Invent or Discover: the art of useful science

www.alanrwalker.com

Synthetic insulin

'As if science was too serious a business to be left to grown-ups.'

(François Jacob)

Squeezed up against the southern slopes of the San Gabriel Mountains, by Los Angeles, lies the town of Duarte with its City of Hope National Medical Center. The hospital was founded in 1928 as a tented sanatorium to care for patients with tuberculosis and over the next fifty years it evolved into a world famous location for research and treatment of cancer. In 1979 one of the researchers there, Keiichi Itakura, applied for a patent concerning 'recombinant DNA'. Few of his colleagues would have appreciated what he wanted, but amongst them Arthur Riggs had applied for a paired patent on 'microbial polypeptide expression'. Within the depths of the dry legalistic texts stirred the creation of an incredible transformation in medicine and industry. Here were instructions, both specifically technical and of general application, to construct the genes responsible for human hormones such as insulin, then manufacture these hormone by using broths of bacteria.

How to *construct* human genes? Yes, build them up them from small chemical units on the laboratory bench. Manufacture them using *bacteria*! Why not, the researchers would have answered? Bacteria can be tricked into producing a protein based on the coding from a gene that has been inserted into them; a synthetic human gene for example. Since bacteria reproduce extremely rapidly and can be grown in industrial vats using simple nutrients, they form vast numbers of tiny factories to pump out the hormone.

Several years before those patent applications, Herbert Boyer up in San Francisco, called Riggs with a proposition. Boyer had conferred with Robert Swanson from a local venture capitalist firm. Swanson wagered he could make money from molecular biology, which was burgeoning from the combined discoveries about DNA and how proteins are synthesized. Boyer had already made his name, with Stanley Cohen, as co-inventor of the first crucial method, called gene splicing. Whilst Cohen was happy to continue in academia, Boyer was lured by exciting and lucrative novelties. He was convinced he could exploit these new ways to manipulate DNA and make something for sale. Something in urgent demand: say, an improved insulin for people with diabetes. Boyer simply asked Riggs if he could make the gene for insulin. Riggs and Itakura were already trying for a simpler gene for the hormone which prevents gigantism, but the usual funders thought this all too ambitious. Instead Swanson and Boyer staked their own cash to start with this hormone, empowering themselves toward insulin by creating a company. They named it Genentech.

* * *

The most iconic photograph in biology shows James D. Watson and Francis H.C. Crick in a corner of a sparse laboratory of the University of Cambridge in 1953, posing proudly in front of their tall spindly model demonstrating the structure of DNA with metal rods and plates. These two outsiders, working on a semi-official, unfunded project, with data scrounged from other people who had the skills to collect it,

were able to solve in a few intense years the structure of the molecule that carries genetic information. What that icon and all the breathless publicity gave no hint of was that they had found only the easy part of 'the secret of life' as it became known.

The hard part occupied far more people and time as it continued to torment the minds of these two pioneers and many more great talents. The question was: how are proteins synthesized? How are they put together in the body, starting with small simple molecules to be constructed into huge molecules of precise function? How does the information in DNA translate into the enzymes and hormones, muscle and tendons that make up a living body? A typical answer given in the early 1950s was that proteins are made by enzymes. Since enzymes themselves are proteins, what makes them – surely not more proteins, and so on into infinite regression?

In one sense the researchers on the structure and function of DNA were faced with the same mystery that had puzzled early geneticists such as Thomas H. Morgan at Columbia University in New York during the 1910s to 1920s. He and colleagues set out to prove that new species are formed by changes, or mutations, in a gene caused by influences such as a pulse of energy from natural gamma-rays. They bred fruit flies, the famous *Drosophila*, in glass jars for speed of experiments, then induced mutations in the flies's genes using X-ray machines. They proved the classic genetics of Gregor Mendel is due to a mechanism of inheritance that is precisely particulate. This transformed their field into the most rigorous branch of biology through the discipline of mathematics – slicing through the muddle of variation that is typical biological data.

However, within this wide success hid a brooding mystery, a question they delicately steered past. A black box. What was inside the box of genes within the nucleus of living cells? If genes really do carry information, how is it translated into physical structure? As Morgan himself wondered in his Nobel Prize speech of 1934: 'There is no consensus of opinion amongst geneticists as to what genes are – whether they are real or purely fictitious – because at the level at which the genetic experiments lie, it does not make the slightest difference whether the gene is a hypothetical unit, or whether the gene is a material particle.'

Times were ripe by the 1930s for proper scientists to join classical biologists in this genetics game. Physicist Max Delbrück and the chemist Linus Pauling became pivotally influential in tackling biology at the level of atoms and photons, specially when they paired up to write together. Delbrück originally studied astrophysics in the University of Göttingen. He had been an assistant to Lise Meitner, who was working in Germany on radiation from uranium; the same woman in the team that discovered nuclear fission. Delbrück came to work at the California Institute of Technology in Pasadena in the 1930s and then at Vanderbilt University, Tennessee. He had intended to work on genetics using *Drosophila* flies at Caltech, but a colleague pointed him to the precision gained by working with the simplest living things. Single celled and uniform microbes, rather than insects that consist of millions of varied cells.

Delbrück's talents as a physicist were at home with new opportunities and collaborators. He expanded his network to include Salvador E. Luria of Indiana University. Luria had originally trained in Rome as a medical doctor specialising in radiology and at Indiana became fascinated by Delbrück's ideas on the significance of mutations caused by penetrating radiation. Delbrück and colleagues had shown in 1935

that mutations induced by X-rays could occur at great dilution of the dose. They could best explain this as an effect of quantum physics. That is, a single X-ray photon, a particle of radiation, might smash into a gene and cause a step change in its structure. This would alter the form or physiology of the organism growing according to the information from its genes, including the mutated gene.

Delbrück and Luria estimated the size of the genetic structure affected by one photon as equivalent to a cube containing only one thousand atoms. At this minute scale, the randomising effect of heat motion at room temperature is huge. What could possibly resist such harmful agitation to remain able to transmit information reliably? When Luria posed questions like this he also hoped that his first doctoral student, James Watson, would carry on to provide some answers.

More microbiologists joined in collaboration with Delbrück and Luria in 1943, forming the nucleus of what became known as The Phage Group. Bacteriophages are literally eaters of bacteria; viruses that invade and replicate within bacteria and a reminder of the ditty about big fleas having 'little fleas upon their backs to bite ′em and so on ad infinitum.' The Group used them as tests for the presence of mutations in bacteria that made the bacteria resistant to the phages. Their central question was: how is genetic information transmitted reliably, without error? This group thrived under Delbrück's leadership and grew to about forty members worldwide. Seminars, writing weekends, a newsletter – Delbrück was a thought provoker determined for the group to answer their big question. Enjoying yourself was mandatory for members as they camped and hiked in the Sierra Nevada mountains, spurred on upward under the spell of Delbrück's vitality.

The group gained access to the facilities of Cold Spring Harbor Laboratory for research during the summers. The rustic village of Laurel Hollow on the north shore of Long Island is the charming setting for this private centre. Originally funded by the Carnegie Institution of Washington, it gained notoriety as also being the base of the Eugenics Record Office, until the racist nature of that project was revealed and the Carnegie Institute stopped such pseudo-science on their site. Now it is one of many research laboratories financed by the National Cancer Institute, thriving with new labs rising amongst the tall broad-leaf trees. For the Phage Group in those troubled times it was a tranquil haven in which to walk along the foreshore, talk, think and to prise open the lid of Morgan's black box of genetics.

The times might have been favourable for such collaborations in the peace and sunshine of California and Long Island, but elsewhere obscure research problems were the least of worries for researchers. Many of them were fleeing from deadly discrimination and strife in continental Europe. Both Delbrück and Luria found themselves classed as enemy aliens, but at least they were left in the peace of their laboratories, concentrating on their bacterial culture plates and equations. By 1943, they demonstrated, in what came to be famous as the Luria-Delbrück experiment, that mutations arise spontaneously in bacterial colonies, at random, and before any selection pressure for resistance to bacteriophages have been applied. Thus, the selection of resistant strains that does occur is by natural selection of the sort that Charles Darwin had proposed.

Delbrück explained the phenomenon mathematically – biology developed into fertile ground for physicists once it was reduced to a level explicable in terms of things as small as atoms and photons. The revelation that encouraged so many talents, with their different techniques, into this field was that biology

could now be done as chemistry, with the ability of that science for precise quantification and one-to-one causal relationships. Searches for simple rules underlying biological complexity were the way forward. Nevertheless, early molecular biologists confronted a barricade of technical problems. No longer could a huge molecule such as DNA, so easily obtained in large amounts in semi-crystalline form, be observed directly using the sturdy apparatus of X-ray machines. The physicists could carry on theorizing or filling in empirical details of the structure of DNA, but more and more they would have to rely on the delicate and subtle techniques of biochemists and microbiologists, or even try their own hand at such craft.

Where to start? Ask smaller questions; the simpler and more accessible to existing techniques the better. Chemists already had contributed much to the understanding of DNA, or deoxyribonucleic acid. As early as 1931 Phoebus Levene at the Rockefeller Institute in New York, wrote an authoritative textbook on the nucleic acids of living cells. He had named the basic chemical units of these acids as nucleotides, comprised of a phosphate group – linked to a sugar – linked to a base. There were four bases, as two pairs: adenine and guanine, cytosine and thymine. They all seemed to form a simple molecule in a loose watery colloid. But Levene's account of nucleic acids contained no mention of any biological function. Nevertheless, understanding of this chemistry and its link to proteins grew: Alexander R. Todd at the University of Cambridge clearly demonstrated that the nucleotides were chemically linked together in a simple and repetitive chain capable of forming an extraordinarily long thin thread.

One obvious question researchers now asked was: what is the code for the smaller sub-units then known to constitute proteins? These were the amino-acids, twenty of them. Answer that and the route to answering the question of how are amino acids assembled to make proteins might reveal itself. The first amino acids were discovered in the early 1800s and biochemists and geneticists had long been busy with the relationship between amino acids and proteins. Vernon Ingram at the Cavendish lab for example, had followed the lead of Linus Pauling and colleagues in California who showed in 1949 that the disease called sickle cell anaemia is caused by a defect at the molecular level. It is inherited following simple genetic rules through a specific alteration in the structure of haemoglobin, the red pigment of blood. Ingram demonstrated that this is due to a mutation which causes the substitution of a single amino acid. This led researchers to propose a relationship between the nucleotide bases in DNA and the formation of amino acids.

*

Another molecule similar to DNA was known: ribonucleic acid or RNA. How it related to DNA in the same cell was unknown. James Watson expressed his opinion as early as 1952 that it is the intermediary between DNA and protein. The problems early molecular biologists found with RNA were multiple however: it yields little information when probed with X-rays. It occurs in several different chemical forms, it is found spread throughout the watery ground material of the cell, and the nuclei of some viruses are composed of RNA rather than of DNA.

An alternative approach to RNA was needed, possibly from people not impeded by such knowledge of difficulties. The Oak Ridge National Laboratory, Tennessee, was better known for its major role in improving the uses of nuclear fission than for research on proteins. Nevertheless, a great place to perform experiments requiring the dangerously radioactive forms of phosphorus and other elements that were needed. Elliot Volkin and Lazarus Astrachan worked there in the mid 1950s. Despite severe technical

difficulties they demonstrating that when phage replicates in bacteria the phage DNA increases greatly and continuously but the RNA only increases for a short time and then stays steady. They found their results puzzling; they could discern no clues to the functional nature of RNA. Most researchers in the mainstream at the time also found these results anomalous, the sort of thing to be expected from outsiders. Fortunately, a few of them squirrelled away the information as a piece to offer to the puzzle later on.

Francis Crick found irresistible the intellectual challenge of what soon became known as the coding problem. James Watson, as Crick's original sounding board, had returned to America but soon Crick found himself sharing an office at the Cavendish with Sydney Brenner. The pair operated a powerful partnership for the next twenty years. Brenner was fresh with a doctorate from the University of Oxford. He had arrived in Britain from Germiston, which sits amongst huge heaps of yellow waste dug from the lucrative goldmine deep below, just east of Johannesburg, South Africa. Not that the Brenner family shared much of the wealth. Recent immigrants from Eastern Europe, they lived in two rooms behind the cobbler shop his father ran, reeking with the smell of freshly tanned leather. By the age of fourteen Brenner gained a scholarship to the medical school of the University of Witwatersrand and then, because he would have been still too young to graduate and practise after the six year course, he opted for an intercalated degree in anatomy and physiology. At last, Brenner the self-taught naturalist, could study in a proper laboratory. He developed an abiding interest in preparing and observing specimens with a microscope. Such skills resurfaced when he moved on from protein synthesis to ask questions about how an organism can possibly form itself, starting from one cell with its set of genetic instructions, into a functioning animal. For that he chose a tiny transparent roundworm, *Caenorhabditis elegans*, and initiated a whole new school of embryology with this easily cultured and observed, even charming, model of animal form and function. That of course is another story, larger and still unfolding.

Crick's first attempt at the coding problem, was to propose a code that could produce 20 amino acids, starting from the total combinations of 4 nucleotide bases and assuming 3 nucleotide bases for each amino acid: thus 4 to the power of 3 equals 64. John Griffith, a mathematician colleague, found a neatly coincidental 20 codes; the 'magic number' as Crick described it for the 20 amino acids. How, they asked, could this code be read as a sequence of 3 letter words without any obvious starts and stops, or punctuation as it were? They called their proposed system a 'code without commas'.

It died young: too much magic and not enough biology. Brenner stringently scrutinized these propositions against theoretical criteria for the general nature of the coding problem. Crucial to Brenner's development of ideas was the work of nearby biochemists, Frederick Sanger and H. Tuppy, in the Department of Biochemistry at the University of Cambridge. In the 1940s Sanger had taken up a well known and important protein, insulin, as a suitable subject for revealing the full sequence of amino acids constituting one protein.

Sanger branched out with a radically new method from that used by the well established method of structural biology. The classical method was based on directing X-rays through materials to produce intricate patterns on photographic plates as the rays were diffracted into many beams of radiation by the crystalline layers within the material. The extremely taxing and visually intuitive task of translating these two dimensional patterns of spots into a three dimensional model, built of bits of metal and plastic, was the work of various schools of X-ray crystallography. A strong school operated in Britain where two of its

leaders were Max F. Perutz and Dorothy M. Hodgkin. They both spent a lifetime of hard labour deciphering the full three-dimensional atomic structure of haemoglobin and insulin respectively, quietly trusting all the while their knowledge would be of direct practical use.

Sanger and his colleagues were content with two dimensions instead of the vast complexities of the three dimensional structure of insulin. Sanger, an unpretentious man of intricately skilled method, approached his biological molecules by asking of them soluble questions. Insulin was an obvious choice as one of the few proteins then readily available in pure form. Sanger's hunch about the suitability of insulin was accurate. He found the amino acids in two neat and parallel lines of amino acids with small interconnections. Neat that is when drawn in two dimensions only. This was far simpler information than the intricate folding and re-folding of these lines into the three dimensional structure of the functioning hormone. Nevertheless it was exceedingly illuminating. Sanger and Tuppy's results revealed to those working on the coding problem that amino acids could string together in any sequence. It seemed that the joins must be in the correct order to constitute a specific and functional protein, but the evidence suggested the fundamental rule that any amino acid could join to any other.

*

One the ironies of this story is that Sanger worked separately from the crystallographic studies on insulin. The techniques he used were novel for his specific purpose and of no easy connection with methods of the structural biologists. Often there is a separation of researchers into groups based as much on the techniques and apparatus at hand in the laboratory they first start work in, as on curiosity about various aspects of the same items of the natural world. Certainly work in this area could be divided between those of the structural school typified by Linus Pauling, concerned with shapes of molecules, and the informational school typified by Max Delbrück, concerned with codes.

Hodgkin and Sanger can likewise be classified apart. The discovery of the structure of DNA centred around the skills of chemist's crystallographs and their mathematical and visual skills to interpret such arcane data. The tightly knit school thrived within a small area over one generation. The crystallographic studies of insulin and haemoglobin were a deeply influential component of that school which sheltered and nurtured those bold enough to apply the methods to DNA. These pioneers were Rosalind E. Franklin and Maurice H.F. Wilkins, toiling in their lab at King's College, London, and providing Watson and Crick with crucial crystallographic data.

About half way through Max Perutz's 24 year search for the three dimensional structure of haemoglobin he published in 1949 a long paper on the progress he and two assistants had made. He despaired: 'The photographing, indexing, measuring, correcting and correlation of some 7000 reflexions was a task whose length and tediousness it will be better not to describe.' He concluded the paper with: 'It remains to be seen whether the X-ray analysis of haemoglobin itself can be carried further or whether future progress lies in the analysis of simple proteins of smaller molecular weight which are now being studied by several workers.' Perutz battled on for many decades, defending his approach of sticking to the facts, as if the facts could speak to researchers with explanations of their own significance. He argued against the style of research advocated by logicians such as Karl Popper: simplify your question then formulate it as a

hypothesis to be tested by experiment. The *art of the soluble*, as the immunologist Peter Medawar titled it. Despite his own pleas, Perutz in this same paper reveals himself imaginatively hypothesising about protein structure.

Whence the origin of these hypotheses, these models both mental and literal? Karl Popper suggested in his *Logic of Scientific Discovery* that whilst the work of scientists consists of proposing and testing hypotheses, the act of conceiving them is beyond logical analysis: '. . . looking at the matter from the psychological angle, I am inclined to think that scientific discovery is impossible without faith in ideas which are of a purely speculative kind, and sometimes even quite hazy; a faith which is completely unwarranted from the point of view of science, and which to that extent, is "metaphysical".' This idea of the core procedure of scientists has been extended by David Deutsch in his book the *Fabric of Reality*. He argues that hypotheses and theories are a form of virtual reality, mental copies of external physical reality, that despite remaining imperfect, nevertheless are an objective form of reality, as activities of matter and energy within human brains.

*

Crick soon thrust on from Sanger's new understanding with a tour de force, a theoretical paper titled with his characteristic bravura: *On Protein Synthesis*. Crick set out eight hypotheses covering the coding problem, the role of RNA, the site of protein synthesis and the relationship between information contained in DNA and in protein. Crick called the latter a central dogma (something without any supporting evidence) which stated that the information went only one way, from DNA to protein. Furthermore, he proposed an adaptor molecule, probably one for each amino acid, which aided the transfer of information from DNA at the site of synthesis of the proteins. Researchers assumed these sites were the microsomes (small bodies within the cell, soon to be called ribosomes) and they contained RNA strands as templates. This RNA had the genetic information needed for synthesis of a particular type of protein. Soluble RNA in the cytoplasm would also be needed. So this ineptly named central dogma was actually a plain hypothesis, but still with a crucial role to play.

The researchers were numbed into frustration by scarcity of experimental methods. Many of them proposing the grand ideas were not adepts at the craft skills needed at the laboratory bench. The molecules that worked to synthesize proteins evolved within the physical structures of tiny living cells. They need membranes to adhere to, a favourable balance of acidity and alkalinity, and nearby energy transformers. They need information from the cell's nucleus. Devising experiments to enter into these ephemerally delicate cells seemed impossible. These RNAs, adaptors, microsomes or ribosomes were to prove a secret world in which to chase after flittering butterflies. Or, as Sydney Brenner expressed it, as a rueful South African exposed to the damp climate of the low Cambridgeshire plains: 'What I could see in them was fog – thick fog, wet, cold, silent. The nonscientist does not often break through the sharp-edged view of the world that science presents to sense the whirling gray uncertainty in which that confidence must take shape.'

*

The French chemist turned bacteriologist who invented pasteurisation for the preservation of wine and milk, the vaccine against rabies, and much else, founded an institute in 1888 to continue his work on infectious diseases. Louis Pasteur became so famous and admired for his work that public subscriptions

and vaccine sales were ample for the institute to operate autonomously from government and university systems. Free from such bureaucracy, it came about that in the early 1950s a trio of researchers at the Institut Pasteur flourished in this functionally anarchic incubator of ideas on protein synthesis. They developed a vigorous node in the spreading network of labs working on the new biology, working in a bustling lab in a corner of the attic. The original building, with its grey mansard roof above yellow stone and pink brick, remains a fine contribution to the elegance of Paris. If you visit it on the Rue du Docteur Roux make sure you arrive at a time to enjoy the fascinating museum within.

Microbiologist André M. Lwoff, biochemist Jacques L. Monod and bacteriologist François Jacob found themselves bound by the terrible times they had lived through. Lwoff had been a departmental head at the institute since the late 1930s and during the enemy occupation he fought for the resistance, setting up intelligence gathering networks under cover of his regular work and contacts across the city. He used his apartment to shelter stranded Allied servicemen. Jacob escaped the invasion, sailing over to London with a boatload of Polish troops. There he joined the Free French Forces, thence as a medical officer on the campaigns in North Africa, eventually returning to France at Omaha Beach, Normandy. Monod volunteered for the armed branch of the resistance; he signed up to the communist party to gain entry to the inner circle planning sabotages, soon rising to chief of staff of operations. Monod had grown up in an artistic family and developed his musical talents to conduct as well as play. When on a Rockefeller scholarship at Thomas Morgan's new genetics lab in California, a musical group for Bach concerts that Monod set up was such a success the group offered him salaried position as their full time conductor. Back at the lab, his colleagues managed to persuade Monod that he could manage being a scientist and musician, but probably not the converse.

Jacques Monod and Linus Pauling dramatically contrast the French and American styles of how to go about the business of research, but in one period they found themselves bound in more ways than their interest in genetics. Pauling, an outspoken pacifist, was temporarily denied by the US State Department a passport to travel to Europe during the crux of the DNA model building. Monod, a documented former communist, was denied entry to America during those years of Senator Joseph McCarthy's Un-American Activities Committee. Strange days.

In Paris the big question around the big lunch table up in the attic was how is the function of genes regulated during protein synthesis? How is a cell induced to start the biochemical expression of characteristics encoded by the genes? Long known was that many single types of cell can produce different proteins during many phases of their working life, or they can produce one protein at varying rates. Was protein synthesis switched on by inducers, or was it switched off by repressors? Did both things happen at different phases of the cell's activity? Monod had spent many years studying enzyme production in one of the simplest organisms, bacteria called *Escherichia coli*. He devised a way to manipulate genetically the bacterium's system for producing an enzyme, beta galactosidase, which controls its use of a type of sugar.

André Lwoff had good contacts in America, exploiting the attraction of working in Paris and generous funding in Marshall Plan style. Arthur Pardee and Monica Riley, a doctoral student, came over and produced two crucial papers; the first became famous under the nickname of the PaJaMo Experiment. The team used the beta galactosidase enzyme of *E.coli* in a series of mating crosses of mutant strains.

Additionally, they inactivated the gene in some of the crosses by use of rays from radioactive phosphorus that they introduced. First they needed to differentiate between an inducer or a repressor that would control the activity of some messenger between the gene for the enzyme and the synthesis of the enzyme. Then they needed to differentiate between the gene acting directly as a template for synthesis of the enzyme, or the gene forming stable intermediates (such as ribosomes) on which the enzymes are synthesized.

Their results were confusing. The bacteria without the functioning gene for beta galactosidase would – when the gene was introduced by mating with fully functioning bacteria – produce the enzyme at maximal rate. But this was incompatible with the proposition of a stable intermediate being responsible for synthesis of the enzyme. The group struggled to rationalize their results to fit their idea that ribosomes were self-contained factories possessing the information needed for assembly of at least one protein each. The messenger remained a good idea but their mental image of inducers and repressors was not helping to design better experiments. After so much work Riley became resigned at the last sentence of her paper of 1960: 'These experiments therefore appear to define an interesting dilemma.'

Interesting indeed. The entire network was gripped by a collective neurosis as they tackled this problem. People talked in despairing and derogatory terms about ribosomes. Thick cold fog continued to swirl. After a visit to London in April 1960 for a symposium, François Jacob travelled on up to Cambridge. A small group gathered in Brenner's office where they brainstormed the problem in a babble of interrogative cross talk. Then someone remembered the anomalous results of Volkin and Astrachan from four years ago.

All of sudden this scrap of information, this piece for the puzzle fitting into those from the Institut Pasteur, flashed a possible answer. Brenner and Crick leapt to their feet and began: 'to argue at top speed in great agitation. A red faced Francis. A Sydney with bristling eyebrows. The two talked at once, all but shouting. Each trying to anticipate the other. To explain to each other what had suddenly come to mind.' Could it be that the information was carried to the stable ribosomes by an unstable intermediate, a messenger RNA? This would inform every ribosome already resident in the cell that now was the time to put together a string of amino acids. The sequence of amino acids would accord with the code of nucleotides about to pass through them – a recording tape through a reading head!

In those days before digital sound recorders they were blessed with a potent metaphor to carry their creative thoughts. Brenner was inspired by the ideas of John von Neumann on a type of self reproducing automaton. Biological evolution probably requires a separation of what he called a universal constructor from the information needed to construct it. When errors occur in copying this information, as mutations, then variants of the self reproducing automaton are created, and those that are viable become the raw material for natural selection to work on.

Francis and Odile Crick had planned a party at their house that evening. Their parties were famously lively but Brenner and Jacob just grabbed some food and huddled in a corner to make plans and discover their coincidental luck of separate invitations to Caltech that June. There they set to work with Matthew Meselson to test the hypothesis that ribosomal RNA was not an intermediate carrier of information from gene to protein. Instead, they proposed that ribosomes are structures that receive genetic information from

the nucleus in the form of an unstable messenger. Cultures of *E.coli* were maintained in media with heavy isotopes of nitrogen, carbon and phosphorus. Then they infected bacteria with bacteriophage and immediately transferred them to medium containing the normal or light isotopes, so that all constituents synthesized from then on would be lighter. They used centrifugation to reveal the data from the suspension of ground up bacteria at the end of the experiments. The papers were published as a pair in *Nature*. Messenger RNA it was to be.

The description in the published paper makes it all sound so neat and precise but Jacob revealed more in his memoir: 'We found ourselves lying limply on a beach, vacantly gazing at the huge waves of the Pacific crashing onto the sand. Only a few days were left before the inevitable end. But should we keep on? What was the use? Better to cut our losses and return home. [. . .] Suddenly, Sydney gives a shout. He leaps up, yelling, "The magnesium! It's the magnesium!"'

Crick at this time not only carried out his role as central provocateur – he and Brenner decided that if they wanted more answers to the parallel problem of the genetic coding system then Crick would have to roll up his sleeves and join Brenner in the lab. With the help of two colleagues they borrowed some bench space and embarked on an exhaustive series of experiments with the usual *E.coli* infected with phage. They followed closely the methods recently developed by Seymour Benzer, at Purdue University, for revealing the fine structure of genes. They studied a specific gene of their bacteriophage; mutations can be induced in this gene by treatment with a chemical dye. Confidently the pair concluded that the code was indeed in triplets, that it does not have commas to separate each triplet, and that there can be more than one triplet code for a particular amino acid. In other words, the code incorporates redundant capacity. Some amino acids have six different codes, leucine for example; whilst methionine has one triplet and phenylalanine has two to code them. Brenner and Crick predicted that the full code for all twenty amino acids would be worked out promptly. The information they supplied more or less cracked the code and their work became known as the Crick-Brenner Experiment.

'At last' they all sighed. Revealed was messenger RNA which temporarily interacted with DNA to become the tape of code representing a particular active gene, and then diffused out to the cytoplasm. Revealed was the way information from the nucleus came in triplets of nucleotides. Visible in the fabric of the cell, as tiny blobs under an electron microscope, were numerous stable ribosomes with their own type of RNA combined with protein as a structural platform. There were varieties of transfer RNA, one for each amino acid, which carried the amino acids to the ribosomes to form proteins. Many more details and confirmations were needed to know the full process of protein synthesis. But the basics were traced as the intermeshing of tenuous connections between bewildering pieces of knowledge, a canvas in which to paint a radically fresh picture.

Confirmation of some essential parts of this story of protein synthesis were meanwhile being worked on by a pair that others in the network had scarcely heard of, working at the National Institutes of Health in Bethesda, Maryland. Marshall Nirenberg had travelled to Moscow, for the vast meeting of the International Congress of Biochemistry in 1961. He possessed hot news of an experiment that Heinrich Matthaei and he had recently completed. They were not members of the clique who assumed intellectual possession of this problem; their new lab was then of unknown status in this field. Thus, Nirenberg's allotted few minutes of presentation fell before a small audience who were not sensitised to grasp its

meaning. Fortunately for all, Matthew Meselson was present and pricking up his ears to absorb the stunning implications. He rushed to inform Francis Crick and Max Perutz, who squeezed in a repeat at the end of the congress for a big audience alerted on the fervid gossip-mill.

This experiment was a simulation of some of the basic processes of protein synthesis in a test tube system derived from extracts of *E.coli* bacteria, with amino acids as building blocks, and a source of energy. To this Nirenberg and Matthaei added a synthetic polynucleotide in the form of polyuridylic acid, which can be represented by the genetic code of three uridine bases. This resulted in the incorporation of phenylalanine into a very simple protein: polyphenylalanine. For its time, this was the work of technical wizards who reproduced sufficiently the intricate interactions of the cell to probe it deeply. Although they posed more questions equally difficult as the original one, they were the first to corroborate the ideas emerging from the network on protein synthesis.

There remained vast detail to be brushed onto this canvas, but sufficient understanding was now available to attract a matching number of researchers eager to do biology in the way chemists experiment. They knew that to work with defined starting materials that interact on a direct molecular basis, without variation and ambiguity, was the only way penetrate the mysteries of cell biology. By 1966 the field was sufficiently advanced for researchers to hold one of the regular Symposia on Quantitative Biology, at the Cold Spring Harbor lab, on the topic: The Genetic Code. For some who had been at the centre of this tumultuous turn-around it was time to move on to fresh questions: how does the brain work; how does an egg develop into an adult, even if just an adult worm? For most of them there was a feeling that molecular biology was just beginning. The basics were established; there was a vast amount of detail still needing elucidation and the practical applications allured as gold nuggets brought up into sunlight.

*

The structure of DNA was a problem solved by a group of people that could fit around a laboratory bench to scrutinise radiographs of nucleic acid. But the problem of protein synthesis was a vast maze spreading across the inner space of microscopic living cells. A jigsaw puzzle in which the pieces were scraps of information, tiny parts of an exceedingly complex picture. Of course, none of the players had sight of any picture printed on the top of the box in which the puzzle came packed. They did not know the appearance of what they sought. Each player had a handful of pieces that they could try to fit to others, or exchange with other players, and all pieces were displayed face-up. No one person was in charge; how could that work? It would provide only the ability of that single person. Instead, the individual initiatives of many players, of varied but high talent, spontaneously coordinated toward the emergence of the big picture. A patch of plain sky-blue here; a group of speckled green stuff over here; who wants to swap for these red stripey pieces? A gorgeous flower garden emerges.

Michael Polanyi proposed this analogy for scientific research in general; as an X-ray crystallographer himself he contributed a small piece to that picture. Moreover, Polanyi had good cause to understand the strange ways of science. Between 1914-16 he published a series of papers proposing a mechanism of the adsorption of gas atoms and molecules onto solid surfaces. Another fifty years passed for his ideas to be accepted, and that understanding was fitted into the development of the lithium ion battery.

Players round a big table all helping each other to solve a big puzzle – relaxed happiness. Read James Watson's autobiographical account of the race to decipher DNA; happy for some but never relaxed and always competitive. All these players collaborated and they all competed. They collaborated in groups that competed with other groups. Individual researchers fiercely strove for autonomy by competing with everyone else to have something new to offer and thus be allowed into the game.

Every one of these players knew well that the puzzle they aimed to solve was finite. There were only so many pieces of the jigsaw that nature had to reveal until the researchers could satisfy themselves with an explanation that would enable them to proceed to new problems. Objective physical reality, out there in the world, is comprised of infinitely many facts, each of which is singular. Nature can only be discovered, it cannot be invented in the way an artist or music composer creates something unique every time they complete a work. Hence, scientists must compete to be first. Like athletes in a race, who must also collaborate by performing at the same stadium, at the same time, to the rules they have collectively devised.

This competitive collaboration is aptly termed an invisible college and is not so much an analogy as a literal collective of researchers working on a particular problem. In fact the physicist Robert Boyle coined the term, about 1646, to describe a group of researchers in Oxford who later formed the Royal Society in London. Their self-elected college was invisible amongst the grandly formal buildings of the university Colleges for the same reason they remain invisible to commercial economics. Invisible colleges operate as networks of informal contacts, with tacit rules and routes of collaboration below the radar of College administrators or company directors, although discretely remaining within institutional norms. The barter economy of an invisible college is based on gifts of information in exchange for recognition in the form of acceptance within the invisible college.

Derek J. de Solla Price is the physicist and commentator on science to thank for this insight. He studied an information exchange for swapping preprints of research papers on a particular problem in biochemistry, and operated through the National Institutes of Health in America. The size of such a college, as a peer group exchanging gifts, is limited to one hundred members or so. The main limiting factor is the number of papers that any single member is prepared to read in a working month. First authors of papers prefer writing them rather than reading those of others.

What characterized the members of the protein synthesis college, this peer group, enabling their success? Curiosity and intellect, driving, tormenting curiosity; the sort of consuming tantalization at today's problem that sees you waking with a head spinning with ideas but unsure of what day of the week it is. And of course, those absent-minded boffins from across the world were as bright as they come, one way or the other. They were quick, sharp, and voluble. They were slowly methodical, quietly using their pattern recognition skills in the delicate but fact laden approach of X-ray crystallography; pragmatically doing experiments with broths of bacteria. All manner of highest talents were needed to fight their way through the fog of what they knew they did not know. Scientists emerging victorious from that fog, however, was not how the recombinant DNA method to synthesize insulin was revealed. Very few of them were *intending* to manufacture anything, let alone synthetic insulin.

*

Understanding the chemistry of amino acids and proteins had developed so fast that the information in the papers of the Sanger team seemed an opportunity for synthetic organic chemists. Could this medically vital and commercially successful drug molecule be synthesized instead of the usual extraction from pancreases cut from carcasses of cattle and pigs? Several groups started work on this but the lead was held by Panayotis Katsoyannis, at the Medical Research Center of the Brookhaven National Laboratory, New York State. His group relied directly on the data of Sanger and team, and demonstrated that their chemically synthesized insulin was biologically active. Unfortunately, it was complex and expensive to make, having 200 different steps in the synthesis. The field remained open for a fresh approach.

By the late 1960s, the knowledge needed to make a connection between the understanding of natural synthesis of proteins and the ability to synthesize specific proteins as a laboratory manipulation came from an unlikely source. The geneticists Werner Arber and Daisy Dussoix, at the newly formed Biophysics Laboratory of the University of Geneva, studied *E.coli* bacteria, specially their ability to resist infection by bacteriophages. They were following a line of investigation started in the 1920s on a classification of pathogenic bacteria based on the infective ability of bacteriophages. The Swiss researchers found the bacteria were able to defend themselves from infection by phages. However, remaining uncertain of the mechanism, they named it from the perspective of the phages, as ‘host controlled modification of bacteriophage’.

This discovery, of what looked like a primitive immune ability of bacteria, was a challenge. Matthew Meselson and his new colleague, Robert Yuan, then both at Harvard University, responded with the advantage of outsiders. Immunity was traditionally the business of medical researchers who studied its highly evolved complexities in humans resisting infection by bacteria. To think of bacteria themselves having immune ability required a big mental somersault. Meselson and Yuan found an enzyme that gave the ability to resist, or restrict, infection of bacteria by phage. This type of enzyme was soon to become known as a restriction endonuclease. It chemically cuts twice through the phosphate backbone of DNA whilst leaving the nucleotide bases intact; a pair of molecular scissors. Other enzymes known as ligases, found by workers in the group of Arthur Kornberg, rejoin these cut chemical bonds. Soon other groups, enthralled by the delicate power of these techniques, corroborated the findings of Meselson and Yuan.

All these researchers realized that restriction endonucleases could cut out desired fragments, which could then be spliced with other useful fragments of DNA. This was possible only if the cut ends of the DNA had sequences of nucleotides that were exactly complementary and thus would naturally fit together according to the structural rules of the molecule. Thus fermented a question in the quickening imaginations of molecular biologists. A transformational question. Can we take the gene for an important protein from one host, say humans, and insert it in a microorganism in such a way that the new host manufactures the protein in a form that can be extracted for use? Researchers struggling up the rocky valley that was natural synthesis of proteins arrived at the pass and looked out over a vista fertile with promise.

Amongst the first to venture into this new land, it took Paul Berg to connect what strands of evidence, what existing methods, enzymes, reagents and protocols needed to be assembled to demonstrate the possibilities. Berg had arrived at this strategic position from undergraduate studies in biochemistry at

Pennsylvania State University that had been interrupted by service as a pilot in the US Navy. Doctoral studies at Western Reserve University stimulated his interest in the action of cellular enzymes, leading to a position at Washington University, in St. Louis, working on enzymic processing of amino acids in Arthur Kornberg's lab. Berg chose a virus as his main experimental workhorse. Simian virus 40 was originally found as a contaminant of cultures of cells that were used for making polio vaccine; the cells were originally from monkeys. Berg became such an enthusiast for this hardworking virus that by the time he moved to California he was displaying SV40 as the registration plate of his automobile. Enormous potential as an experimental tool lay in the ability of this virus to infect human cells without causing harm. Berg and colleagues dubbed it a cloning vector; in plain English a carrier of copies of DNA. They aimed at a practical goal: genetic therapy for people using the products of molecular biology.

David Jackson, Robert Symons and Paul Berg, all then at Department of Biochemistry, Stanford University Medical Center, were the first group to demonstrate what was soon to become known as recombinant DNA technology. As they stated in the introduction of their paper of 1972: 'Our goal is to develop a method by which new, functionally defined segments of genetic information can be introduced into mammalian cells. It is known that the DNA of the transforming virus SV40 can enter into a stable, heritable, and presumably covalent association with the genomes of various mammalian cells.'

They started with the circular DNA of SV40, cutting it to produce linear strands. Onto these they inserted genes from a specific bacteriophage and the cluster of genes of *E.coli* that code for the galactose enzyme. Then they closed the strands into circles again, forming the biochemical equivalent of an unusual but natural form of DNA known as a plasmid. In other words, they invented a way to manufacture plasmids containing desired genetic information. Could the genetic information now be placed in a system where it would produce desired proteins?

Back in 1952 Joshua Lederberg had introduced the term plasmid to unify diverse old descriptions of strange units of inheritance that were found occurring naturally in bacteria, but were not part of the normal mechanism of inheritance. He clarified that plasmids are self reproducing circles of DNA that transfer between bacteria of the same species. They are a form of genetic information in the cells that are not essential to the normal life of the bacteria, but do possess the information needed by bacteria to produce enzymes that can destroy antibiotics. Thus plasmids transfer antibiotic resistance to those strains of bacteria that contain them. A huge problem for the practice of medicine was about to become a huge opportunity for experimenters.

Who would grasp the chance first? Stanley N. Cohen at the Department of Medicine, Stanford University, Palo Alto, and Herbert W. Boyer at the Department of Microbiology, University of California, San Francisco, met for the first time at a conference in Honolulu in 1972; a crux now commemorated with a statue. Over a late night snack they discovered they needed each other to prove that plasmids constructed in the laboratory were sufficiently functional when inserted into *E.coli* bacteria to use the protein synthesis mechanisms of their new host. Cohen knew how to manipulate plasmids from *E.coli* that could be distinguished experimentally by their resistance to tetracycline antibiotic. Boyer knew how to use two restriction enzymes to cut the DNA molecule – he had recently discovered them in *E.coli*.

Back in California, in labs only forty miles apart at either end of the busy San Francisco peninsular, they got to work on combining the properties of this plasmid and the special enzymes. The enzymes cut DNA in a precisely crucial way: not so much scissors, more a chisel cutting a mortise joint so that the two strands of DNA are left as exposed and staggered ends that can be thought of as sticky. That is, the cut ends of the DNA will stick to other similar ends of DNA that have exactly complementary sequences of nucleotides with which to bond. Cohen and Boyer invented gene-splicing.

The resulting new piece of nucleic acid was a clone, in the sense that the original small piece of DNA could be replicated clonally by dividing in bacteria. They were able to make a new plasmid from the gene of *E.coli* that codes for tetracycline resistance, then insert this into a strain *E.coli* without the resistance. They then proved that the new host bacteria were resistant, demonstrating the possibility of transferring the capacity to manufacture a new protein into another organism. Cohen and Boyer invented recombinant biotechnology.

Boyer became famous enough by 1981 to appear on the front cover of *Time* magazine: the painting, although respectful, shows a combination of an elderly cherub and a youthful Albert Einstein, complete with mass of dark wavy hair and walrus moustache. Looks can be deceptive: at college he had played football as a linesman, then as an independent researcher in San Francisco he not only struggled for money for his obscure probings into the life of bacteria, which seemed to funders of no medical importance, but other workers in the same department banished him to decrepit labs so as not to contaminate their own experiments. His background was an upbringing in Pittsburgh, where typical work for his relatives and friends was in the mining and railroad businesses. Assigned to give a presentation about DNA in his small liberal arts college, where he had signed up to a pre-medical curriculum, he became fired up with inspiration and, as recounted to author Stephen Hall, he thought ‘To hell with medicine. Who needs all these sick people to take care of? I want to do something interesting.’

Stanley Cohen, another east coaster, educated at the medical school of the University of Pennsylvania, never fell in love with the money making potential of gene splicing but at the meeting in Honolulu he understood as well as Boyer the potential of being able to manipulate bacteria in this way. These tiny but tough organisms replicate fast, very fast, and will do so within industrial vats if supplied with a broth of simple nutrients. Each bacterial cell would be a miniature factory if the inserted genetic material carried the code for a useful protein.

Soon Cohen and Boyer demonstrated they could transfer pieces of DNA using gene splicing across the evolutionary distance between bacteria and that favourite laboratory animal, the African clawed toad. It was looking like time to push this technology up to the bizarre level of a manufactured symbiosis between human medicine and bacterial genetics. Francisco Bolivar, of the University of Mexico and at the time working with Boyer and team at the University of California, published a pair of papers first demonstrating that artificial plasmids could be constructed for cloning for various practical purposes. First isolate the genetic information for a particular protein, then place it within an artificial plasmid, infect host bacteria with the plasmid, and grow the bacteria and finally extract the desired protein from them. Cohen and Boyer patented these methods as soon as possible; they meant business. Who were they in a race with?

*

Most of the early action to exploit recombinant DNA was at the University of California, San Francisco. Names of prominent post-doctoral workers were, in Boyer's lab: Herb Heyneker, Francisco Bolivar, Raymond Rodriguez, Patricia Greene and Mary Betlach. In contrast was William Rutter, the head of biochemistry department and expert on function of the pancreas, the body's source of insulin. He had teamed up with Howard Goodman who knew how to splice genes together. They wanted to know more about how the gene for insulin worked. In the Rutter and Goodman team were Peter Seeburg, Axel Ullrich, and John Shine. In the beginning all these people, from two different departments in the same building at the University of California spontaneously formed a collective trying to answer intertwined questions on insulin and its gene. A tribute to the international nature of research – in addition to their American hosts this collective hailed from Netherland, Germany, Mexico and Australia.

The easy going interchanges did not last long. More than academic reputations were at stake. Propositions of synthesizing or cloning genes of medically important proteins such as insulin were clearly in the air, with the scent of large profits about them. None of them, however, realized the size of the biotechnology industry about to mushroom from their discoveries.

In 1976, at a meeting on insulin sponsored by the Eli Lilly company and held in Indianapolis, the hot topic was gene splicing as a route to making synthetic insulin. Lilly was founded in 1876 and insulin had become one of their major products (up to 85% of US sales, then worth $160 million annually to Lilly). However, they felt that supply from the pancreases of cattle and pigs was failing to meet demand from the 1.5 million people with diabetes America at that time. Lilly predicted crisis within 20 years, and had already considered chemical synthesis of insulin. As an alternative they set up a working group on the commercial scope for genetic engineering to synthesize insulin. They invited many molecular biologists from universities throughout America to the group.

The people from Lilly were disappointed, finding the completely novel methods were without sufficient practitioners. The company had long experience of collaboration with academics since the 1920s when they licensed, from the University of Toronto, the process to extract and purify natural insulin. Now they seemed hesitant. Already they had been involved with a group who had demonstrated it was possible to synthesize insulin by chemical methods using the coding information of Frederick Sanger. Lilly took no further interest in it once they realized how costly was the complexity of this method. Instead they adopted the approach of getting involved with the two groups in California and another group headed by Walter Gilbert over at Harvard. The Lilly company failed initially to drive into the recombinant DNA field for manufacture of pharmaceuticals, but later became one of the main manufacturers of biologically synthesized insulin. The balance between commercial confidence in their existing very effective natural product and confidence in this new technology, untested in the marketplace, was very fine.

By 1975 Herbert Boyer was more optimistic but still needed much positive thinking because his approaches to commercial companies and other molecular biologists as partners were getting nowhere. However, as optimistic as Boyer was Robert A. Swanson, a venture capitalist working for a firm in California, who had identified synthetic insulin by recombinant DNA as a possible route to high profits. He tried to interest the Cetus company in California, also individual molecular biologists such as Paul Berg, but got nowhere. Stephen Hall summed up the attitude of most researchers: 'No one had a good

word to say about the idea. Molecular biology at that time had a long and illustrious tradition of basic research – untainted, as the biologists themselves would put it, by commercial interests. Venture capitalists were held in about as high esteem as ambulance chasers.'

Robert Swanson was born in New York, and went on to study chemistry then business management at Massachusetts Institute of Technology. His career started with Citicorp Venture Capital then he moved to Kleiner & Perkins in San Francisco. An early photograph shows him side by side with Boyer. The younger Swanson is business suited in grey and white, whilst Boyer wears a pink open-neck shirt, leather waistcoat and blue jeans. Both look confidently pleased as they pose in front of a mass of biotech plumbing – owners of a company succeeding with astonishing success.

Swanson was 28 years old, with little capital of his own but impatient and very ambitious. He aimed for a full research, development and manufacturing company based on the emerging molecular biology. Correct selection of initial products was crucial: suitable products had to have an existing, large and profitable market; they had to be likely candidates for a simple and quick passage through the regulatory process. Furthermore, the product had to be easy to sell because they needed, as a start-up company, to avoid what the pharmaceutical business calls missionary marketing. Boyer was the man Swanson was looking for and soon they created the company Genentech. They both staked $500 on it.

Despite the lesser business savvy of the other groups then joining this hottest of areas, there was increasing competition along the usual academic tracks of who could actually demonstrate the best technique for constructing a gene for insulin and manipulating the bacteria. There were at least four groups on this in America, all aware it would be a race, a sharp sprint bursting from the starting-blocks.

Boyer made the implausible decision to construct the gene from synthetic DNA; a chemistry kit project. The alternative biological approach was to isolate messenger RNA for the gene then make a copy of it, resulting in a complementary DNA gene. The biological method was obvious and sure whilst the chemical method was more advanced, little tried and tested. Boyer was prescient, basing his decision on fear of the adverse publicity already surrounding recombinant DNA and genetically modified organisms. Recently this had created enormous problems for molecular biologists, specially on the east coast.

Gripping accounts have now been written of the epic struggles of Walter Gilbert and colleagues to work in accordance with new regulations, which obliged them to find a lab with microbiological safety precautions of the specification needed to contain the most dangerous microbes. They went to the British government's lab known informally as Porton Down; originally set up for research on warfare microbes, thus equipped with ultra-secure containments. But the complex logistics of the research project let slip a peculiar error; their efforts ran out into the sand. Meanwhile the Rutter and Goodman group became mired with problems over a misunderstanding of their certification by the National Institutes of Health to work with a complementary DNA plasmid. This trapped them in a controversy at a time that was already becoming difficult because of intra-departmental competition. In contrast synthetic DNA insulin genes were not technically of human origin, so Boyer squeezed his method past these paranoid barriers. At least there seemed to be no regulations specifically forbidding his method.

In 1976 Boyer contacted colleague Arthur Riggs at the City of Hope National Medical Center, at Duarte, on the edge of Los Angeles. Was he interested in working on synthetic insulin? Riggs replied that he was writing a grant for work with Keiichi Itakura on synthesizing the hormone somatostatin and their approach was to synthesize chemically the gene, and then to insert it into bacteria for mass expression. For them this project was a trail run with a small simple protein leading to synthesis of more complex proteins with larger commercial markets. Somatostatin regulates the production of human growth hormone and absence of it leads to gigantism, a rare problem. Natural somatostatin is produced in the hypothalamus and can only be obtained from cadavers. However, the potential funders found the proposal implausibly ambitious. Against the fears of Swanson about unnecessary complications and diversions, Boyer decided their new company should support this project instead. He felt that Genentech needed the highly specialized skills of Riggs and Itakura. Soon a powerful team focused on a highly specific goal: using somatostatin as a demonstration project, invent synthetic insulin and be the first to patent and manufacture it.

Roberto Crea, David Goeddel, Herbert Heyneker, Dennis Kleid and others were typical of the production workers. Mostly research fellows on the post-doctoral treadmill of meagerly paid short-term contracts, working crazy hours set as much by the physiological demands of their microbial cultures as their own ambitions to make a name for themselves and get a proper job. At that time DNA nucleotide bases were commercially available, extracted from salmon sperm, costing from $9.80 per pound weight for cytosine through to $2800 for guanine. They had to be joined to the sugar containing groups (nucleosides) to form the nucleotides. The process is a simple chemical one, but the reactions never proceed to 100% completion thus each step had to be separately purified to remove other reactants that would spoil the next stage. The nucleotides then had to be strung together as triplets coding for the relevant amino acid. For somatostatin this amounted to 42 nucleotides for its 14 amino acids. Finally, these had to be formed into a double helix structure. It was the patents of Itakura and then Riggs, as a pair, filed in 1979 and granted 1982 that defined the basic method for synthesizing both somatostatin and insulin. For insulin the patentees made clear their method used directly the information of Frederick Sanger and colleagues on the amino acid sequence of natural insulin.

* * *

There exists no single starting place or time for this story but the development of X-ray crystallography in the 1920s is a sufficient anchor. From there to Boyer and Cohen in California took about sixty years, of which less than twenty were occupied by the invention of recombinant DNA technique and its first commercial application. What was happening in the first forty years? The entire story, up to the point of researchers like Berg, Cohen and Boyer realizing the potential of the techniques to manipulate plasmids, was acted out by science researchers in universities and similar laboratories. They sought answers to fundamental questions of physics, chemistry and biology with nothing more than token consideration of any practical outcome. As Bruce Alberts and co-authors of the first key textbook of molecular biology explained: ‘The development of this technology was neither planned nor anticipated. Instead, steady advances in the ability of researchers to manipulate DNA molecules were made on many different fronts . . .’

The revolution in biology that was dawning in 1980 was illuminated for all to see when Genentech issued 1.1 million shares on October 14, at $35 each. After one hour they were selling for $89 and by the end of trading they had settled down to $71.5. A good day's work: $38.5 million, then unprecedented in the history of Wall Street.

Sources and notes

Biography and history

Jacob, F. 1988. *The statue within: an autobiography*. Basic Books Inc., New York. [*'As if science was too serious . . .'*: pg 258. The last two chapters of his autobiography reveal much about the special nature of the Institut Pasteur during the work of the protagonists there on the discovery of protein synthesis.]

Watson, J.D., 1968. *The double helix. A personal account of the discovery of the structure of DNA*. Weidenfeld & Nicholson, London. [A lively and infuriating autobiographical impression of early work on structure of DNA.]

Crick, F.H.C., 1988. *What mad pursuit: a personal view of scientific discovery*. Weidenfeld & Nicholson, London. [As the title says, a personal view; see Ridely 2006 for more analysis of the protein synthesis problem.]

Ridley, M., 2006. *Francis Crick: discoverer of the genetic code*. Atlas Books, Harper Press, London. [Many details on the story of the protein synthesis problem.]

Watson, J. D., 2000. *A passion for DNA*. Oxford University Press, Oxford.

Watson, J.D., 2003. *DNA: The secret of life*. William Heinemann, London. [This has a clear diagram of protein synthesis on pg. 77.]

Wilkins, M. H. F. 2003. *The third man of the double helix; the autobiography of Maurice Wilkins*. Oxford University Press, Oxford.

Luria, S.E., 1984. *A slot machine, a broken test tube*. Harper & Row, New York.

Judson, H.F. 1979. *The eighth day of creation: makers of the revolution in biology*. Jonathan Cape Ltd, London (first edition 1975, Simon & Schuster, New York.) [An extraordinary journalistic account of the people and discoveries of DNA and protein synthesis, based on contemporary observations and many interviews with the protagonists.]

Brenner, S., 2001. *My life in science*. BioMed Central, London.

Hall, S., 1987. *Invisible frontiers: the race to synthesize a human gene*. Oxford University Press, Oxford. [A gripping journalistic account, based on interviews with 96 participants, of the competitive commercialization of synthetic insulin using recombinant DNA technology.]

Ferry G., 1998. *Dorothy Hodgkin, a life*. Granta Books, London.

Ferry, G., 2008. *Max Perutz and the secret of life*. Pimlico, London.

Cairns, J., Stent, G.S. & Watson, J.D. (Editors), 1966. *Phage and the origins of molecular biology*. Cold Spring Harbor Laboratory Press, New York State. [Discussion articles by 35 contributors to the early days of the phage-DNA network.]

Portugal, F.H. & Cohen, J.S., 1977. *A Century of DNA. A history of the discovery of the structure and function of the genetic substance*. MIT Press, Cambridge, Massachusetts.

Fredrickson, D.S., 2001. *The recombinant DNA controversy: a memoir*. ASM Press, American Society for Microbiology, Washington. [A detailed memoir of this dismal saga from the person who was then director of the National Institutes of Health.]

What is a gene?

Morgan, T.H., 1933. The relationship of genetics to physiology and medicine. *Nobel Prize Lecture*. ['*There is no concensus . . .*': at halfway through this lecture.]

Delbrück, M.L.H., Timoféëff-Ressovsky N.W. & Zimmer, K.G., 1935. Über die Natur der Genmutation und der Genstrucktur. *Nachrichten aus der Biologie der Gesellschaft der Wissenschaften Göttingen,1*, 189-195.

Pauling, L. & Delbrück, M., 1940. The nature of the intermolecular forces operative in biological systems. *Science, 92*, 77-79.

Pauling, L. & Corey, R.B., 1953a. Structure of the nucleic acids, *Nature 171*, 346-346

Pauling, L. & Corey, R.B., 1953b. A proposed structure for the nucleic acids. *Proceedings of the National Academy of Sciences, 39*, 84-87.

Luria, S.E. & Delbruck, M., 1943. Mutations of bacteria from virus sensitivity to virus resistance. *Genetics, 28*, 491-511.

Platt, J.R., 1964. Strong inference. *Science, 146*, 347-353. [On doing biology like chemistry.]

Johnson, N.F., 2009. *Simply Complexity: a clear guide to complexity theory*. Oneworld Publications, Oxford. [For commentary on the need for reductionist approaches to discovery.]

von Neumann, J., 1970. The theory of self reproducing automata. In: A.W. Burk, *Essays on cellular automata* (Editor), University of Illinois Press, Urbana.

Levene P.A. & Bass L.W., 1931. *Nucleic Acids*, American Chemical Society Monograph Series 56, Chemical Catalog Company, New York . [Pgs 262-286, a descriptive account of the organic chemistry of nucleic acids.]

Todd, A.R., 1957. Synthesis in the study of nucleotides. *Nobel Prize Lecture*.

Pauling, L., Itano, H.A., Singer, S.J. & Wells, I.C., 1949. Sickle cell anemia, a molecular disease. *Science, 110*, 543-548.

Ingram, V.M., 1957. Gene mutations in human haemoglobin: the chemical difference between normal and sickle cell haemoglobin. *Nature, 180*, 326-328.

Volkin, E. & Astrachan, L., 1956. Phosphorus incorporation in *Escherichia coli* ribonucleic acid after infection with bacteriophage-T2. *Virology 2*, 149-161.

Translating genetic information into protein

[*'. . . to argue at top speed . . .'*: Jacob, 1988, pg 311.]

Gros, F., et al., 1961. Unstable ribonucleic acid revealed by pulse labelling of *Escherichia coli*. *Nature, 190*, 581-585.

Brenner, S., Jacob, F. & Meselson, M., 1961. An unstable intermediate carrying information from genes to ribosomes for protein synthesis. *Nature, 190*, 576-581.

[*'We found ourselves lying limply on a beach . . .'*: Jacob, 1988, pg 316.]

Crick, F.H.C., et al., 1961. General nature of the genetic code for proteins. *Nature, 192*, 1227-1232.

Crick, F.H.C., 1958. On protein synthesis. *Symposia of the Society for Experimental Biology, 12*, 138-263.

[*'What I could see in them was fog . . .'*: Judson, 1975, pg 427.]

Styles of research

Popper, K.R., 2002. *The logic of Scientific Discovery*. Routledge, Milton Park, England [*' . . . looking at the matter from the psychological angle . . .'*: pg 16. First published as *Logik der Forschung*, in 1935.]

Deutsch, D., 1997. *The Fabric of Reality*, Penguin Books, London.

Kendrew J.C., 1967. How molecular biology started. *Scientific American, 216*, 141-144.

Perutz, M. F., 1949. An X-ray study of horse methaemoglobin. II. *Proceedings of the Royal Society of London Series A: Mathematical and Physical Sciences 195*, 474-499. [*'The photographing, indexing . . .'*: pg. 475]

Polanyi, M., 1969. *Knowing and Being*. Routledge & Kegan Paul, London.

de Solla Price, D.J. & Beaver, D.B., 1966. Collaboration in an invisible college. *American Psychologist, 21*, 1101-1117.

Mullins, N.C., 1972. The development of a scientific speciality: the phage group and the origins of molecular biology. *Minerva, 10*, 51-82.

Experiments on natural protein synthesis

Monod, J., 1979. *Chance and Necessity: an essay on the natural philosophy of modern biology*. Fontana Books, London. [See chapter 6 for an example of Monod's thinking about DNA and proteins.]

Johnson, R.H., 1952. Passports and visas. *Science, 116*, 178-179.

Pardee, A.B., Jacob, F.& Monod, J., 1959. The genetic control and cytoplasmic expression of "inducibility" in the synthesis of β-galactosidase by *E.coli*. *Journal of Molecular Biology, 1*, 165-178.

Riley, M., et al., 1960. On the expression of a structural gene. *Journal of Molecular Biology, 2*, 216-225. [*'These experiments therefore . . .'*: pg. 225.]

Matthaei J.H. & Nirenberg M.W., 1961. Characteristics and stabilization of DNAase sensitive protein synthesis in *E.coli* extracts. *Proceedings of the National Academy of Sciences of the United States of America 47*, 1580-1588.

Nirenberg, M.W.& Matthaei, J.H., 1961. The dependence of cell-free protein synthesis in E.coli upon naturally occurring or synthetic polyribonucleotides. *Proceedings of the National Academy of Sciences of the United States of America, 47*, 1589-1602.

[*Caenorhabditis elegans*: Look this up on Wikipedia for a video of this worm wriggling, with its individual cells visible.]

Crick, F.H.C., 1957a. Molecular structure of deoxyribose nucleic acid and nucleoprotein and possible implications in protein synthesis. *Biochemical Society Symposia, 14*, 25.

Crick, F.H.C., 1957b. Codes without commas. *Proceedings of the National Academy of Sciences of the United States of America, 43*, 416-421.

Brenner, S., 1957. On the impossibility of all overlapping triplet codes in information transfer from nucleic acid to proteins. *Proceedings of the National Academy of Sciences of the United States of America, 43*, 687-694.

Sanger, F. & Tuppy, H., 1951. The amino acid sequence in the phenylalanyl chain of insulin. 2. The investigation of peptides in the phenylalanyl chain of insulin. *Biochemical Journal, 49*, 481-490. [The last of the quartet of papers in that year describing the amino acid sequence of insulin.]

Katsoyannis, P.G., 1966. Synthesis of insulin. *Science, 154*, 1509-1514.

Katsoyannis, P. & Tometsko, A., 1969. Process for joining the A and B chains of insulin. *United States Patent* 3420810.

Arber, W. & Dussoix, D., 1962. Host specificity of DNA produced by *Escherichia coli*. I. Host controlled modification of bacteriophage λ. *Journal of Molecular Biology, 5*, 18-36.

Arber, W., 1965. Host-controlled modification of bacteriophage. *Annual Review of Microbiology, 19*, 365-378.

Meselson, M. & Yuan, R., 1968. DNA restriction enzyme from *E.coli*. *Nature, 217*, 1110-1114.

Smith, H.O. & Wilcox, K.W., 1970. A restriction enzyme from *Hemophilus influenzae*. I. Purification and general properties. *Journal of Molecular Biology, 51*, 379-391.

Boyer, H.W., 1971. DNA restriction and modification mechanisms in bacteria. *Annual Review of Microbiology, 25*, 153-176.

Recombinant DNA method

Berg, P., et al., 1961. Enzymic synthesis of amino acyl derivatives of ribonucleic acid. *Journal of Biological Chemistry, 236*, 1726-1730.

Berg, P., 2004. Asilomar and recombinant DNA. *Nobel Prize Lecture*.

Jackson, D.A., Symons, R.H. & Berg, P., 1972. Biochemical method for inserting new genetic information into DNA of Simian Virus 40: circular SV40 DNA containing lambda phage genes and the galactose operon of *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America, 69*, 2904-2909. [*'Our goal is to develop . . .'*: pg. 2904.]

Lederberg. J., 1952. Cell genetics and hereditary symbiosis. *Physiological Reviews, 32*, 403-430.

Cohen, S.N., et al., 1973. Construction of biologically functional bacterial plasmids in vitro. *Proceedings of the National Academy of Sciences of the United States of America, 70*, 3240-3244.

[*'To hell with medicine . . .'*: Hall, 2002, pg. 60.]

Morrow, J. F., et al., 1974. Replication and transcription of eukaryotic DNA in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America, 71*, 1743-1747.

Bolivar, F., et al., 1977a. Construction and characterization of new cloning vehicles: I. ampicillin-resistant derivatives of the plasmid pMB9. *Gene, 2*, 75-93.

Bolivar, F., et al., 1977b. Construction and characterization of new cloning vehicles: II. a multipurpose cloning system. *Gene, 2*, 95-113.

Cohen, S.N. & Boyer, H.W., 1980. Process for producing biologically functional molecular chimeras. *United States Patent* 4237224.

Cohen, S.N. & Boyer, H.W., 1984. Biologically functional molecular chimeras. *United States Patent* 4468464.

[*'No one had a good word to say . . .'*: Hall, 2002, pg 66.]

Stern, S., 1995. Incentives and focus in university and industrial research: the case of synthetic insulin. In: N. Rosenberg, A.C. Gelijns & H. Dawkins *Sources of medical technology: universities and industry* (Editors), pgs 157-187, National Academy Press, Washington DC.

Crea, R., et al., 1978. Chemical synthesis of genes for human insulin. *Proceedings of the National Academy of Sciences of the United States of America, 75*, 5765-5769.

Goeddel, D.V., *et al*., 1979. Expression of chemically synthesized genes for human insulin. *Proceedings of the National Academy of Sciences of the United States of America, 76*, 106-110.

Itakura K., 1982. Recombinant DNA cloning vehicle. *United States Patent* 4356270.

Riggs, A.D., 1982. Method for microbial polypeptide expression. *United States Patent* 4366246.

Alberts, B., et al, 1994. *Molecular biology of the cell.* (3rd edition) Garland Publishing, New York and London. [*'The development of this technology . . .'*: pg. 298.]