A DICTIONARY OF GENETIC ENGINEERING

STEPHEN G. OLIVER

JOHN M. WARD

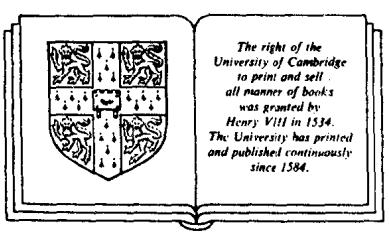
A DICTIONARY OF GENETIC ENGINEERING

STEPHEN G. OLIVER

Department of Biochemistry and Applied Molecular Biology
University of Manchester Institute of Science and Technology

JOHN M. WARD

Department of Biochemistry, University College, London



CAMBRIDGE UNIVERSITY PRESS
Cambridge
London New York New Rochelle
Melbourne Sydney

How to use this dictionary

The definitions in this dictionary are arranged in strict alphabetical order. If you wish to look up a term which includes Greek letters or Roman or Arabic numerals, you should first transliterate it into Latin script. For example:

ϕX174	will be found under p-h-i-x-o-n-e-,
Qβ	will be found under q-b-e-t-a,
λ	will be found under l-a-m-b-d-a,
5′	will be found under f-i-v-e-p-r-i-m-e,
RP4	will be found under r-p-f-o-u-r,
cI	will be found under c-o-n-e.

Figures

A

actinomycete A member of the bacterial class, Actinomycetales. These Gram-positive, spore-forming, mycelial bacteria are abundant in soils and composts. Many actinomycetes produce volatile fatty acids which give earth its characteristic odour. In their natural environment they are responsible for the breakdown and recycling of substances such as cellulose, chitin and keratin. Actinomycetes, especially *Streptomyces* species, produce most of the world's antibiotics and many different species are cultivated on a large scale for the commercial production of clinically useful antibiotics. In genetic engineering certain streptomycetes have been used to develop a host-vector system for cloning.

activator (i) In molecular biology, a protein which binds to a site on DNA upstream of a gene and activates transcription from that gene. (ii) In enzymology, a small molecule which binds to an enzyme and increases its catalytic activity.

agarose gel An inert matrix used in the electrophoretic separation of nucleic acid molecules on the basis of their size or conformation. Gels may be cast as tubes or slabs although the latter now predominate. Molecules are visualised in the gel by the ultraviolet fluorescence of ethidium bromide which is either included in the running buffer or used to stain the gel after electrophoresis. (*See* comb, LGT agarose, power pack, Tris-acetate buffer, Tris-borate buffer)

Agrobacterium rhizogenes A species of Gram-negative, rod-shaped soil bacteria closely related to *Agrobacterium tumefaciens*. *A. rhizogenes* often harbours large plasmids, called Ri plasmids, which are closely related to Ti plasmids. The combination of *A. rhizogenes* and an Ri plasmid can cause a tumourous growth known as hairy root disease in certain types of plants.

Agrobacterium tumefaciens A species of soil bacterium which, when it contains a Ti plasmid, can infect the stems of many plants and form crown gall tumours.

agropine A rare amino acid derivative which is produced by a certain type of crown gall tumour. The genes responsible for the synthesis of agropine are part of the T-DNA from a Ti plasmid.

alkaline hydrolysis The use of a high pH to degrade or hydrolyse a bond. DNA is not hydrolysed at high pH while RNA will be degraded to

mononucleotides. RNA has a 2′ hydroxyl group which, at high pH, will attack the 3′ phosphodiester bond. DNA has no hydroxyl at the 2′ position and is thus stable to alkaline hydrolysis.

alkaline phosphatase An enzyme which removes the 5′ terminal phosphate groups from linear DNA molecules. It is used to prevent the religation of plasmid vector molecules following cleavage with a restriction endonuclease. This increases the chance that intact circular molecules generated by the ligase reaction are recombinant.

α-peptide A short (185 amino acids) amino-terminal fragment of the enzyme β-galactosidase. The α-peptide can combine with, and restore the activity of, a non-functional β-galactosidase enzyme which has a defective N-terminus. M13mp phage cloning vectors make use of this complementation as they carry the gene for the α-peptide.

amber A mutation which creates the stop codon UAG in the coding region of a gene. This leads to the synthesis of a truncated protein. These mutations can be suppressed by certain mutant tRNA species which permit the incorporation of an amino acid in response to the UAG stop codon thus allowing the protein to be completed. Amber mutations are deliberately incorporated into certain λ phage cloning vectors so that they can only be propagated in hosts which will suppress the amber mutation. This is a type of biological containment. In genetic notation, amber is abbreviated to *am*, thus an amber mutation in gene *S* is written as *Sam*.

Ampicillin resistant Ap^r; Ampicillin sensitive Ap^s Resistance or sensitivity to the lethal effects of the antibiotic ampicillin. Ampicillin is a β-lactam antibiotic and resistance is (often) mediated by a class of enzymes called β-lactamases which are secreted either into the periplasmic space of Gram-negative bacteria or into the medium in Gram-positive bacteria. The cloning vector pBR322 contains an ampicillin resistance gene.

amplification An increase in the copy number of a gene or plasmid. (*See* chloramphenicol amplification)

angle rotor, fixed-angle rotor A centrifuge rotor in which the wells holding the tubes are drilled at an angle to both the axis of rotation and the lines of centrifugal force. Angle rotors were originally used simply to pellet material in the technique of differential centrifugation. They are now used routinely for density gradient centrifugation. The gradients formed in angle rotors are not linear, but permit a great deal of resolution over a narrow density range.

anneal A verb meaning to hybridise nucleic acid molecules.

antibiotic A substance, produced by one organism, which inhibits or kills another organism. Most antibiotics are active against bacteria and are produced by fungi or streptomycetes. (*See* ampicillin resistance, tetracycline resistance)

antibiotic resistance Resistance to the lethal effects of an antibiotic. There are five main mechanisms of antibiotic resistance: (i) inactivation of the antibiotic; (ii) a reduction in cellular uptake or an increase in cellular excretion of the antibiotic; (iii) production of an altered target protein which no longer binds the antibiotic; (iv) overproduction of the target protein such that the antibiotic is 'titrated out'; (v) elaboration of some alternative enzyme or pathway which is not susceptible to the antibiotic.

anticodon The three nucleotides in a tRNA molecule which base pair with the complementary nucleotides forming the codon within mRNA. It is this codon–anticodon interaction, taking place at the ribosome, which ensures that the correct amino acid is inserted into the growing polypeptide chain.

anti-terminator A type of protein which enables RNA polymerase to ignore certain transcriptional stop or termination signals and read through them to produce longer mRNA transcripts.

Ap^r, Ap^s Ampicillin resistant, ampicillin sensitive.

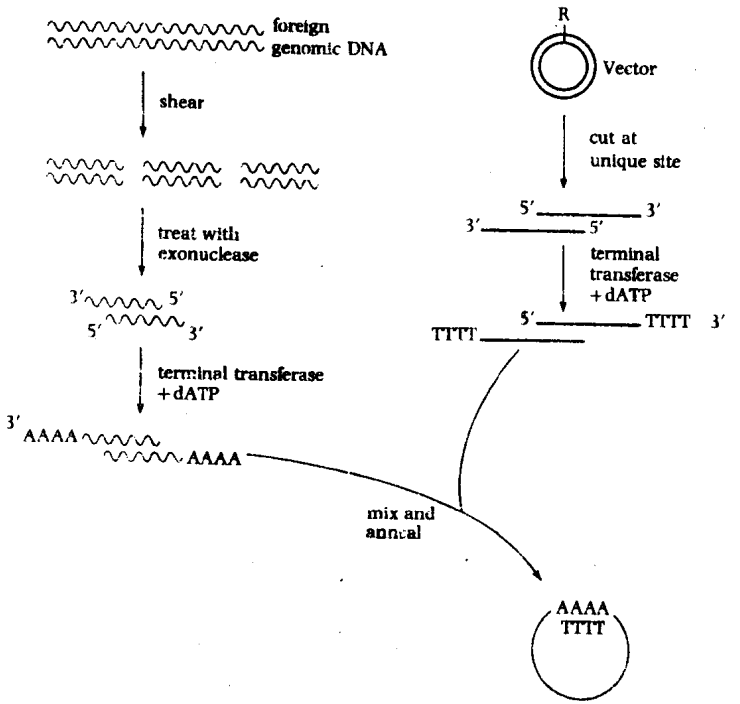
Arabidopsis thaliana Thale Cress. A small, fast-growing dicotyledonous plant, a member of the Cruciferae. It is a favoured experimental organism among plant molecular biologists.

ARS Autonomously replicating sequence (or segment). A term, commonly used in yeast molecular biology, for a DNA sequence which will support independent replication of a plasmid in a host yeast cell. Some *ARS* sequences may be cloned from either the yeast itself or from some other organism. They are thought to represent origins of DNA replication, although it is not clear that they are actually used as such in their parent genomes. Recombinant plasmids which rely on an *ARS* sequence for their replication are intrinsically unstable in yeast (see YRp). The term may, in principle, be applied to sequences which promote plasmid replication in any organism.

A's and T's method A method whereby random DNA fragments can be cloned into a vector molecule. Random DNA fragments generated by mechanically shearing or sonicating genomic DNA are treated with λ-exonuclease in order to generate 3′ single-strand tails. These tails are

then extended by the addition of deoxyadenosine residues using the enzyme terminal transferase (calf thymus terminal deoxynucleotidyl transferase). The vector molecule is cut at a unique restriction site to generate 3′ tails again. These tails are extended with deoxythymidine and terminal transferase. The foreign and vector DNA now bear complementary 3′ tails and can be annealed together. If the tails are long enough the molecule will be sufficiently stable not to require treatment with ligase before being used to transform a host organism. While certain advantages attach to the cloning of truly random fragments, the technique has the drawback that it is difficult to retrieve the insert from the recombinant molecule. However, the lower melting temperature of the oligo dA·dT junction sequences may be exploited by partially denaturing and then cutting with a single-strand-specific endonuclease such as S1.

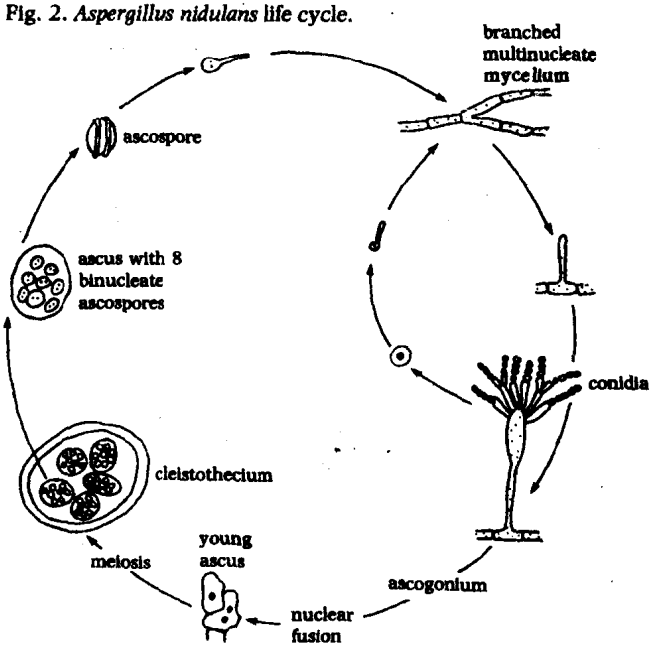
Fig. 1. A's and T's method.



Aspergillus A filamentous fungus of both industrial and genetic importance. Two imperfect (non-sexual) species, *A. niger* and *A. oryzae*, are used in the production of citric acid, industrial enzymes and fermented foods. The perfect (sexual) species *A. nidulans* has been an

important research tool in both biochemical and mitochondrial genetics. Its life cycle is given below.

Fig. 2. *Aspergillus nidulans* life cycle.



attenuator A sequence which is found upstream of a bacterial operon which encodes enzymes involved in amino acid biosynthesis. The attenuator regulates the expression of such operons by determining whether the mRNA molecules containing their transcripts will be completed or not. The attenuator sequence contains a short open reading-frame which includes several codons specifying the amino acid which the gene products of the regulated operon synthesise. If the concentration of this amino acid in the cell is limiting, then a ribosome will pause in its translation of the attenuator sequence in the nascent message. The continued presence of the ribosome in the attenuator region favours the formation of one of two possible secondary structures in the attenuator transcript. This secondary structure permits the RNA polymerase to extend the mRNA chain to include the operon transcript. If, on the other hand, there is a plentiful supply of the amino acid in the cell then the alternative secondary structure is formed within the attenuator mRNA. This structure is recognised as a terminator by the RNA polymerase and transcription is halted (attenuated) before the operon is transcribed. In this way, mRNA for

the amino acid biosynthetic enzymes is only produced when there is a high demand for amino acid biosynthesis.

autoradiography A method for detecting the location of a radioisotope in a tissue, cell or molecule. The sample is placed in contact with a photographic emulsion, usually an X-ray film. The emission of β-particles from the sample activates the silver halide grains in the emulsion and allows them to be reduced to metallic silver when the film is developed. In genetic engineering, autoradiography is most commonly used to detect the hybridisation of a radioactive probe molecule to denatured DNA in either the Southern transfer or colony hybridisation procedures.

***Ava* I** A type II restriction enzyme from the blue-green alga *Anabaena variabilis*. *Ava* I recognises the DNA sequence below and cuts at the sites shown by the arrows:

```
      ↓
5′  C Py C G Pu G  3′
3′  G Pu G C Py C  5′
          ↑
```

where Py and Pu stand for pyrimidine and purine. (A complete list of restriction enzymes can be found in Appendix 1)

B

Bacillus A genus of rod-shaped, Gram-positive, spore-forming bacteria which are widespread in nature. Some members of the genus are thermophiles, e.g. *Bacillus stearothermophilus* and can grow at 60 °C, others are pathogenic for Man and animals, e.g. *Bacillus anthracis* the causative agent of anthrax. Many *Bacillus* species secrete large amounts of extracellular enzymes which, in their natural habitat, allow them to degrade high-molecular-weight substrates such as starch and protein. *Bacillus* species are of great interest in biotechnology as several are used for the production of enzymes of industrial importance and other species produce antibiotics which can be used clinically.

Bacillus subtilis A Gram-positive, spore-forming, rod-shaped bacterium. The genetics and physiology of *B. subtilis* have been extensively studied and this organism has become widely established as a vehicle for genetic engineering. There are many cloning vectors available, both phage and plasmid, which will replicate in *B. subtilis*. It is non-pathogenic for Man and animals and can be consumed in large quantities with no ill effects. This, together with the fact that *Bacillus* species excrete considerable amounts of extracellular proteins, led to the organism becoming one of the main host systems for the cloning of foreign genes.

bacterial alkaline phosphatase *See* BAP.

bacteriocin A toxin or antibiotic, produced by one class of bacteria, which kills other, usually closely related, bacteria. (*See also* colicin, Col factor, Col E1)

bacteriophage A virus which infects bacteria. Commonly known as a phage. (*See* lysogeny, lytic infection, plaque, T4)

BAL-31 A nuclease from the bacterium *Alteromonas espejiana* BAL-31. One of its three activities is the simultaneous degradation of the 3′ and 5′ termini of duplex DNA. BAL-31 is used to create deletion mutants *in vitro* by progressively shortening restriction fragments from both ends. The shortened molecules can then be religated with T4 DNA ligase.

***Bam* HI** (pronounced bam to rhyme with ham) A type II restriction enzyme from the bacterium *Bacillus amyloliquefaciens* H which recognises the DNA sequence below and cuts at the sites indicated by the arrows:

```
      ↓
5′  G G A T C C  3′
3′  C C T A G G  5′
          ↑
```

The four base pair sticky ends are complementary to those produced by the enzymes *Sau* 3A, *Bgl* II, *Xho* II, *Mbo* I and *Bcl* I and so fragments produced by any of these enzymes can be cloned into a *Bam* HI site. The popular cloning vector pBR322 has a single *Bam* HI site in the tetracycline resistance gene. Many vectors are constructed with a *Bam* HI site to facilitate cloning of a wide variety of DNA fragments generated by the enzymes listed above. (A full list of restriction enzymes can be found in Appendix 1)

banjo A descriptive term for a stem-loop structure in a nucleic acid molecule.

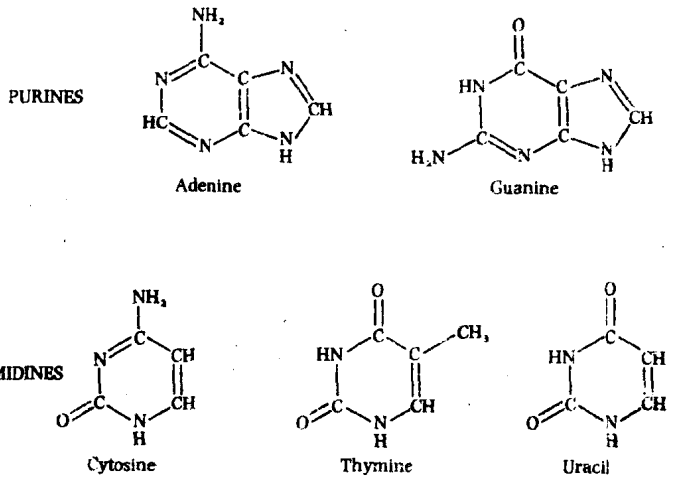
bank, gene bank A collection of recombinant DNA molecules containing inserts which together comprise the entire genome of an organism.

Also used as a verb, as in 'We'll bank *Aspergillus* in YIp5 and test for *ARS* activity in yeast.'

BAP (pronounced bap, to rhyme with tap) Bacterial alkaline phosphatase. An enzyme, isolated from *E. coli*, which removes 5′ terminal phosphate groups from DNA chains. It is used to prevent the recircularization of vector molecules during gene cloning experiments.

base The heterocyclic compounds which are the constituents of all nucleic acids. There are five common bases. Three, adenine, guanine

Fig. 3. Base structures.



and cytosine, are found in both DNA and RNA; thymine is found only in DNA and uracil only in RNA. A base plus a sugar (deoxyribose in DNA, ribose in RNA) is referred to as a nucleoside. A base plus sugar plus phosphate(s) is a nucleotide. The structure of the five common bases is given above.

base pair A pair of nucleotides held together by hydrogen-bonding which are found in double-stranded nucleic acids. DNA contains the base pairs A=T and G≡C, while RNA contains A=U and G≡C. (The lines indicate the number of hydrogen bonds.) The size of a nucleic acid molecule is often given in terms of the number of base pairs (bp) it contains. (*See* kb)

Benton–Davis technique *See* plaque hybridisation.

Berk–Sharp mapping, S1 mapping *See* S1 nuclease.

β-galactosidase An enzyme which will cleave lactose into glucose and galactose. The most commonly used β-galactosidase gene is from the *Escherichia coli lac* operon.

β-lactamase(s) A class of enzymes that inactivate β-lactam antibiotics (the penicillins). These enzymes are either periplasmic (in Gram-negative bacteria) or extracellular (in Gram-positive bacteria). The ampicillin resistance gene of pBR322 encodes a certain type of β-lactamase.

***Bgl* II** (pronounced buggle or baygel) A type II restriction enzyme from the bacterium *Bacillus globigii*. The enzyme recognises the DNA sequence shown below and cuts at the sites indicated by the arrows:

```
      ↓
5′ A G A T C T 3′
3′ T C T A G A 5′
        ↑
```

The sticky ends produced by *Bgl* II are complementary to the ends produced by the enzymes *Bam* HI, *Bcl* I, *Xho* II, *Mbo* I and *Sau* 3A. Thus, fragments produced by any one of these enzymes will have single-strand extensions which can anneal with the sticky ends on fragments produced by any other of the enzymes above. (A full list of restriction enzymes can be found in Appendix 1)

bifunctional vector or **plasmid** A DNA molecule able to replicate in two different organisms, e.g. in *E. coli* and yeast or *E. coli* and *Streptomyces*. These molecules are thus able to 'shuttle' between the two alternative hosts and are therefore also known as shuttle vectors.

It is usual to 'grow' DNA in *E. coli* for the genetic transformation of the alternative host.

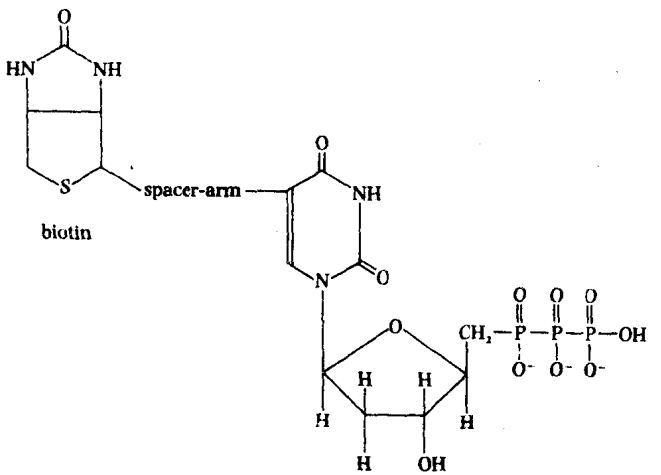
Biodyne™ A form of activated nylon filter which can be used in place of nitrocellulose in the Southern blotting procedure. The advantage of biodyne membranes is that they are more robust than nitrocellulose and the same filter may be used in a large number of consecutive hybridisations.

biological containment A strategy which reduces the risks of recombinant molecules being propagated within microorganisms found in the general environment. Biological containment involves the use of vector molecules and host organisms which have been genetically disabled such that they can only survive in the peculiar conditions provided by the experimenter and which are unavailable outside the laboratory. (*See* physical containment)

biotinylated-DNA A DNA molecule labelled with biotin by incorporation of biotinylated-dUTP into a DNA molecule. It is used as a non-radioactive probe in hybridisation experiments such as Southern transfer. The detection of any hybrids uses a complex of streptavidin-biotin-horseradish peroxidase which will give a fluorescent green colour where hybrids have formed.

biotinylated-dUTP A nucleoside triphosphate with the vitamin biotin attached, via a spacer arm, to the dUTP (*see* Fig. 4). It can be incorporated into a DNA molecule by nick translation and then the biotinylated-DNA formed from such a reaction can be used as a probe in a hybridisation experiment. It is a non-radioactive alternative to labelling with ^{32}P.

Fig. 4. Biotinylated d-UTP.



Birnboim–Doly procedure A rapid method for the purification of plasmid DNA, often used to screen recombinant colonies to determine the size of a DNA fragment inserted into a vector. The technique uses an alkaline denaturation step followed by a rapid renaturation to achieve the removal of the majority of the chromosomal DNA and most of the cell's RNA. It can be used for mini-preps (*q.v.*) or scaled up to give large quantities of fairly pure plasmid DNA. The DNA produced at the end of the procedure is usually pure enough to be digested by a restriction enzyme.

blot (i) As a verb, this means to transfer DNA, RNA or protein to an immobilising matrix such as DBM-paper, nitrocellulose or biodyne membranes. (ii) As a noun, it usually refers to the autoradiograph produced during the Southern or Northern blotting procedure. As in 'This blot demonstrates that the transforming DNA has been inserted into the chromosome.' (*See* Southern, Northern, Western blots)

blunt ends, flush ends Certain restriction enzymes, e.g. *Hae* III, generate DNA fragments which are perfectly base-paired along their entire length. The ends of such molecules are known as blunt or flush ends since they do not carry single-stranded extensions. Blunt ends may be generated artificially by removing single-stranded extensions with S1 nuclease. Blunt-end ligation is the process of joining two DNA molecules with blunt ends using DNA ligase. The process requires higher concentrations of both DNA and ligase than does the ligation of molecules with cohesive ends. The process is known colloquially as 'blunt-ending' and the term is also used as a verb as in: 'We trimmed off the 3′ extensions and blunt-ended the fragment into the *Hae* III site on our vector.'

bovine papilloma virus, bpv A group of viruses which cause warts (papillomas) in cattle and which will replicate in a wide variety of mammalian cells. These viruses do not lyse their hosts but replicate as plasmids with a copy number of 10–200 per cell. This stable association with their host has led to derivatives of bpv being used as cloning vectors for mammalian cells.

bp An abbreviation for base pair when used as a measure of the size of a double-stranded nucleic acid.

bpv *See* bovine papilloma virus.

broad host range A term used to describe a plasmid or phage which can replicate in a large number of different species.

B. subtilis *See Bacillus subtilis.*

buffer A solution containing a mixture of a weak acid and a base which resists changes in pH and is therefore able to provide a favourable environment for enzymic reactions.

buoyant density The intrinsic density which a molecule, virus or subcellular particle has when suspended in an aqueous solution of a salt such as CsCl, or a sugar, such as sucrose. The buoyant density of DNA is *ca.* 1.7 g cm^{-3} and different species have a characteristic buoyant density which reflects the proportion of G·C base pairs in the molecule. The greater the proportion of G·C the greater the buoyant density of the DNA molecule. (*See* density gradient centrifugation)

C

calf intestinal alkaline phosphatase (CIAP, sometimes pronounced 'chap' or 'sip') An enzyme which removes 5′ terminal phosphate groups from DNA molecules (*see* BAP). It has the advantage over the equivalent bacterial enzyme (BAP) that it can be inactivated by heat treatment at 70 °C.

cAMP *See* cyclic AMP.

CaMV *See* cauliflower mosaic virus.

canonical sequence A prescribed or archetypical nucleotide or amino acid sequence to which all variants are compared.

cap A structure found at the 5′ end of eukaryotic mRNA molecules. It consists of the modified base 7-methylguanosine joined in the opposite orientation, i.e. 5′ to 5′ instead of 5′ to 3′, to the rest of the molecule via three phosphate groups:

^{m7}G(5′)ppp(5′)Nmp....

The 7-methylguanosine cap is added to the primary transcript in the nucleus while it is being spliced and polyadenylated.

CAP, catabolite activator protein A protein studied mainly in *Escherichia coli* but with analogues in other bacteria. CAP binds cyclic AMP (cAMP) and then activates transcription from a large set of genes and operons involved in the catabolism of carbon compounds. These genes such as the *lac* operon and the *mal T* gene are thus under the positive control of the CAP–cAMP complex. The CAP–cAMP complex appears to stimulate the initial binding of RNA polymerase to the promoter. RNA polymerase has a low affinity for a CAP-activated promoter in the absence of the CAP–cAMP complex and hence the cellular concentration of cAMP controls the expression of such genes.

CAP binding site The nucleotide sequence upstream of the coding sequence of a bacterial gene or operon to which the catabolite activator protein can bind. This site only occurs before genes or operons which are under the positive control of this protein. In the presence of cyclic AMP the catabolite activator protein stimulates transcription from promoters which have a CAP binding site.

cap site The probable transcription initiation site of a eukaryotic gene. The cap is added to the 5′ end of the mRNA molecule; most eukaryotic mRNAs have an A as the first nucleotide and the cap is added to that.