

How Genes Manufacture Plants and Animals

JERRY BERGMAN

ABSTRACT

Molecular biology researchers have revealed a world far more complex than the macroscopic and microscopic worlds. Research in molecular biology has revealed a world of enormous order, information, design and intelligence that can be rationally understood only by design theory. No evidence exists for a naturalistic evolutionary mechanism which would adequately account for this order and complexity. In many ways the gap between minerals and one-cell life is greater than that between bacteria and humans.

HOW THE RECIPE FOR LIFE WORKS

Molecular biology has revealed that the physical foundation of life consists of information coded by four bases — adenine, thymine, guanine and cytosine. These four compounds are grouped in sets of three called codons in the DNA molecule(s). These sets determine the amino acid order needed to manufacture functional strings of amino acids, called polypeptides, which often assemble into proteins. The code set that produces one protein is called a gene; each gene including introns and control sequences is on the average 40,000 base pairs long. The codons serve to arrange the 20 common amino acids into the order needed to produce functional polypeptide chains which fold according to the specific amino acid positions on the chains.¹

These chains, called the **primary structural level**, are in turn assembled to form three dimensional units such as alpha helix or beta sheet structures, a complexity referred to as the **secondary structural level**. These spring-shaped coils and corrugated sheets then fold up into complex three-dimensional shapes referred to as the **tertiary level**. The last level, called the **quaternary level**, involves various polypeptide units associating together to form functional structural units such as haemoglobin and insulin. Proteins work because of their shape and charge pattern, which result from all four of the complexity levels listed above.

PROTEIN FOLDING

The core structure of the 20 different amino acids is

identical. For this reason they bind in a consistent fashion to form long chain-like structures. The 'R' groups or chemical side chains are unique to each one of the 20 amino acids. They stick out from the chain like little hairs and become locked into other side groups, and also into the backbone of the protein. This process, called **folding**, produces the required three-dimensional protein shapes and charge patterns necessary for the protein parts to function in the body. The folding of the amino acid chains into functional proteins takes place because of the placement of three types of bonds that form **between** the amino acid side chains. The bond types are disulphide linkages, hydrogen bonds and salt bridges.² These are all weak bonds compared to the bonds which hold the links of the chain together. The chain link bonds are all strongly covalent, and this difference is critical in how proteins are regulated. It also gives proteins great flexibility and mobility without which life could not exist. Molecular biologists who have tried to design new proteins have found it very difficult, because proteins are enormously complex. Even with modern computers and only 20 amino acids, years of work has so far achieved only limited results.³

Proper folding of the amino acid chain produces the specific three-dimensional shape needed to function. Preventing wrong folding is just as important as causing the correct folding, and this is no simple feat. Correct folding is aided by large complex molecules called chaperone proteins. Proper folding must occur because proteins function according to their shape, just as a key works in a lock because of its specific shape. As filing off a notch on a key would likely cause it to not function,

likewise, even small changes in the protein often cause it to malfunction.

All cells contain a comparatively large amount of water — the average human is about 64 per cent water — and we now know that a water environment is critical for proper protein function. Most proteins fold properly only in a water environment, a process called hydrophobic interaction and a characteristic called hydrophobicity, because their hydrophilic (water loving) and hydrophobic (water repelling) areas are of critical importance.

Ten amino acids have electrically charged side groups that are attracted to water, and the other ten lack an electrical charge. Consequently, those amino acids that lack a charge tend to cluster in the dry **inside** of the folded up protein molecule. The attraction of water-repelling amino acids to each other, and the hydrogen bonding networks between the backbone and side chain, produce alpha helices and beta sheets, the two basic secondary level amino acid shapes.

The other common environment existing in cells that proteins are made to work in is **fat**, such as that found in the cell membrane. Proteins made for this environment fold much differently. They must be designed to work with hydrophobic R groups **outward**, the opposite of water environment amino acid chains. Thus proteins assembled to work in a fat environment must be designed very differently from those designed to work in a water environment.⁴

These factors result in hundreds of thousands of different protein types, many unique to their owners. The plant and animal kingdoms contain 'literally billions' of protein species, all complex, intricately folded and perfectly fitted to the life-form each is part of.⁵ Interestingly, research on the catalytic properties of several enzymes put in various non-aqueous polar solvents has shown that this environment influences most proteases to run in **reverse** — that is, they catalyse polypeptide chain **formation** instead of cleavage. For example, chymotrypsin can be used to synthesise peptides, the reverse of its normal role in the cell.

Many argue that because life can be broken down into a simple three letter DNA code, this apparent simplicity negates the requirement for a creator.⁶ This apparent simplicity, though, must be viewed in relation to all of the other levels of biological complexity. Because a beautiful painting can be reduced to a large number of three different coloured dots, or an enchanting musical consists at its core of a mere eight basic different sound variations called notes, does not subtract from the beauty or complexity of the finished work. Likewise, the masterpieces of a Shakespeare or Milton can be 'reduced' to 26 basic symbols, but this does not negate the genius and insight existing in their literature.

The **Encyclopaedia Britannica** can be reduced to a series of binary code symbols on a magnetic tape — magnetised or not magnetised bits — a fact which likewise does not detract from the enormous amount of information,

profundity, and insight existing at the higher level. Research in genetics and molecular biology has revealed a huge amount of complexity existing in the human body and all life,⁷ although it is largely contained in the 3-letter codes.

PROTEINS: AN INTELLIGENT SOLUTION

Mechanical parts can be manufactured by the removal of unwanted material, as accomplished by a milling machine or lathe. They can also be **moulded**, as by the injection moulding process, or **stamped**, as in a tool and dye press, or manufactured by some combination of these techniques. None of these methods is feasible for achieving the precision needed for parts used in living cells. Only a chain 'growth' process, that self-folds under the proper conditions to produce the correct three-dimensional shape, is feasible for life. Further, the parts produced cannot be rigid as are the parts of a mechanical watch, but must be flexible within limits so that a cell can function **even if** it is physically distorted.

The explosion of knowledge about molecular biology has caused the level of understanding of human anatomy and physiology of only a few decades ago to seem enormously simplistic. Most 19th century biologists concluded that cells were simple small lumps of albuminous carbon not much more complex than clay.⁸ Scientists now realise that cell complexity surpasses that of any machine. It involves multi-thousands of proteins, all interacting with each other in complex ways.

The complexity of proteins is illustrated by allosteric enzymes, those designed to change shape reversibly in response to a chemical signal. Allosteric proteins are also called regulatory proteins, because changes in their shape are used to inactivate their enzyme activity. An example is a negative feedback loop, where the product of an enzyme is more likely, as it accumulates, to fill a regulatory slot on the enzyme, which changes the enzyme's shape and consequently inactivates the enzyme. When the concentration is lowered, the product falls off, activating the enzyme again and allowing the chain to operate to produce more of the product until enough is again produced.⁹

Control can be local, as in the example above, or at a higher level, such as at the machinery production level — the DNA which produces the enzymes in the first place. The first feedback system is rapid and sensitive; the second is slower but has more consequences. The second involves repressors which sit on genes and block their transcription to mRNA. An example is the repressor which falls off when tryptophan is low, allowing RNA polymerase to copy the DNA which makes the mRNA that codes for the enzymes which manufacture tryptophan from sugar. When enough tryptophan exists in the cell, the tryptophan fits on the repressor, reactivating it and allowing it to suppress mRNA production.¹⁰

Allosteric receptors are also built into the cell membrane. When the proper molecule binds with it, the

Base	Thymine (T)	Cytosine (C)	Adenine (A)	Guanine (G)	Uracil (U)
Nucleoside (base + sugar)	Thymidine	Cytidine	Adenosine	Guanosine	Uridine
Nucleotide (base, sugar + phosphate)	Thymidylic acid	Cytidylic acid	Adenylic acid	Guanylic acid	Uridylic acid

Table 1. The 'letters' of the universal code of life.

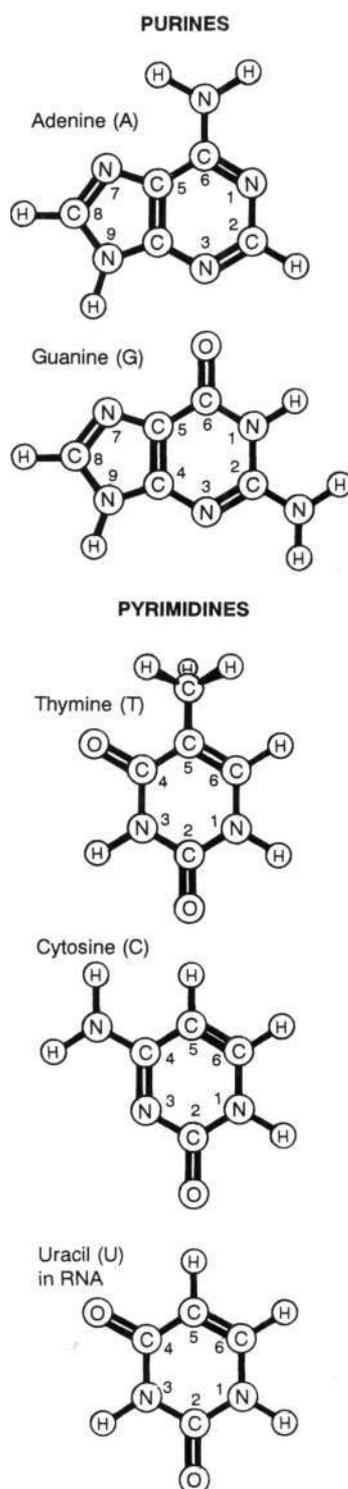
receptor changes shape, releasing an internal signal which triggers cellular changes. By this means information is communicated to cells, and they are thereby regulated.

Modern research has demonstrated that the simplest organisms, the prokaryotes (for example, bacteria), are enormously complex and require millions of base pairs for their blueprints. Even viruses, though their nucleic acid codes only for a few proteins, carry a huge amount of information. The amount of information contained in an *Escherichia coli* plasmid, a small circular DNA unit, is so large that it created storage and analysis problems until the advent of high-speed computers.

THE STRUCTURE OF DNA

DNA consists of a long chain of specifically-ordered nucleotide nitrogenous bases. The four bases include two **purines**, guanine (G) and adenine (A), and two **pyrimidines**, thymine (T) and cytosine (C). Two basic types of backbones exist to hold the code together — deoxyribose nucleic acid (DNA) and ribonucleic acid (RNA), a single-stranded copy of DNA. DNA was discovered by Oswald Avery in the 1940s and named **deoxy**, because the sugar in its backbone lacks an oxygen. RNA, which has a ribose sugar phosphate backbone and uses uracil (U) instead of thymine, carries the code stored on DNA in a molecule called messenger RNA (mRNA) to the protein manufacture sites on structures called ribosomes (see Table 1 and Figures 1 and 2). Ribose has a hydroxyl group (OH) which makes it **less** stable than deoxyribose. This contributes to the short half-life of RNA of only 1-5 minutes.

DNA is assembled by a clipping enzyme, which removes a pyrophosphate



from free nucleoside triphosphates present in the cytoplasm and then bonds the resulting nucleotides into a chain by dehydration synthesis.¹¹ The energy produced by the cleavage provides the energy necessary for linking new nucleotides to the DNA chain. The nucleotides are held together by phosphodiester bonds formed by the enzyme DNA ligase bonding the C-3 free hydroxyl group (OH) of the sugar ring to the C-5 phosphorylated group. DNA is acidic because it contains phosphate groups, causing it to be negatively charged, a fact used to separate the DNA strands according to size by electrophoresis. The glucoid (sugar) and phosphate linkages give the molecule directionality, namely from the 5' to the 3' end.¹²

Bases pair up in a complementary manner to form the double DNA strand. These are thymine and adenine, held together by two hydrogen bonds, and cytosine plus guanine, which are held together by three hydrogen bonds. Exceptions include the rare **tautomer** isomer held together by an unstable bond that can cause a mutation if an unrepaired mismatch results.¹³ A tautomerism is a reversible interconversion of structural isomers that usually involves the transfer of a proton.

The correct bonding pattern conforms to **Chargaff's rule**, named after its discoverer, Edwin Chargaff. The number of guanines equals the number of

Figure 1. The molecular structures of deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) are built using the nitrogenous bases adenine and guanine (purines), and thymine, cytosine and uracil (pyrimidines), which are the 'letters' of the genetic code, plus a sugar 'backbone'.

cytosines, and the number of adenines equals the number of thymines. If C equals about 22 per cent of the total number of bases, obviously G must equal about 22 per cent and the remaining bases equal 56 per cent. Thus T must equal 28 per cent and A likewise is 28 per cent. The two strands of DNA that are bonded together are **anti-parallel**, meaning they lie opposite each other so that a 5' to 3' strand is bonded to a 3' to 5' strand as follows, for example:

5'TAGCT3'
3' ATCGA 5'

The first 'letter', or codon, of the code was discovered by Marshall Nirenberg and Sohan Matthaei in 1961. They used a triplet of uridines to produce polyphenylalanine, a chain made up of the amino acid phenylalanine. Later other letters such as UUC were used to determine what they coded for, and all 64 of the triplet codes were identified by 1965.

The DNA structure is best described as a double helix, and the most common DNA type is the B form. If the B strand direction is facing up (5' to 3'), the sugar-phosphate framework structure would curve to the **right**, and thus it is labelled **right-handed**. The total outside diameter of the B form spiral is close to 18 Å (1 Å = 10⁻¹⁰m). Because each base pair is nearly flat, one complete revolution of a DNA strand pair (about 34 Å long) consists of 10 bases stacked one on top of the other. Two types of openings exist in the side profile of the helix, a **major** (wide deep) groove and a **minor** (narrow shallow) groove.

These grooves provide **accessibility** for the proteins and enzymes that constantly hover around the chromosome so as to mend, bend, copy, wind and repress genes. They also enable effective transcription so that a gene is properly expressed in the phenotype. The regulation proteins discussed below can bind either in the minor or major groove, depending on the protein's size, shape, charge pattern and other factors.

The 'A' form of DNA spontaneously forms in many circumstances, including when DNA is dehydrated or is hybridised with RNA, or even when the DNA is mixed with alcohol. The 'A' form, which is also right handed but slightly wider in diameter than the 'B' form, contains 12 base pairs per pitch.

CEN Tech. J., vol. 11, no. 2, 1997

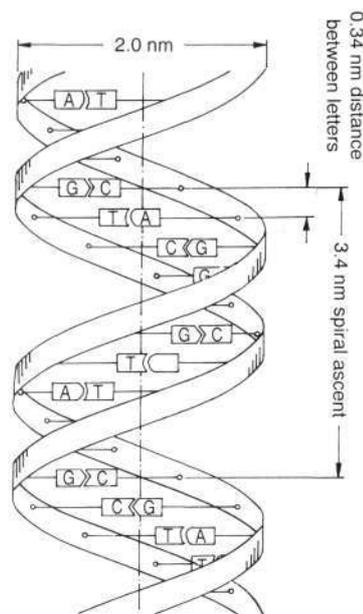
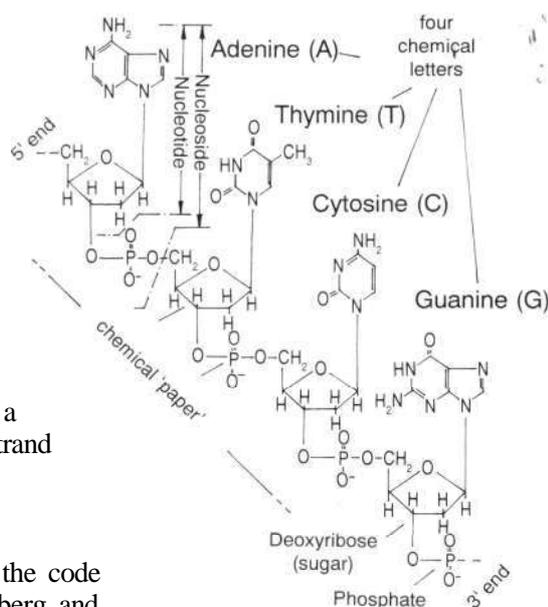


Figure 2. The way in which genetic information is stored.

Top: The chemical 'paper' is shown in the form of a long sugar-phosphate chain with the four chemical letters A, T, C and G.

Bottom: The actual structure and dimensions of the DNA molecule. (Adapted from: Gift, W., 1997. *In the Beginning Was Information*, CLV, Bielefeld, Germany, p. 92.)

The last known DNA form is called Z-DNA, and is **left handed** in contrast to both the A and B forms. Named Z-DNA because the side of the molecule forms a 'zig zag'-shaped Z pattern, it exists naturally in small stretches in some eukaryotic cells and can be used to produce antibodies.¹⁴ The Z form has only one type of groove instead of both major and minor types as formed in A and B types of DNA. It is believed to be part of a system that aids in regulating gene activity — Z-DNA is 'on', the other forms are 'off'.¹⁵

The B-DNA form can be caused to become Z form by (1) methylation of C or G bases, (2) supercoiling, or (3) protein binding.

The methylation process bonds methyl groups (CH₃) to the 5-carbon atoms of the pyrimidine rings on the opposite side of the 3 base pair hydrogen bonds. Highly methylated DNA is usually associated with genetically silent regions of the chromosome — meaning that these areas are turned off and therefore do not produce protein.¹⁶

The **outer** region of a DNA strand is called the backbone and consists of sugar and phosphate groups. It has a negative charge (is an anion) because of the phosphate and is **hydrophilic**. The **inner** region of the bases is **hydrophobic**, causing a stacking interaction because of the hydrophobic interaction between the adjacent base pairs (A, T, C, G), and because they also share electrons some

bonding results. Hydrophobicity occurs because the flat bases produce a stabilising force analogous to that achieved by squeezing two plates of glass together under water. The hydrogen bonding between the base pairs is also critical for producing the stability necessary for DNA molecules to exist in a cell environment.

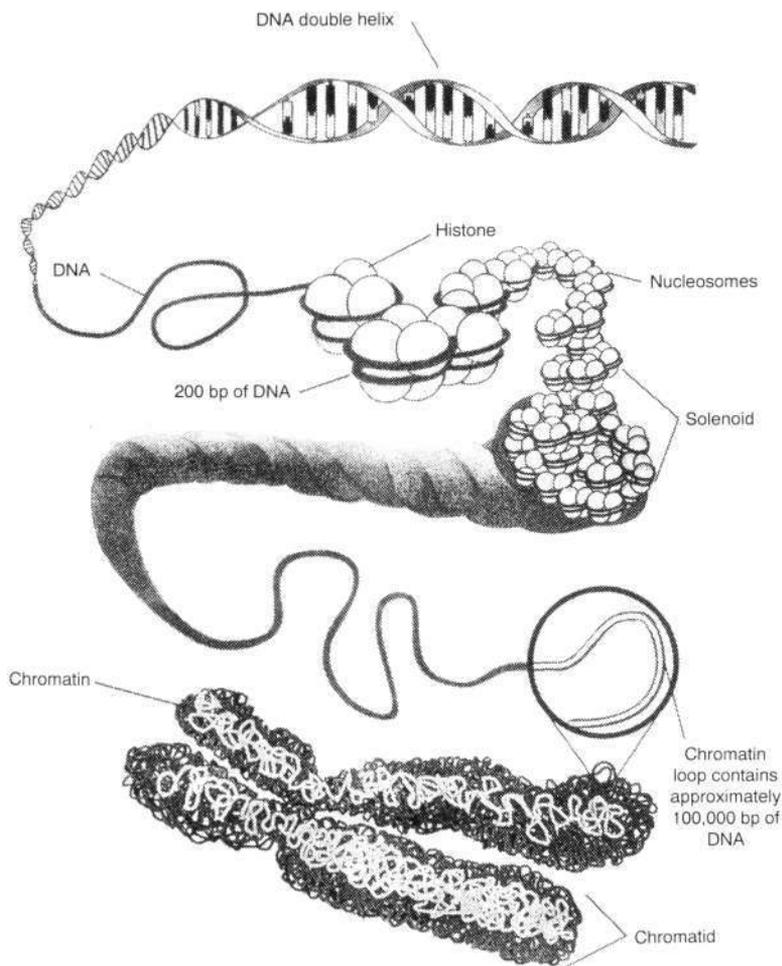


Figure 3. Patterns of DNA coiling. DNA is wound around histones to form nucleosomes. These are organised into solenoids, which in turn compose chromatin loops. (Adapted from: Jorde, L. B., Carey, J. C. and White, R. L., 1995. *Medical Genetics*, Mosby, St Louis, Missouri, Figure 2-15.)

THE DNA PACKING SYSTEM

Each human cell contains about six feet (1.8 m) of DNA. If a muscle cell were the size of a basketball, its DNA would be as thin as a fisherman's monofilament line and fully 125 miles (200 km) long. All this DNA must be packed so the regulator proteins that control making copies of DNA to manufacture protein have access to it. The discovery of the nucleosome, the basic structural packing and safe storage unit for DNA strands, was a major advance in understanding how each human cell's six feet (1.8 m) of DNA is prevented from becoming a tangled mess and is packed neatly in a chromosome structure. The packing reduces the length of DNA by a factor of 1 million (see Figure 3).

The nucleosomes are roughly spherical, about 100 Å (10 nm) in diameter, and consist of eight protein molecules called histones. About 150 pairs of bases are wrapped around the inner histones, and proteins other than histone

function as a further support for the folding system and to help stabilise the structure. Fully **half** of the eukaryote chromosome mass consists of histones. Interestingly, the histones are almost identical in all plants and animals, indicating that little deviance in their design is possible.

After DNA duplicates, the strands are first coiled around histones, forming a complex of histone plus non-histone proteins and DNA called **chromatin** (see Figure 3 again). The histones serve the critical role of protecting DNA from digestion by nucleases.¹⁷ It is only when DNA coils up completely by wrapping around various protein structures into chromosomes that DNA can be seen with a light microscope. The chromosome pairs are attached by a **centromere**. The last step in DNA duplication is for the centromere to separate as the cell's spindle fibres pry the chromosome pairs apart and pull one of each pair toward each end of the cell.

When fully packed, the DNA is supercoiled, a regular method of packing similar to repeatedly twisting the ends of a rubber band in opposite directions. If the chromosome is stained, distinct bands of colour appear that reveal differences. The darker areas around the nuclear membrane, called the **heterochromatic** regions, are inactive, tightly condensed genes. The lighter areas in the centre of the nucleus, called **euchromatic** regions, are potentially active areas which can produce mRNA for polypeptide synthesis. The lighter stained regions consist of loosely coiled genes — a condition necessary to allow the machinery that manufactures mRNA to move into the area. The areas actively producing mRNA are called chromosomal puffs because they look somewhat-like cotton puff-balls. Other active areas are called lampbrush chromosomes when they resemble the loops in lampbrushes. These traits appear and disappear as genes are activated and repressed. The loops and puffs respectively are most common during the active stages of cellular production.

Chromosomes often exist as **homologous** pairs (meaning the same shape). One chromosome, called a homologue, from each parent exists in most normal cells of sexually reproducing animals. Humans have 23 homologous chromosome pairs, or 46 chromosomes in total. The most common chromosome assembly is **diploid**, meaning that it consists of paired chromosomes. This assembly is typical of all familiar mammals, including humans, birds and many plants. Some life forms, such as fungi, male honey bees, and all asexual one-celled life, have only a **single set** of chromosomes, a condition called **haploid**. The third arrangement is **polyploid**, meaning the cell has more than two copies, a system found in plants such as potatoes (fortunately for genetic pioneer Mendel,

peas are diploid).

Not all DNA is stored in chromosomes: non-chromosomal DNA, called extranuclear DNA, is stored mostly in plasmids, mitochondria and chloroplasts. These genes sometimes follow slightly different coding rules from those of chromosomal DNA.

OTHER CLASSES OF DNA

We can also classify DNA according to its abundance in the cell. Over 500 nucleotide pairs (excluding introns) are required to specify the average protein, and as many as 300,000 genes exist in humans.¹⁸ The total number of base pairs in humans is estimated at 3 billion and, for comparison, the number in *E. coli* is about 2 million. Importantly, a relationship does not always exist between the morphological complexity of the organism and the number of DNA base pairs.¹⁹ The code does not just produce proteins, but is also responsible for possibly hundreds of other functions including control, repair and production of tRNA and rRNA. The basic DNA types include the following:

- (1) **High repetitive** — often found in satellite DNA, these sections are usually less than 100 base pairs long and can be as short as only two base pairs. These short sequences are then repeated up to a million times per haploid genome. It is called satellite DNA because it is easily separated by centrifugation. One stretch of 300 base pairs appears nearly one million times in humans. We do not yet understand the function of this DNA, an important research topic because it is such a large part of our total DNA.
- (2) **Repetitive** — an estimated 30 per cent of human DNA consists of sequences repeated at least 20 times.
- (3) **Middle repetitive** — is used for both rRNA and tRNA coding, and often 10 to 100 copies exist in humans.
- (4) **Single copy** — this form often codes for the genes that encode proteins.
- (5) **Unique or moderately repetitive DNA** makes up about half of the total haploid DNA and codes for most enzyme functions.
- (6) **Pseudogenes** have nucleotide sequences very similar to functional genes — most show a 75 to 80 per cent homology — but evidently do not code directly for protein as do functional genes.²⁰

Small amounts of the DNA in most humans may have come from viral infections, and viral DNA may be in a person's cells for as long as the person lives — AIDS is an example. These sections may be cut out by the DNA editing system, or the cell may shut down the virus genes by a mechanism called **repressive tolerance**.

To determine the amounts of each of the above DNA types, we first obtain a sample of DNA, then do a graph of the 'melt curve'. Heating causes DNA to separate into single strands, and each DNA type has a characteristic 'melt temperature' (T_m) which can be used for classification. Most

common DNA will also reanneal (that is, reform a double strand structure), and when it comes back together, sections will reanneal at a **different** rate, depending on their DNA traits. The renaturation rate differences can be used to identify different DNA types. Gradually increasing the temperature changes the light absorbance by as much as 20 to 30 per cent. This changes the hyperchromic effect, and this hyperchromicity difference can be used to differentiate the melted from the paired DNA form. The melted form absorbs more light because of its uncoupled chromophores (DNA bases) and tight winding, called the stacking interaction effect.

HOW DNA FUNCTIONS TO MAKE PROTEIN

Replication is the process of producing an exact copy of the DNA before cell division. The process is started by an initiator protein which finds the correct place to begin copying. It then guides the 'unzipper' protein called helicase to cause the double DNA strand to separate at the origin of replication, forming a fork area.²¹ The process of unwinding involves speeds estimated at approximately 8,000 rpm!

Since the DNA duplex kinks back on itself as it unwinds, a cut is needed to allow relief of the twisting pressure. The cut is made and repaired by an untwister enzyme called topoisomerase. Type I cuts only one strand; type II, such as gyrase, cuts both strands. Gyrase helps twist the DNA doublet into a supercoil, and other topoisomerases untwist supercoiled DNA so yet other enzymes can read or copy the stored instructions on DNA. Topoisomerase contains a 5' OH group, the latter linked to the tyrosine residue.²² Synthesis is in a 5' to 3' direction template, and the strand is read only in a 3' to 5' direction.

Enzymes that copy DNA work only on flat, untwisted sections. Researchers use knot theory to understand the physics of DNA twisting.²³ DNA polymerase synthesises two complete new strands of DNA, resulting in two sets of double strands, each of which forms a DNA pair called a duplex. Each new pair has one old and one new strand, a system called semi-conservative. When DNA is in the single strand form, it is fragile and must be strengthened, stabilised and prevented from becoming a tangled mess by single-stranded DNA binding proteins (SSB) which line up along the sugar/phosphate framework.²⁴

Many free nucleotides exist in the area of the cytoplasm where the new strand is being synthesised. The energy to drive the assembly process is derived from DNA polymerase splitting two of the three phosphate groups that are attached to each nucleotide. Next, DNA ligases (a stitcher repair protein) connect the nucleotides together into one continuous strand.²⁵

To translate the DNA code into amino acid chains, a messenger RNA or mRNA is first made from DNA by an enzyme called RNA polymerase. First a small section of DNA unzips, revealing the actual message called the sense strand and the template called the anti-sense strand. Then

the copy is made of the gene of interest only, producing a relatively short RNA segment. The knots and kinks provide crucial topological stop-and-go signals for the enzymes. Next, the mRNA is processed for its next role and then leaves the nucleus to do its work in the cytoplasm.

After mRNA is made, the DNA duplex is zipped back up. The process of mRNA production is called **transcription** and begins 20 to 30 bases away from a specific DNA sequence called the promoter (see Figure 4). These bases bind the RNA in one way only, thus transcription can occur in one direction only.

RNA is usually single stranded and much shorter than DNA, often from 50 to 1,000 nucleotides, but its length can be greater than 20,000 nucleotides.

RNA also uses hairpin structures (called such because they consist of RNA binding to itself to form loops of RNA in such a way that they resemble a hair pin) to help stabilise it and aid in effective translation. These structures consist of small folds in the RNA strand which protrude out from the strand. They slow down the translation process so the ribosomes can read the stop codes more effectively to terminate the translation process.²⁶

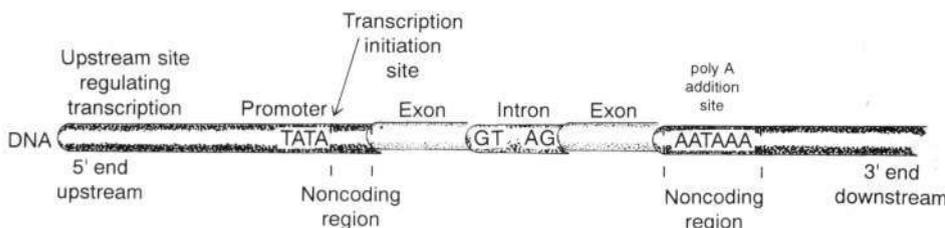


Figure 4. Details of gene structure showing promoter and upstream regulation (enhancer) sequences and a poly-A addition site. (Adapted from: Jorde, L. B., Carey, J. C. and White, R. L., 1995. *Medical Genetics*, Mosby St Louis, Missouri, Figure 2-13.)

DNA usually codes for **polypeptides**, small proteins which are assembled into

- (1) structural proteins like collagen,
 - (2) enzymes,
 - (3) hormones,
 - (4) molecular carriers such as haemoglobin,
 - (5) regulator proteins discussed above, and
 - (6) antibodies that are part of the body's defence system.
- The genetic code is read in blocks of three bases that are non-overlapping and called a reading frame. As an example, the DNA code TTT (the RNA code is UUU) codes **only** for phenylalanine.

The triplicate code used is 'degenerate', meaning **multiple** codons can often code for one amino acid. Leucine, for example, is coded for by TTA, TTG, CTT, CTC, CTA, and CTG. The third base of the code does not matter in several cases, because only the first two of the three bases determines the amino acid, a condition called the wobble effect. This reduces enormously the damaging effect of mutations. Accidental changes in the code are called mutations, and according to evolution are the creator of all life-forms. If a TTA mutates to TTG, no changes will result

in the protein, thus this mutation is called neutral.

Usually only one code, methionine, is used for a 'start' signal to begin the reading process.²⁷ It is almost always the first amino acid in the code (for DNA, the methionine code is ATG, and for RNA it is AUG). Three stop or termination codons exist, namely UGA, the most common, and also UAA and UAG. Translation stops when these codes are reached, because no tRNA exists with an anti-codon to bond to these codes. Also a protein called releasing factor competes for a mRNA base pair, but does not bind if a tRNA exists for a triplicate. The releasing factor results in the hydrolysis of the ester bond between the tRNA and the polypeptide chain, releasing the completed polypeptide.²⁸

An **open** reading frame is a series of codes, uninterrupted by a stop codon, that is properly read from the start codon to the stop codon. The series of triplets will then code for the proper amino acid sequence. An example is: AUG (the start codon is first) then UUU-AUC-GGG-UAA-GCG codes ending with UGA, a stop codon. An insertion mutation adds a base, and a loss of a base is called a deletion mutation. Both types shift the reading frame, often producing an amino acid chain that does not function

or is harmful, and therefore is cut up by the cellular proteolytic system so the amino acids can be recycled. If no other gene produces the original protein and it is essential for life, the mutation is lethal.

The open reading frame can be distinguished by computer analysis from the **other two** wrong reading frames if each set is compared by reading in sets of three, starting from the first base, then the second, and last, the third base in the set. The functional gene area can usually be identified by a **lack** of stop codons: stop codons will occur randomly in about 3 out of 64 codons (4.7 per cent), but significantly fewer stop codons will result if the **correct** reading frame is identified. This procedure reveals the information content of the genes and suggests that all of the eukaryote genome has a function. Sometimes there can be two reading frames covering the same sequence.²⁹ This complexity is not what evolution would predict.

Control of mRNA production is of crucial importance. Like a machine which produces parts, the correct parts must be produced in the needed amounts at the correct time. This is controlled by regulatory proteins. Those proteins serve either to start or stop the process of making proteins. They must travel into the nucleus when needed from where they are made on ribosomes, and must find a specific gene, and then bind to it in order to block it from making mRNA. By this means DNA is prevented from making proteins. Other regulatory proteins bind to a gene to signal it to start producing mRNA, as occurs in the various steps of cell division.³⁰

PROTEIN SYNTHESIS

Protein synthesis is controlled by a structure called a ribosome which consists of about 50 proteins and special RNA called ribosomal RNA. The three types of RNA involved are ribosomal RNA (rRNA), messenger RNA, (mRNA) and transfer RNA (tRNA). The short half-life of mRNA of only about 1-5 minutes helps control protein production.³¹ Because of this short half-life, about half of all RNA synthesised is mRNA, yet it represents only 3 per cent of cellular RNA.

Ribosomes are composed of a larger and smaller section. The whole structure functions as a frame to control the conversion of the base code into a set of amino acids, a process called **translation**. The mRNA attaches itself to the smaller ribosome subunit, then the first tRNA is matched with the code. Next, the larger ribosome subunit appropriately connects to the smaller, and the second tRNA bonds with its matching mRNA code. As the mRNA slides through the ribosome, the tRNA lines up amino acids according to the codon and anti-codon match (see Figure 5). At least one tRNA molecule exists for each of the 20 amino acids, and as noted above, because some can bind to several codons, 64 are not required. The tRNA is a single strand of RNA which folds back on itself so that its bases can hydrogen bond to form the shape required to do its job in the cell.³² Like a protein, the tRNA must fold up to produce the required conformation, and is therefore sometimes humorously said to be an RNA that wanted to be a protein.

The tRNA structure also uses several slightly modified bases such as inosine. The bases are modified after the tRNA strand is produced by special enzymes, including methylases, deaminases, thiolases, pseudouridylyating enzymes and transglycosylases.³³ The correct amino acid bonds to its mate tRNA with the aid of enzymes called tRNA synthetases and energy from ATP.³⁴ At least 20 different activating enzyme tRNA synthetases exist. They function by first an ATP docking in a slot on the enzyme, then an amino acid docks nearby and is 'charged' by the ATP losing two of its phosphates. The energised amino acid can now bond to the correct tRNA, which soon docks so that the proper end of tRNA can bind.³⁵ The tRNA is said to be charged when it has an amino acid attached to its binding site, and uncharged when it loses the amino acid.

To form the protein chain, a peptide bond forms between each added amino acid and the existing chain by dehydration synthesis. Concurrently, the bond between the tRNA and its amino acid is then broken by hydrolysis. The empty tRNA floats away to be recharged, so it can carry yet another amino acid to be bonded to the growing chain. This process occurs until a stop codon is reached and the two parts of the ribosome separate, dropping off the messenger rRNA. The mRNA is usually read by more than one ribosome (a polysome) simultaneously for efficiency. Eventually, the mRNA is broken down by a ribonuclease and its parts,

mostly nucleotides, are recycled into new RNA. The cell uses this as one way of achieving tight control of protein synthesis.

THE IMPLICATIONS FOR CREATION

The data gained from genetic research has been extensively analysed using computer programmes, revealing an enormous amount of purposeful information. Considering the fact that an estimated 10 million species

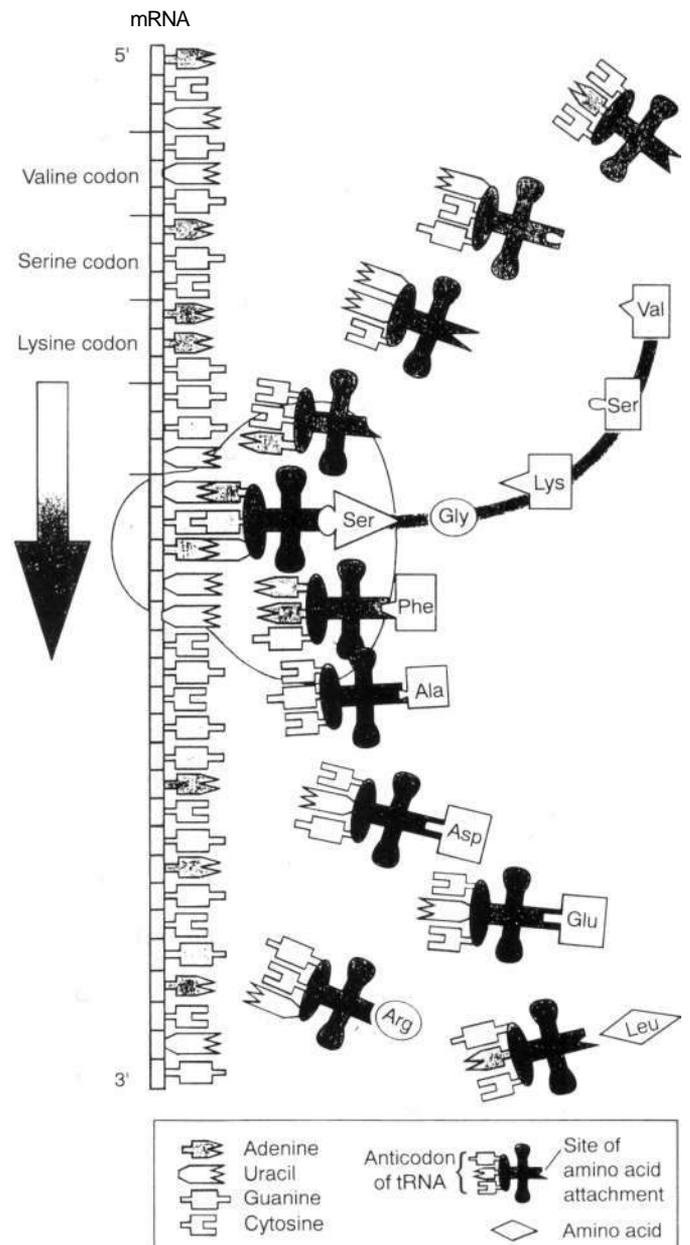


Figure 5. Translation of mRNA to amino acids. The ribosome moves along the mRNA strand in the 5' to 3' direction, assembling a growing polypeptide chain. In this example the mRNA sequence GUG-AGC-AAG-GGU-UCA has assembled five amino acids Val, Ser, Lys, Gly and Ser into a polypeptide. (Adapted from: Jorde, L. B., Carey, J. C. and White, R. L., 1995. *Medical Genetics*, Mosby, St Louis, Missouri, Figure 2-10.)

of life exist, a presently incalculable amount of genetic information exists in the plant and animal world. If this variety and that found in extinct animals were all mapped, it is likely that the existing world computer storage capacity could not contain all of the data. A **single** human genome contains more than four times as much information as a complete set of the **Encyclopaedia Britannica**. Multiplying this by both the estimated 10 million species and the information found within each individual species illustrates the enormous amount of information that exists in genes.³⁶

In researching the genetic structure of many animals, scientists were struck by the **incredible similarity** between all life at the molecular level. This phenomenon, which is termed 'the unity of all life', has resulted from the discovery that all living creatures use DNA and RNA to build life from the same four nucleotides, with hardly any exceptions. And all life uses these nucleotides to produce proteins by very similar methods. The same 20 amino acids, the same basic translation apparatus, including ribosome tRNA, mRNA, and hundreds of virtually identical enzymes, are used by all life. Ribosomes from bacteria will translate human mRNA into human protein and vice versa.³⁷

Even many of the proteins in the body, including enzyme proteins that serve as communicators and transporters, are *'quite similar in most creatures throughout the living world'*.³⁸ The fact that bacterial ribosomes are so similar to those in humans argues that the alleged 3.5 billion years of evolution between the two has produced, at best, minor inconsequential changes. It also argues for **stasis** at the molecular level, except for largely inconsequential changes called mutations and genetic drift. In fact, this is one of the most elegant arguments for creationism, and one of the clearest disproofs of the possibility of macroevolution.

Genetics does not argue for evolution as many evolutionists assume. It argues that certain complex systems are necessary for life, and all of these multi-thousands of precision parts must simultaneously exist for life to exist. In order for a television to work, it must contain certain parts. Without these basic parts a TV, from the low price black and white models to expensive colour with stereo sound, will not work.

Likewise, life will not work without certain parts. And all life from bacteria to humans must have all of the basic parts to function. An unbridgeable chasm exists between non-life and the simplest form of life, the bacteria. And the gap existing between non-life and bacteria is in many ways similar to the gap between a pile of sand, carbon, water, and a few minerals, and a basic black and white television set. Likewise, the gap between bacteria and humans is in many ways similar to the gap between an inexpensive black and white television set and a deluxe colour television set with VCR and stereophonic sound.

SUMMARY

Biochemistry has revealed not the simplicity once expected to exist at the chemical foundation of life, but an enormous level of complexity of which this review only skims the surface. The design and order that prevails at the higher level is found to a far greater degree as we move to the molecular level, revealing a degree of complexity that increases enormously as we move from the gross anatomy to the microanatomy, and lastly, to the molecular level. Further, the structure found at the molecular level vividly reveals the inadequacy of natural selection and mutations in achieving a role greater than fine tuning, and in reducing the level of DNA corruption that occurs with time. Natural selection therefore has a conserving role, not a creative one. Humans have now accumulated over 4,000 known mutations, and the number would be much greater were it not for natural selection.³⁹

The code of life consists of only four bases, and this code is only one of many levels of complexity, **all** of which are necessary and must be functionally integrated before life can exist. Humans are far more than a sequence of 3 billion letters written in DNA, and even these instructions *'are far more complicated than was ever thought'*, and indeed, the complexity and ingenuity of the gene design is being revealed daily.⁴⁰

A good illustration of the increasing realisation of complexity is the dramatic rise in the estimate of the number of genes in the human genome that has occurred in the past two decades. In 1970, Klotz noted that Spuhler *'on the basis of careful mathematical computation'* believed the number of human genes to be *'between 20,000 and 42,000'*.⁴¹ Other contemporary estimates, Klotz noted, were similar⁴² — obviously quite a contrast to those given today, which range from 200,000 to over 300,000 genes. Molecular biology, which is only a few decades old, is only just beginning to reveal the complexities of life.

A concern is the assumption that the basis of life is DNA, and that this code **is life**. Actually, the genes that code the proteins are only a **small part** of the total DNA code in multicellular organisms, possibly less than five per cent. However, DNA by itself is impotent and requires an extraordinarily complex structure to express itself. Likewise, a computer is not just a series of on and off switches, but a complex machine which is designed to use the stored information — and the entire computer is necessary for this stored information to be of any use. We have discussed here only a few of the basic parts necessary to translate the code into life.

A most appropriate analogy of the genome is the relationship between computer hardware and software. The software directs the operations of the hardware, but in some ways the hardware is the more complex part. Both parts are necessary and must be integrated to an enormously high degree for the machine to work. They originally were developed as a unit. The fine-tuned interrelationship is

critical. With the computer hardware-software analogy, the genome is the software and the associated cellular chemistry is the hardware. Both are necessary and have to be present from the start. This is an insurmountable problem for evolutionary theories of origins.

REFERENCES

1. Levine, J. and Suzuki, D., 1993. **The Secret of Life: Redesigning the Living World**, WGBH, Boston.
2. Feigl, D. M., Hill, J. W. and Boschmann, E., 1991. **Foundations of Life: An Introduction to General, Organic, and Biological Chemistry**, Third edition, Macmillan Publishing Company, New York.
3. Langreth, R., 1995. Take one amino acid. **New Scientist**, **147(1993)**: 32-35.
4. Hoagland, M. and Dodson, B., 1995. **The Way Life Works**, Random House, New York.
5. Langreth, Ref. 3, p. 32.
6. Denton, M., 1985. **Evolution, a Theory in Crisis**, Adler and Adler, Bethesda, Maryland.
Denton reviews these views.
7. Yockey, H. R., 1992. **Information Theory and Molecular Biology**, Cambridge University Press, Cambridge.
8. Farley, J., 1979. **The Spontaneous Generation Controversy from Descartes to Oparin**, The Johns Hopkins University Press, Baltimore, p. 73.
9. Hoagland and Dodson, Ref. 4.
10. Hoagland and Dodson, Ref. 4, p. 136.
11. Bhagavan, N. V., 1992. **Medical Biochemistry**, Jones and Bartlett Publishing, Boston.
12. Freifelder, D., 1982. **Physical Biochemistry: Applications to Biochemistry and Molecular Biology**, Second edition, W. H. Freeman and Company, New York.
13. Lewis, R., 1994. **Human Genetics**, William C. Brown Publishers, Dubuque, Iowa.
14. Mader, S., 1993. **Biology**, William C. Brown Publishers, Dubuque, Iowa.
15. Mader, Ref. 14.
16. Watson, J. D., Gilman, M., Witkowski, J. and Zoller, M., 1992. **Recombinant DNA**, Second edition, W.H. Freeman and Company, New York.
17. Guttenplan, J. and Oratz, M., 1994. **Biochemistry**, Sulzburger and Grahman, Publishers, New York.
18. Watson *et al.*, Ref. 16, p. 149.
19. Guttenplan and Oratz, Ref. 17, p. 196.
20. Adams, R. L., Knowler, J. T. and Leader, D. P., 1992. **The Biochemistry of the Nucleic Acids**, Chapman and Hall, New York.
21. Hoagland and Dodson, Ref. 4.
22. Kendrew, Sir J. and Lawrence, E. (eds), 1994. **The Encyclopedia of Molecular Biology**, Blackwell Science Ltd, Osney Mead, Oxford.
23. Skerrett, P. J., 1994. The ties that bind. **Popular Science**, May, pp. 114-120 (p. 119).
24. Starr, C. and Taggart, R., 1995. **Biology: The Unity and Diversity of Life**, Wadsworth Publishing Company, Belmont, California.
25. Starr and Taggart, Ref. 24.
26. Zubay, G. L., Parson, W. W. and Vance, D. E., 1995. **Principles of Biochemistry**, William C. Brown Publishers, Dubuque, Iowa, p. 710.
27. Osawa, S., 1995. **Evolution of the Genetic Code**, Oxford University Press, New York.
28. Singleton, P. and Sainsbury, D., 1994. **Dictionary of Microbiology and Molecular Biology**, Second edition, John Wiley and Sons, New York, p. 712.
29. Aw, S.-E., 1996. The origin of life: a critique of current scientific models. **CEN Tech. J.**, **10(3)**: 300-314.
30. Hoagland and Dodson, Ref. 4, p. 164.
31. Zubay *et al.*, Ref. 26.
32. Mader, Ref. 14, p. 256.
33. Zubay *et al.*, Ref. 26.
34. Mader, Ref. 14.
35. Hoagland and Dodson, Ref. 4, pp. 114-115.
36. Micklos, D. and Freyer, G. A., 1990. **DNA Science**, Carolina Biological, Burlington, North Carolina.
37. Hoagland and Dodson, Ref. 4, p. 122.
38. Hoagland and Dodson, Ref. 4, p. 122.
39. Bhagavan, Ref. 11.
40. Jones, S., 1993. **The Language of Genes**, Anchor Books, New York.
41. Klotz, J. W., 1970. **Genes, Genesis and Evolution**, Second revised edition, Concordia Publishing House, St Louis, Missouri, p. 248.
42. Klotz, Ref. 41, p. 248.

Jerry Bergman has seven degrees, including in biology, psychology, and evaluation and research, from Wayne State University (Detroit), Bowling Green State University and other colleges. He has taught at Bowling Green State University (Ohio) and at the University of Toledo. He is now a professor of science at Northwest College, Archbold (Ohio) and is working on a third Ph.D., this one in molecular biology.