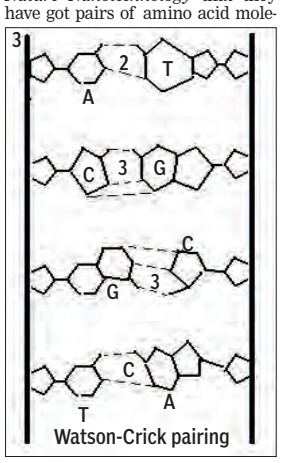
DNA architecture reaches out

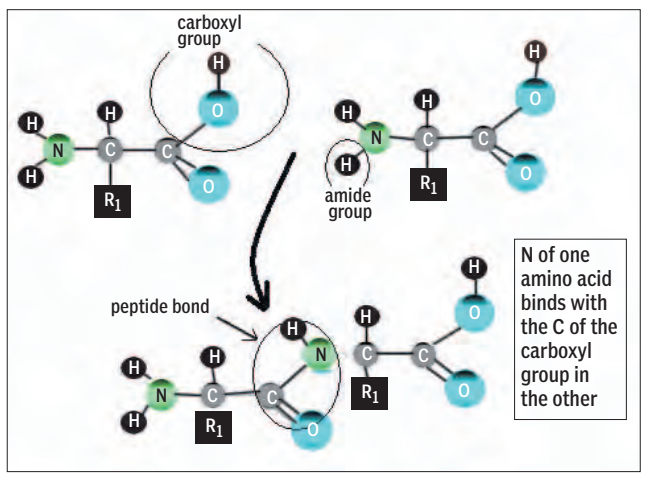
THE ASSEMBLAGE ITSELF INSPIRES THE CREATION OF OTHER STRUCTURES THAT SELF-ASSEMBLE, WRITES
S ANANTHANARAYANAN

The function of the DNA molecule is to help form a specific sequence of amino acids, each one picked out of a total 20 amino acids. Each such sequence is a separate protein and the DNA, in a chain that can be thousands of units long, spells out millions of proteins and as many characteristics of living cells. But along with this function, one strand, or half a DNA molecule, after cell division, has the ability to assemble its complementary strand alongside itself to create a complete molecule.

While there has been progress in using laboratory-made DNA sequences to create protein-like molecules "on order" — the so called "DNA Origami" — it is tempting to extend the self-replicating property of DNA to other structures. This could get the structures to attach to specific complementary molecules and form sheets or three-dimensional arrays. One possibility is with short sequences of amino acids, which are called *peptides*, and which connect to each other to form proteins. Or Berger, Lihi Adler-Abramovich, Michal Levy-Sakin, Assaf Grunwald, Yael Liebes-Peer, Mor Bachar, Ludmila Buzhansky, Estelle Mossou, V Trevor Forsyth, Tal Schwartz, Yuval Ebenstein, Felix Frolow, Linda JW Shimon, Fernando Patolsky and Ehud Gazit of Tel-Aviv University, the Institut Laue — Langevin, in Grenoble, Keele University in the UK and the Weizmann Institute of Science, Israel, report in the journal *Nature Nanotechnology* that they have got pairs of amino acid molecules to self-assemble into stable structures, which also showed interesting optical properties.



Watson-Crick pairing



DNA — Deoxyribo Nucleic Acid — is a long chain molecule built around a backbone to which are attached a series of chemical groups called "bases". There are only four kinds of bases, and these are *adenine* (A), *guanine* (G), *thymine* (T) and *cytosine* (C). Every group of three consecutive bases forms a "triad" and each triad is the code for a particular amino acid. A DNA strand of nine bases would thus have three triads and code for three amino acids, connected in that order, although real DNA consists of hundreds of thousands of bases.

Now, as each base can be of four kinds, a sequence of three bases can be formed in 4x4x4=64 ways. But many of these ways code for the same, important amino acids, so that an error in one of the bases would leave the meaning of the triad unchanged, and only 20 amino acids, plus a signal for the start and the end of a sequence, are actually coded for.

As we now have the means to synthesise the components of DNA and also ways to cut and splice segments of existing DNA, we have the ability to "create" DNA sequences and this ability can be applied to generate large quantities of tailor-made organic chemicals, with the help of the DNA template.

The other feature of the DNA molecule is that it consists of a pair of strands. While one strand has bases, A, G, T or C, attached along its length, the bases that appear on the complementary strand depend exactly on what bases there are in the first strand. Thus, where there is A, it would pair with T, and vice-versa, and where there is G, it would pair with C, and vice-versa. At the time of cell division, the DNA molecule splits into two, both portions being complements of each other and in the new cells the lone strands are able to assemble the second strand by controlling the order of A, T, G, C according to the T, A, C, G in the first strand.

The way this pairing happens is that A and T each have a pair of hydrogen bonds with which they can connect, while G and C have three hydrogen bonds each and G and C hence form a couple. This architecture makes sure that a dangling A, for instance, would only link to a T


Ehud Gazit

and if the base next to the A were a G, then the base next to the T would have to be a C, and so on.

The other kind of specific linking is of the amino acids that the triads code for. *Amino acids* are so called because they consist of an *amine*, or nitrogen containing portion and an *acid portion*. The amine side ends with NH_2, while the acid portion ends with *COOH*. In the *peptide bond*, the nitrogen of one amino acid binds with the carbon in the COOH portion of the second amino acid. The amine portion of the second amino acid can then bind with the COOH portion of a third amino acid, and so on.

In this way, the DNA, which replicates with the help of just four building blocks and the A-T and G-C pairing, called the *Watson-Crick base pairing*, generates a tool kit of just 20 amino acids that can couple among themselves with versatility and stability. And these 24 entities bring about the stupendous diversity of living things.

Self-assembly

While structures that use peptides, or the combination of amino acids, are stable and can be versatile and flexible, they do not have the selectivity of the base pairing mechanism of DNA. There was, hence, an effort to combine the strengths of the two architectures, and this led to the idea of the *PNA* — which is not a Deoxyribo Nucleic Acid but a Peptide Nucleic Acid. The PNA, the name is a misnomer as it is not an acid, is a chain of amino acids, like glycine, the simplest, linked by peptide bonds so that it forms a peptide backbone and has nucleobases (viz, A, T, G, C) as side chains. Such a structure could then show both the stability of the peptide as well as the possibility of the Watson-Crick base pairing. One of the simpler peptides, called diphenylanaline, consisting of only two amino acids, is known to self-assemble into tubular nanostructures that are stable, in fact they show a metal-like rigidity, can withstand thermal and chemical stress and even display optical and piezoelectric properties. To better understand the mechanism of how simple peptides may combine, the Tel-Aviv based group of scientists examined how all 16 of the simplest, two-unit PNAs could self-assemble in solution.

The sixteen di-PNAs are AA, AC, AG, AT, CA, CC, CG, CT, GA, GC, GG, GT, TA, TC, TG and TT. The group synthesised these di-PNAS and assayed for favourable conditions for self-organisation in different solvents. Assembly was found to arise in three of the 16 di-PNAs — *CG*, *GC* and *GG* — all three containing guanine. In these cases, well-organised architectures, including long rods (tens of micrometres long) for CG and GC and spheroids with a diameter of two-three micrometres for GG were seen. These observations have pointed to the conditions, like the number of hydrogen bonds, which are necessary for structures of di-PNAS to be stable. And the fact of guanine, which is a key component in many natural nucleic acid structures, being present has led the team to more conclusions.

A study of the crystals of the structures, using X-Ray scattering, revealed that cytosine and guanine, with Watson-Crick type hydrogen bonding, led to stacking interactions between molecules. It is, hence, genuinely a case of DNA style selection brought to play to create stable structures built from peptides. It was also discovered that the structures displayed fluorescence, or the absorption and emission of light, with features like wide span of wavelengths, that make the structures suitable for use as organic light emitting material in optoelectronics, optical communication, display, etc.

THE WRITER CAN BE CONTACTED AT
simplescience@gmail.com

INTEGRAL COMPONENTS

TAPAN KUMAR MAITRA EXPLAINS THE FLOW OF GENETIC INFORMATION

DNA is the genetic material of cells and organisms and we have come to understand its structure, chemistry and replication, as well as the way it is packaged into chromosomes and parcelled out to daughter cells during mitotic and meiotic cell divisions.

To explore how DNA is expressed — that is, how the coded information it contains is used to guide the production of RNA and protein molecules — the nature of the genetic code and the way in which information stored in DNA guides the synthesis of RNA molecules is a process call *transcription*. How RNA molecules are then used to guide the synthesis of specific proteins is a process known as *translation*. Finally, an elaboration on the various mechanisms used by cells to control transcription and translation leads to the regulation of gene expression.

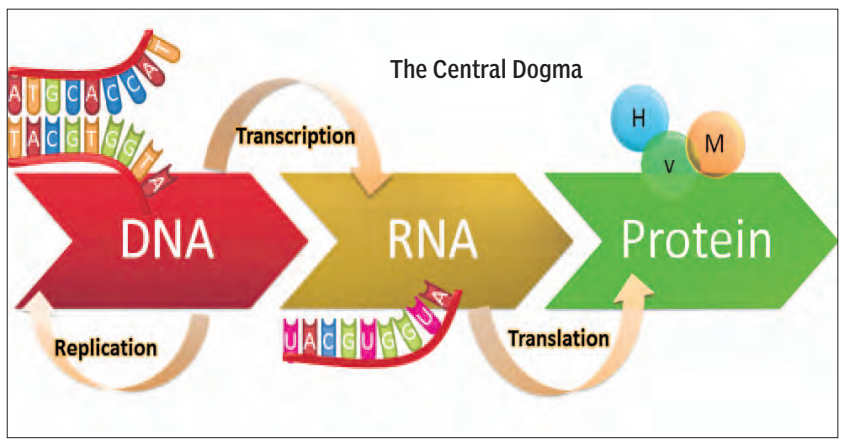
To put these topics in context, let us start with an overview of the roles played by DNA, RNA and proteins in gene expression. At the beginning of this unit, the flow of genetic information cells generally proceeds from DNA to RNA to protein DNA (more precisely, a segment of one DNA strand) which first serves as a template for the synthesis of an RNA molecule, which in most cases then directs the synthesis of a particular protein. (In a few cases, the RNA is the final product of gene expression and functions as such within the cell.) The principle of directional information flow from DNA to RNA to protein is known as the Central Dogma of molecular biology, a term coined by Francis Crick soon after the double-helical model of DNA was first proposed. This principle is summarised thus: The flow of genetic information involves replication of DNA, transcription of information carried by DNA into the form of RNA, and translation of this information from RNA into protein. The term transcription is used when referring to RNA synthesis using DNA as a template to emphasise that this phase of gene expression is simply a transfer of information from one nucleic acid to another, so the basic "language" remains the same. In contrast, protein synthesis is called translation because it involves a language change — from the nucleotide sequence of an RNA molecule to the amino acid sequence of a polypeptide RNA that is translated into protein and is called messenger RNA (mRNA) because it carries a genetic message from DNA to the ribosomes, where protein synthesis actually takes place.

In addition to mRNA, two other types of RNA are involved in protein synthesis: ribosomal RNA (rRNA) molecules, which are integral components of the ribosome, and transfer RNA (tRNA) molecules, which serve as intermediaries that translate the coded base sequence of messenger RNA and bring the appropriate amino acids to the ribosome. Note that ribosomal and transfer RNAs do not themselves code for proteins; thus, genes coding for these two types of RNA are examples of genes whose final products are RNA molecules rather than protein chains.

During the years since Crick's first statement of the central dogma, it has been refined in various ways. For example, many viruses with RNA genomes have been found to synthesise RNA molecules using RNA as a template. Other RNA viruses, such as HIV, carry out reverse transcription, whereby the viral RNA is used as a template for DNA synthesis – a "backward" flow of genetic information.

But in spite of these variations on the original model, the principle that information flows from DNA to RNA to protein remains the main operating principle by which all cells express their genetic information.

THE WRITER IS ASSOCIATE PROFESSOR, HEAD, DEPARTMENT OF BOTANY, ANANDA MOHAN COLLEGE, KOLKATA, AND ALSO FELLOW, BOTANICAL SOCIETY OF BENGAL, AND CAN BE CONTACTED AT tapanmaitra59@yahoo.co.in



Chemical waste disposal

REGULATED, ECO-FRIENDLY METHODS ARE IMPERATIVE, SAYS **PARTHASARATHI CHAKRABORTY**

The three-day international conference on waste management and recycling concluded in Sofia, Bulgaria, on 13 March with particular focus on chemical waste disposal and recycling technology transfer. In this regard, waste disposal and its sustainable management assumes alarming proportions in India, the huge chemical pollution in the Ganga river for decades remaining a matter of grave concern. Waste discharge from a large number of factories, including pharmaceuticals, chemicals, paper mills, leather, sugar and cotton industries, have aggravated the situation and it's a crying shame that this sacred river is being frivolously used as a drain.

Almost all the developing countries are confronting chemical waste disposal problems while India has a large legacy of abandoned hazardous sites, unregulated waste dumping and accords low priority to environmental issues. It must take chemical waste disposal issues very seriously and develop more cogent ideas for underground disposal, considering the technical barriers, legal basis, waste characterisation and conditioning. Chemical waste from school and college laboratories, research institutions and hospitals needs to be handled very carefully and disposed of properly. In formal educational institutions, little or no education and training is imparted on chemical safety. There are different ways to achieve disposal, especially for chemical waste from pharmaceutical factories and processing plants. These include incineration, recycling, neutralisation and land disposal.

Landfills are considered safe because they are deep underground and non-biodegradable. Adequate measures should be taken to ensure recycling and safety. The waste should not be allowed to contaminate oceans and rivers. Although millions of chemical compounds are known, only a small number has been tested as far as toxicity is concerned and chemical hazards may occur during storage, transportation, manufacture, handling, experimentation or even during disposal of waste. Toxic chemicals come in the form of gases, liquids, solids or dust and often find their way into the human system through breathing, swallowing or contract with the skin. Some corrosive chemical ingredients may be responsible for acute skin diseases and prolonged exposure to certain carcinogenics may lead to cancer. Bhopal in India experienced the worst disaster in the world resulting from the leakage of a highly toxic chemical gas, methyl isocyanate, commonly known as MIC — an ingredient necessary for preparing pesticides.

Thousands died and several thousand went blind. Moreover, the problem of occupational health is not necessarily confined within the four walls of factories. It has been reported that these hazards can often result in abortion, low fertility or even impotence. Some chemicals are responsible for birth defects in newborn babies and prolonged exposure to lead, mercury, organic lead compounds and radiation from X-Rays, carbon monoxide, chloroprene used in the manufacture of rubber, cadmium, chromates, fiber glass, silica, pesticides, etc, can cause occupational hazards for workers in chemical and pharmaceutical industries.


The Bhopal gas disaster remains a chilling reminder.

The disposal of radioactive waste out at sea, in space, deserts or deep geological formations have received considerable attention from nuclear scientists. Although the terms low level and high level for waste are entirely theoretical, different radioactive wastes are produced simultaneously at various stages of the production of nuclear power. Storing radioactive waste at a depth of 500-1,000 metres underground will evidently ensure long-term safety and burying such waste for a fairly long period will presumably not allow for the return of any residual radioactivity to the environment.

However, the mechanism of handling, storing and managing radioactive waste is complicated and it transpires that setting up underground laboratories for carrying out measurements, experiments and grading of such waste is indispensable. We need to maintain an ecological balance and provide a decent sustainable habitat for future generations and every living creature.

THE WRITER, A FORMER READER IN CHEMISTRY AT PRESIDENCY COLLEGE, KOLKATA, HAS BEEN ASSOCIATED WITH THE UNIVERSITY GRANTS COMMISSION

PLUS POINTS

Cervantes's tomb

Researchers in Spain say they've found remains that may be those of Miguel de Cervantes Saavedra, often referred to simply as Cervantes, on the grounds of a convent in Madrid. The *Don Quixote* writer died in 1616 and the location of his tomb had been a mystery since the late 17th century, when the Convent of the Discalced Trinitarians was rebuilt. Investigators searching for the literary giant's remains found a box that contained bones from 10 adults and five children in the ground beneath the crypt of a church at the convent. "We believe that some of the remains of Miguel de Cervantes are among the fragments," forensic anthropologist Francisco Etxeberria said on 17 March.

The forensics team investigating the site will now try to reconstruct the bodies and perform DNA testing to confirm the identities of the deceased. "It would of course have been better to find his remains complete, but I still have the feeling today of having reached the end of a journey," Fernando de Prato, the historian who led the search for Cervantes, told *The New York Times*.

Officials in Madrid are reportedly seeking to build a monument within the church where the remains were found to commemorate the writer in time for the 400th anniversary of his death come April.

BOB GRANT/THE SCIENTIST

Ancient beasts

Unable to isolate DNA from ancient fossils of the so-called South American ungulates, researchers turned to collagen, a major structural protein of bone that survives some 10 times longer than nucleic acids. Comparing the protein's sequence with that of extant species, they placed the extinct beasts — which included two 12,000-year-old specimens of *Toxodon*, a sort of hippo-rhino hybrid with rodent-like teeth, and two camel-like *Macrauchenia* specimens, which could not be dated — among the Perissodactyla, a clade made up of horses, tapirs and rhinos. The results, which refute earlier suggestions that the beasts belonged among elephants and manatees as part of the group *Afrotheria*, were published on 18 March in *Nature*.


Macrauchenia, a humpless camel with an extended snout.

"Compared to DNA, there's absolutely tons of (collagen)," study co-author Ian Barnes, a molecular evolutionary biologist at the Natural History Museum in London, said. He and his colleagues suggested that protein sequencing could be a complementary approach to DNA sequencing for studying ancient species, and one that allows them to look further back in time: to date, the oldest DNA ever recovered is between 450,000 and 800,000 years old, but proteins could be recovered from specimens millions of years old.

THE SCIENTIST

Power nap

At last, nodding off in the office or at school can be justified for its medical benefits, specifically for its positive effect on the memory. Scientists at Saarland University in Germany have found that taking a 45-60 minute power nap can boost a person's memory fivefold.

The study, published in the journal *Neurobiology of Learning and Memory*, tested the power of napping on 41 students. The students were taught 90 words and 120 unrelated word pairs, and then split into two groups. One group was allowed to take a nap, while the other was given a DVD to watch.

Both groups were tested on the words they had been taught, with the napping group performing the best.
Professor Axel Mecklinger, who supervised the study, said, "A short nap at the office or in school is enough to significantly improve learning success. Wherever people are in a learning environment, we should think seriously about the positive effects of sleep.

"Even a short sleep lasting 45 to 60 minutes produces a fivefold improvement in information retrieval from memory."

THE INDEPENDENT