [49]. The experimental formation of bifunctional enzymes by gene fusion offers very interesting opportunities for exploring enzyme structure and function.

Are enzymes still evolving?

One must admit that most enzyme evolution took place before 1876 and, if the micropaleontologists are to be believed, the really fundamental events took place more than 3×10^9 years ago. However, enzymes went on evolving as the evolution of species proceeded and there is no reason to suppose that all possibilities have been exhausted. We might be able to detect instances of evolution of enzymes during the last 100 years by examining examples of biological adaptions to environmental changes to see if any new enzyme activities have emerged during this time. The micro-organisms are the most likely candidates for this and there are at least two areas in which marked alterations in environmental factors can be discerned.

The first of these is the development of new drugs and antibiotics and we can ask whether the introduction of any of these has had significant effects on the enzymes of bacterial populations which have been exposed to them. Antibiotic resistant organisms are now commonplace and in

recent years it has been found that many bacteria owe their antibiotic resistance to genes carried on an extrachromosomal element or plasmid. Plasmids carrying antibiotic resistance may be freely transmissible through bacterial populations but although this increases the total number of antibiotic-resistant organisms, it does not necessarily involve any new enzyme activities. Shaw [50] has described a way in which resistance to the antibiotic chloramphenicol may have arisen by a change in enzyme activity. Among enteric bacteria chloramphenicol resistance is associated with plasmids carrying genes for chloramphenicol transacetylase. However, some wild type strains of *Proteus* and *Providencia* are chloramphenicol-sensitive but contain low levels of chloramphenicol transacetylase activity. From these sensitive strains it is possible to select resistant strains producing chloramphenicol transacetylase of much higher activity and with apparent K_M values for chloramphenicol some 20–40-fold lower than that of the enzyme from the sensitive wild type strain. It is possible that the normal substrate of the transacetylase from the sensitive parent is a normal cell metabolite and that a mutational change allows the enzyme to become effective in destroying chloramphenicol. The gene determining the *Proteus* and *Providencia* transacetylase appears to be chromosomal and Shaw suggests that such a chromosomal gene could have given rise to the plasmid-borne resistance genes now prevalent among the enterobacteria [50].

The other obvious environmental changes are due to the development by the chemical industry of very many novel organic compounds. Some of these like DDT are so poorly degraded as to constitute potential biological hazards but over long periods of time enzymes have evolved to break down many very complex organic molecules which are synthesised biologically. One could reasonably predict that novel compounds, which were not too dissimilar to natural products, could be degraded by known pathways and that one or more mutational changes in an enzyme might increase its capacity to deal with the novel compound [51].

Evidence that such changes in enzyme specificity are possible comes from following the stepwise evolution in the laboratory of an acetamidase to a phenylacetamidase in a strain of *Pseudomonas aeruginosa* [52]. Where an entirely new pathway is needed it is sometimes possible to 'borrow' enzymes by first removing the existing restraints of regulatory control. Mortlock and colleagues have obtained mutants of *E. coli* and *K. aerogenes* which are able to grow on pentoses and pentitols not utilized by the parent strains and in most cases use is made of existing enzymes which may undergo mutation to make them more effective in their new roles [53,54].

Together with comparative studies on enzymes of similar function which have evolved (divergently or convergently) in nature, these novel enzymes which have been derived by experimental evolution in the laboratory may in the next 100 years assist in the further elucidation of the relationships of enzyme sequence, structure and mechanism.

References

[1] Garrod, A. E. (1909) in: Inborn Errors of Metabolism, p. 80, Hodder and Stoughton, London.
[2] Bateson, W. (1909) in: Mendel's Principles of Heredity, p. 233, University Press, Cambridge.
[3] Carlson, E. A. (1966) in: The Gene: A Critical History, p. 27, Saunders, Philadelphia.
[4] Sturtevant, A. H. (1913) J. Exp. Zoo. 14, 43–59.
[5] Morgan, T. H. (1922) Proc. Roy. Soc. B 94, 162–197.
[6] Muller, H. J. (1922) Amer. Naturalist 56, 32–50.
[7] Knight, B. C. J. G. (1937) Biochem. J. 31, 731–737.
[8] Lohmann, K. and Schuster, P. (1937) Biochem. Z, 294, 188–214.
[9] Beadle, G. W. and Ephrussi, B. (1936) Genetics 21, 225–247.
[10] Beadle, G. W. (1974) Ann. Rev. Biochem. 43, 1–13.
[11] Beadle, G. W. and Tatum, E. L. (1941) Proc. Nat. Acad. Sci. USA 27, 499–506.
[12] Fruton, J. S. and Simmonds, S. (1958) in: General Biochemistry, 2nd Edn. p. 820, John Wiley, New York.
[13] Beadle, G. W. (1945) Chem. Rev. 37, 15–96.
[14] Troland, L. T. (1917) Amer. Naturalist 51, 321–350.
[15] Dubos, R. J. (1945) in: The Bacterial Cell, pp. 28, 187, Harvard University Press, Cambridge, Mass.
[16] Griffith, F. (1928) J. Hyg. 27, 113–159.
[17] Avery, O. T., Macleod, C. M. and McCarty, M. (1944) J. Exptl. Med. 79, 137–158.
[18] Tatum, E. L. and Lederberg, J. (1947) J. Bact. 53, 673–684.
[19] Jacob, F. and Wollman, E. L. (1958) Symp. Soc. exp. Biol. 12, 75–92.
[20] Cairns, J. (1963) J. Mol. Biol. 6, 208–213.
[21] Hershey, A. D. and Chase, M. (1952) J. gen. Physiol. 36, 39–56.
[22] Commoner, B. (1964) Nature 202, 960–968.

[23] Watson, J. D. and Crick, F. H. C. (1953) Nature 171, 964–969.
[24] Fildes, P., Gladstone, G. P. and Knight, B. C. J. G. (1933) Brit. J. exp. Path. 14, 189–196.
[25] Dubos, R. J. (1940) Bact. Rev. 4, 1–16.
[26] Gale, E. F. (1943) Bact. Rev. 7, 139–173.
[27] Jacob, F. and Monod, J. (1970) in: The Lactose Operon, pp. 1–4, Cold Spring Harbour Laboratory.
[28] Monod, J., Cohen-Bazire, G. and Cohn, M. (1951) Biochim. biophys. Acta. 7, 585–599.
[29] Rotman, B. and Spiegelman, S. (1954) J. Bact. 68, 419–429.
[30] Hogness, D. S., Cohn, M. and Monod, J. (1955) Biochim. Biophys. Acta 16, 99–116.
[31] Jacob, F. and Monod, J. (1961) J. Mol. Biol. 3, 318–356.
[32] Hayes, W. (1953) J. gen. Microbiol. 8, 72–88.
[33] Hayes, W. (1953) Cold Spr. Harb. Symp. quant. Biol. 18, 75–93.
[34] Jacob, F. and Adelberg, E. A. (1959) Compt. rend. séances Acad. Sci. 249, 189–191.
[35] Gilbert, W. and Müller-Hill, B. (1966) Proc. Nat. Acad. Sci. USA 56, 1891–1898.
[36] Gilbert, W. and Müller-Hill, B. (1967) Proc. Nat. Acad. Sci. USA 58, 2415–2421.
[37] Gilbert, W. and Maxam, A. (1973) Proc. Nat. Acad. Sci. USA 70, 3581–3584.
[38] Nirenberg, M. W. and Matthaei, J. H. (1961) Proc. Nat. Acad. Sci. USA 47, 1588–1602.
[39] Benzer, S. (1957) in: The Chemical Basis of Heredity (McElroy, W. D. and Glass, B. eds.) pp. 70–93, Johns Hopkins Press, Baltimore.
[40] Sarabhai, A., Stretton, A. O. W. and Brenner, S. (1964) Nature 201, 13–17.
[41] Yanofsky, C., Carlton, B. C., Guest, J. R., Helinski, D. R. and Henning, U. (1964) Proc. Nat. Acad. Sci. USA 51, 266–272.
[42] Englesberg, E., Irr, J., Power, J. and Lee, N. (1965) J. Bact. 90, 946–957.
[43] De Crombrugghe, B., Chen, B., Gottesman, M., Pastan, I., Varmus, H. E., Emmer, M. and Perlman, R. L. (1971) Nature New Biol. 230, 37–40.
[44] Zubay, G., Gielow, L. and Engelsberg, E. (1971) Nature New Biol. 233, 164–165.
[45] Magasanik, B., Prival, M. J., Brenchley, J. E., Tyler, B. M., Deheo, A. B., Streicher, S. L., Bender, R. A. and Paris, C. G. (1974) Current Topics in Cellular Regulation 8, 119–138.
[46] Beyreuther, K., Adler, K., Geisler, N. and Klemm, A. (1973) Proc. Nat. Acad. Sci. USA 70, 3576–3580.
[47] Crawford, I. P. (1975) Bact. Rev. 39, 87–120.
[48] Ames, B. N., Goldberger, R. F., Hartman, P. E., Martin, R. G. and Roth, J. R. (1967) in: Regulation of Nucleic acid and Protein synthesis (Koningsberger, V. V. and Bosch, L. eds.) pp. 272–287, Elsevier, Amsterdam.
[49] Yourno, J., Kohno, T. and Roth, J. R. (1970) Nature 228, 820–824.
[50] Shaw, W. V. (1974) Biochem. Soc. Trans. 2, 834–838.
[51] Dagley, S. (1972) in: Degradation of Synthetic Organic Molecules in the Biosphere pp. 1–16, 336–347. National Academy of Sciences, Washington.
[52] Clarke, P. H. (1974) in: Evolution in the Microbial World (Meadow, P. M. and Pirt, S. J. eds.) Soc. Gen. Microbiol. Symp. 24, 183–217. University Press, Cambridge.
[53] Leblanc, D. J. and Mortlock, R. P. (1971) J. Bact. 106, 90–96.
[54] Oliver, E. J. and Mortlock, R. P. (1971) J. Bact. 108, 287–293.