A brewer's biochemistry

By **Professor Charles Bamforth**
Department of Food Science and Technology,
University of California, Davis, USA,

Part 4: Nucleic acids

This is the fourth in a series of articles aiming to position malting and brewing in biochemical terms for the benefit of those who have received no training in this area of science. Readers lacking a formal scientific training will find basic chemical principles described in the first article of the series.

There are two types of nucleic acid in living organisms:

- deoxyribonucleic acid (DNA)
- ribonucleic acid (RNA)

In all the types of organism that a brewer is interested in (barley, hops, yeast, spoilage organisms, themselves) the DNA comprises the information store for the cells. RNA constitutes the machinery than transcribes and translates that information into cellular activity.

The essential events surrounding these molecules are:

- How is the information passed on from one cell to its daughters when the cell divides, e.g. during growth?
- How is the message in DNA translated into cellular activity – i.e. precisely how does the code get interpreted and acted upon?

For each question to be answered clearly, we need to know the basics of nucleic acid structure.

Structure of nucleic acids

Nucleic acids are polymers built up of three different types of units:

- sugar
- phosphoric acid
- bases

The sugars are ribose in RNA (ergo *ribo*nucleic acid) and deoxyribose in DNA (*deoxyribo*nucleic acid). Reference to article 2 in this series will remind you of the structure of ribose, a sugar with 5 carbons and five places in its ring structure (i.e. it is a pentose).

Deoxyribose, as you might guess, is just ribose missing one of its oxygen atoms (on carbon atom number 2).

Linking carbon atom number 3 on one sugar ring to carbon atom number 5 on the next (go to article 2 for the numbering crib) is the phosphoric acid. In this way long chains are formed.

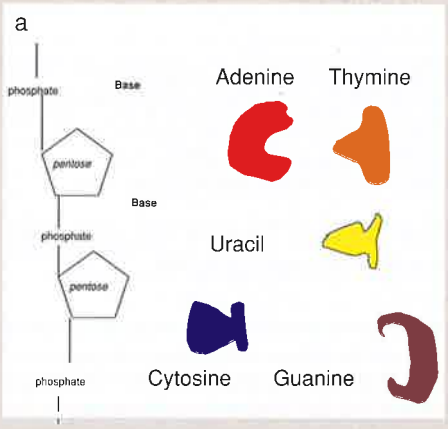


Figure 1a: The basic structural elements of nucleic acids. The configuration of sugars, phosphates and bases and a schematic representation of the importance of base "shape" to allow fitting of cytosine and guanine or of adenine with either thymine (in DNA) or with uracil (in RNA).
Figure 1b: allowable fits between bases.
Figure 1c: The pairing of bases on adjacent strands of DNA and an illustration of the double helix. In the type of illustration shown on the left the vertical lines represent the sugar and phosphate components.

Attached to carbon atom number 1 are the bases. These are the most complicated bits of the picture, and we won't bother with their precise chemical structure. Suffice to say that both DNA and RNA each contain four types of base. In DNA they are called adenine, thymine, cytosine and guanine. In RNA thymine is replaced by uracil.

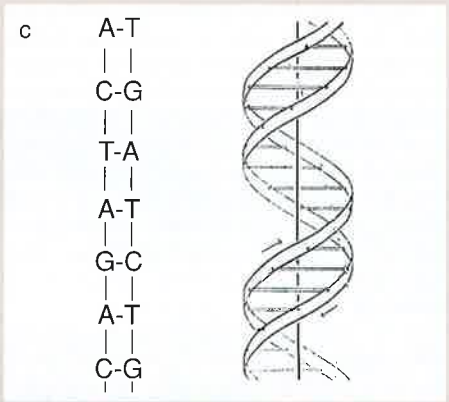
The only other structural feature I am going to mention is that RNA molecules comprise individual, single chains whereas DNA molecules consist of two of these chains intertwined in the form of a double helix (Fig 1).

(Many folk have heard of the proposal of this by the American James Watson and Englishman Francis Crick working in Cambridge, England in 1953.) The two chains are linked relatively loosely through an association between the bases on the separate strands. More specifically adenine always latches on to a thymine on the opposite chain, whereas cytosine always reaches out to a guanine. These couples fit together snugly and make for a well-ordered molecule.

The importance of base-base recognition

This selectivity of binding of the bases is the very key to the genetic code. It allows us to answer the questions posed earlier.

When a cell divides this is preceded by a replication of the entire DNA in the cell. In a process catalysed by specific enzymes the double helix is unwound and to each strand is linked in a new strand, leading eventually to the production of two "daughter" strands, each identical to the "parent" (Fig 2).



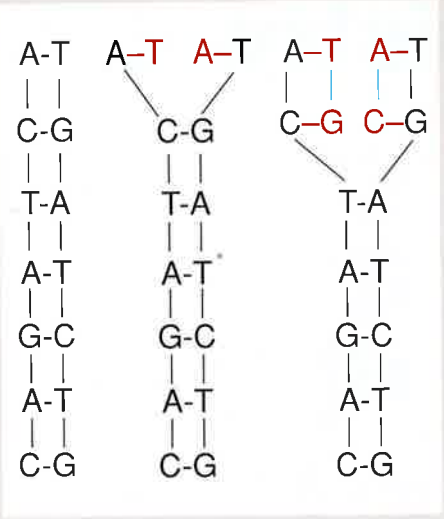
You will appreciate that if there is, say, an adenine in position 1 on one strand, then when the new strand starts to be linked in, the machinery knows to attach a thymine to it. If the next base is a guanine then the machinery knows what to do: "ah! cytosine".

And so, in this way, the genetic code is replicated and passed on when a cell divides.

In the same way, the message from DNA is passed on to RNA to initiate the triggering of cellular functions. An enzyme unzips the DNA molecule and, using one of the DNA strands as the template, an enzyme stitches together molecules of RNA.

This time, if it sees an adenine it splices a uracil in to the chain, and it also pops in ribose instead of deoxyribose.

This process of converting the master code in DNA to the messenger (RNA) is called *transcription*. There are many individual messengers transcribed from a single DNA molecule, each of the messengers carrying the entire code needed for one protein. (You will recall from article 1 that the workhorses of the



cell are the proteins – they are the enzymes, the structural units, the transporters etc. Thus it is the proteins that are the manifestation of the code, whether they are the haemoglobin of a red blood cell, or the flagellum of a motile bacterium, or the hordeins of barley or the surface antigens on yeast.) The sequence of DNA that is transcribed into an individual messenger RNA molecule is called a *gene*.

The process whereby the code carried by the RNA is shifted into a protein is called *translation*. The key is the sequence of bases (Fig 3). Each sequence of three bases indicates a specific amino acid. It is called a *codon*. Thus, for instance, if the reading machinery (another enzyme) sees guanine, guanine, guanine one after the other then it calls for the amino acid glycine, which is actually delivered by another type of RNA called transfer RNA. (There is one type of *transfer* RNA for each amino acid – clever, isn't it!)

If the next three bases are uracil, cytosine, and guanine then the reading machinery alerts the serine carrier to get itself to the party, and

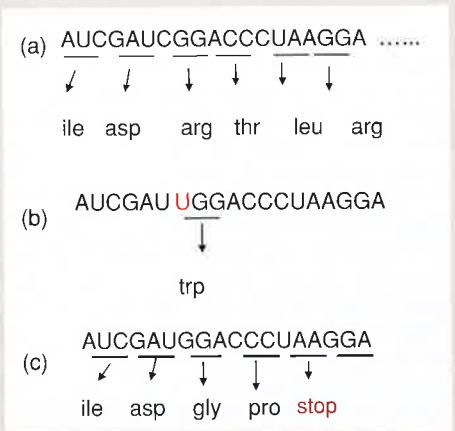


Figure 4. The impact of mutagens.
(a) the non-mutated message
(b) message obtained in messenger RNA where a base change has occurred (U for C)
(c) message obtained in messenger RNA where a base (C) has been deleted. Note that UAA is the codon that says "stop reading" – so every codon after that point would not be read.

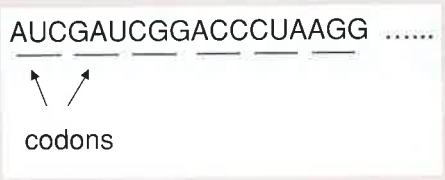


Figure 2 (left). The replication of DNA
An enzyme starting at one end of the double helix starts to unzip the two strands, and new complementary bases, together with deoxyribose and phosphate are pieced in.

Figure 3 (above). The basis of the genetic code translated into messenger RNA.

the enzyme splices it on to the glycine. And so on, until the protein molecule is completed.

And so you can see how everything hinges on the precise sequence of the bases in the DNA. One small change – just in one base – could wreak havoc because the wrong amino acid will be "screwed" into the protein and this can easily alter the structure of the protein and hence its function (as we saw in the first article in this series).

Chromosomes

The DNA in the cells that concern us (apart from bacteria – and most brewers *hope* not to be troubled by them!) is located in the chromosomes in the cell nucleus. We've all got 'em. You and I have 46 chromosomes in most of our cells – in fact two complete sets of 23 different chromosomes.

It's just our reproductive cells that have one set of 23 chromosomes – in the production of eggs and sperm just one copy of each chromosome is packaged. When these cells get together in a fertilised egg, then one copy of each chromosome comes from each parent, returning to the 46-chromosome complement. (If it were otherwise then every time a young'un was born they'd have twice as many chromosomes and in all the aeons since man has been around they'd have accumulated a fair few chromosomes by now!)

Saccharomyces cerevisiae has 16 chromosomes. Those strains that are used for brewing tend to be *polyploid*, which means that they have multiple copies of each chromosome – perhaps 3 or 4. Some are *aneuploid*, which means that they have different numbers of copies of the various chromosomes – perhaps 3 of one and 4 of another.

Genetic modification

I have no intention of leading this article into any areas of controversy. I'll simply state some facts and if the reader chooses to make a judgement one way or another, well that's up to them.

The first thing to say is that DNA is a reasonably sensitive molecule. Its bases can react with various outside influences and become changed. These outside factors include ultra-violet radiation and diverse chemical compounds that are called *mutagens* because they mutate the DNA and, as a consequence, influence its behaviour. The carcinogens act in this way.

The mutagens might lead to a loss of bases or to a change in bases. For example nitrous acid converts adenine to something called hypoxanthine. The latter group binds to cytosine and not thymine.

And so, when DNA replicates, the new strand will contain cytosine instead of thymine, and this change will carry on to the next generation: the cytosine will now be matched by a guanine on the partnering strand, so we now have a cytosine-guanine pair where there used to be an adenine-thymine pair. When that mutated DNA was translated into RNA, then the message passed on would have been different.

Let's assume that the altered base message in messenger RNA was in a triplet that should have read cytosine-guanine-guanine (Figure 4a). That codes for arginine. By exchanging the cytosine for uracil, the reading machinery is kidded in to putting a tryptophan residue into the protein (Figure 4b). The likelihood is that this will change tremendously the resultant protein, and likely as not it won't function.

The situation is even worse if the mutation leads to a deletion of one of the base pairs (Figure 4c). The effect is to shift the reading frame, so that all the amino acids coded after the point where the deletion occurred are probably

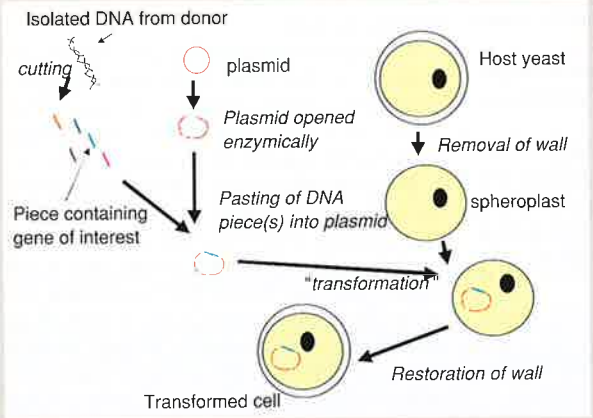


Figure 5 Recombinant DNA technology
The yeast to be transformed has its wall removed using selected enzymes (the wall would be a barrier to DNA entry). Other enzymes are used to "chop" the donor DNA into pieces. The gene of interest will be on one of these pieces. The various pieces are mixed with an opened up ("cut") small carrier ring of DNA called a plasmid under conditions where the ends of the individual DNA pieces and the opened ends of the plasmid will seal together. The plasmid originates from yeast. This plasmid has certain properties that allow it to be retained by a yeast cell to which it is introduced. The wall-free yeast is mixed with the pool of plasmids. Those yeast cells that have received the DNA of interest (and no other) can be selected. The wall is then regenerated. Voila: transformed yeast.

going to be wrong.

Don't get the impression that all mutation is a bad thing. Ultimately it's the reason why we're all different and why evolution takes the course it does. The impacts of mutation can be beneficial. If the altered DNA codes for a ***better*** protein – for instance, changing a single amino acid might increase the heat tolerance of an enzyme – then it would be an advantage, e.g. to an organism adapting to growth in higher temperature locations. (This type of thing can be done in the lab – it's called protein engineering.) Of such changes over a relative eternity did long necks on giraffes develop.

Mutation has been used by brewing scientists in the past to eliminate unwanted activities, e.g. removal of the tendency to develop hydrogen sulphide. And it can be used profitably in research studies into the understanding of metabolism.

We've all heard of cloning. What is it?

If my molecular biologist friends (and I do have some) will forgive me, it's essentially a scissors and paste job on DNA (Fig 5). Take a donor organism and a recipient organism. Extract and chop up (using a certain type of enzyme purchased at exorbitant price) the DNA from the former and attach the various pieces to something called a *vector*.

This is frequently either a virus or a bit of circular DNA that can move into and out of cells, called a *plasmid*. The plasmid-donor DNA is added to the host organism, that has had its wall removed to make it more penetrable. Inside the cell the bits of DNA (including the gene we are interested in) will either integrate into the host chromosomes or will remain in the vector form, provided there is a positive pressure to keep them there. That might take the form of tagging on to the plasmid some other factor. A popular one is the gene for copper resistance.

If the host cells are then grown in the presence of a high level of copper, then only the cells containing the plasmid will survive – and they will also possess the gene we are interested in. Alternatively the host organism might be one used because it is deficient in something, for example the ability to use a certain amino acid because they lack a certain enzyme.

The plasmid or vector would then be engineered so as to possess the gene for the missing enzyme and so a receiving organism now able to use that amino acid must have taken up that gene and, whatismore, must retain it if it is to continue to grow on a diet that deliberately excludes the amino acid that it can't make.

The cells that have successfully received the gene of interest (and whichever other genes are necessary for reasons just mentioned) are purified. If they are microbes then they can be used (law permitting) straight off. If they are cells derived from plants then there is the issue of regenerating the whole plant. Too complex to tackle here.

One thing you might have clocked though is that it must easier to jiggle with the DNA of simple organisms like bacteria than with the likes of yeast. It's all about ploidy. If you try to isolate a mutant of a polyploid organism you have problems: if you knock out one gene, there is at least one other copy of the same gene to fall back on. And the shortage of mutants makes it a challenge to select organisms that have received an extra gene – go back two paragraphs.

You tell me whether the targeted approaches described in this section are more or less controlled than mixing completely separate pools of DNA as occurs in more traditional approaches to developing new organisms. Note that I have **not** asked you to comment on which is the more pleasurable technique. ■