Second Edition

# Dictionary of DNA and GENOME TECHNOLOGY

Paul Singleton

WILEY-BLACKWELL

# Dictionary of DNA and Genome Technology

## Second Edition

Paul Singleton



WILEY-BLACKWELL
A John Wiley & Sons, Ltd., Publication

First edition 2008
Second edition 2010

Wiley-Blackwell is an imprint of John Wiley & Sons, formed by the merger of Wiley's global Scientific, Technical and Medical business with Blackwell Publishing.

*Registered office*: John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

*Other Editorial Offices:*
9600 Garsington Road, Oxford, OX4 2DQ, UK

111 River Street, Hoboken, NJ 07030-5774, USA

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com/wiley-blackwell



***Library of Congress Cataloging-in-Publication Data***

Record on file

ISBN 978-0-470-74732-2 (H/B)  978-0-470-74731-5 (P/B)

A catalogue record for this book is available from the British Library.

Set in 8/9 Times by Integra Software Services Pvt. Ltd., Pondicherry, India.

Printed in Great Britain by CPI Antony Rowe, Chippenham, Wiltshire.

First Impression  2010

# Dictionary of DNA and Genome Technology

Second Edition

For Mobby

# Preface

This edition has been written to accomodate the recent spate of innovations and developments in DNA technology. It covers a wide range of new methods and adaptations, as well as terminology, and reflects current thinking in both *in vitro* and *in vivo* studies. The opportunity was also taken to update and extend entries from the 1st edition. Every effort has been made to present up-to-date information – much of it newly published in mainstream journals within the last 6–12 months.

The dictionary is designed for use throughout the biomedical sciences, particularly in areas such as:

- biotechnology
- diagnosis (hereditary and infectious diseases)
- drug development
- epidemiology
- forensic science
- gene therapy
- genetically modified (GM) foods
- genomics
- industrial enzymes
- microbiology
- molecular biology
- oncology
- systems biology
- taxonomy
- vaccine development

References to published papers and reviews are cited throughout the dictionary. Some of these are the source(s) of information on which particular entries are based; they can provide additional details (e.g. protocols), and they also permit the reader to make his or her own assessment of a given source. Other references are included in order to indicate further information which is relevant to a given entry.

Some pervasive topics – for example, PCR, gene fusion, forensics, phages, overexpression, retroviruses, typing, microarrays, stem cells – are treated more extensively. These essay-style entries are intended to offer the newcomer a broad working knowledge of the area. Such entries may also be of use to the over-specialized researcher.

Commercial systems and materials are widely used in DNA technology, and a range of entries describe these products. This kind of entry may be useful, for example, when a paper offers no information about

a given product other than "used according to the manufacturer's instructions" – leaving unanswered questions on the product, the protocol or the principle. Accordingly, these entries give brief overviews of products – and their uses – and in many cases they also cite relevant papers describing work in which a given product has been used. Where the names of products and systems are known to be trademarks this has been indicated.

*Notes for the user* on the following pages will facilitate use of the dictionary. The attention of the reader is drawn, in particular, to the item on alphabetization.

**Paul Singleton**
Clannaborough (UK), September 1st 2009

# Notes for the user

## Alphabetization

The headwords in any dictionary can be listed in either of two distinct ways:

1. According to the way in which the terms are actually written. This approach, which is called the 'word-by-word' approach, is used in this dictionary.
2. According to the way in which the terms *would* appear if all the spaces and hyphens etc. were deleted. This is the 'letter-by-letter' approach.

The ***order*** in which headwords are listed in any dictionary depends on the particular approach used; for example:

***Approach 1***

A site
A-tract
AAS
abacavir
abasic site
ABC excinuclease

*bla* gene
black–white screening
blue–white screening
Bluescript®
branch migration
branched DNA assay

col plasmid
colchicine
*cos* site
cosmid
Cre–*loxP* system
CREB

***Approach 2***

AAS
abacavir
abasic site
ABC excinuclease
A site
A-tract

black–white screening
*bla* gene
Bluescript®
blue–white screening
branched DNA assay
branch migration

colchicine
col plasmid
cosmid
*cos* site
CREB
Cre–*loxP* system

Note that a one-letter descriptor (such as the 'A' in 'A site') is treated as a word for the purposes of alphabetization in approach 1.

In practice, neither approach is foolproof. For example, in the first approach, the order in which a given term is listed may depend on whether or not a hyphen is regarded as a necessary part of the term. Thus, on deleting a hyphen – and closing-up the intervening space – the characters on either side of the hyphen become contiguous; in this case, the character which followed the hyphen is important as a primary determinant of alphabetical order. In the second approach, strict adherence to the basic 'letter-by-letter' rule would lead – for example – to the following order:

| | |
|---|---|
| factor I | (1) |
| factor II | (2) |
| factor III | (3) |
| factor IV | (4) |
| factor IX | (9) |
| factor V | (5) |
| factor VI | (6) |
| factor VII | (7) |
| factor VIII | (8) |
| factor X | (10) |

To some, this order may seem reasonable, even preferable. To others, it runs counter to common sense: Roman numerals are generally seen (and used) as the equivalent of numbers – and ought therefore to be arranged in numerical order.

When a Greek letter is a significant component of an entry heading – as e.g. in λ phage – it is treated as a word and is listed in the relevant alphabetical position indicated by the English name (i.e. alpha, lambda etc.). However, β-galactosidase and β-lactamases are listed under G and L, respectively; this rule applies also to entry headings starting with letters such as L-, *p*-, *N*-, *O*- etc. which precede the names of certain chemicals. In some cases a headword is given in *both* possible locations, with suitable cross-referencing; this has been done simply in order to assist readers.

## Cross references

Words in SMALL CAPITALS refer the reader to entries elsewhere in the dictionary. Such cross references are included e.g. to extend the reader's knowledge into related fields or topics. Cross references may be particularly useful for directing the reader to allied, or parallel, subjects whose relationship to the entry being read may not be immediately obvious.

In some cases a complete understanding of a given entry, or a full appreciation of its context, depends on information contained in other entries – which are indicated by cross reference(s). Dictionaries are often arranged in this way because it avoids the need to repeat information. If it is especially important to follow-up a cross reference, then the cross reference is followed by '(q.v.)'. In other cases, in which the purpose of a cross reference is simply to link one topic with another, the cross reference may be preceded by 'See also. . ..' or 'cf.'.

## External references

References to papers, articles or reviews in journals are given in square brackets. The names of journals are abbreviated to save space. The abbreviated journal name is followed by the year of publication, the volume number (and frequently the issue number), and page number(s). This information is sufficient to enable the reader to obtain any given reference. (When a paper has been cited as an advance publication it may be referred to by its digital object identifier (doi) number; this number allows the reference to be followed-up.)

## Commercial products

Many of the commercial products are listed under their trade names. In general, these products are widely used in studies on all aspects of DNA-based technology and are cited in many research papers. It should be noted that the inclusion of any given product in the dictionary is not based on an evaluation of that product, and implies no comparison of that product with any similar product(s) marketed by other companies. It is obviously not possible to include every product currently on the market. Importantly, any details of a product given in the dictionary are those details which are to hand at the time of writing; companies are continually modifying and updating their products, so that the reader should refer to the manufacturer's literature for details of any modifications.

# Ready reference

## The Greek alphabet

| | | | | | |
|---|---|---|---|---|---|
| A | α | alpha | N | ν | nu |
| B | β | beta | Ξ | ξ | xi |
| Γ | γ | gamma | O | o | omicron |
| Δ | δ | delta | Π | π | pi |
| E | ϵ | epsilon | P | ρ | rho |
| Z | ζ | zeta | Σ | σ | sigma |
| H | η | eta | T | τ | tau |
| Θ | θ | theta | Y | υ | upsilon |
| I | ι | iota | Φ | φ | phi |
| K | κ | kappa | X | χ | chi |
| Λ | λ | lambda | Ψ | ψ | psi |
| M | μ | mu | Ω | ω | omega |

## Amino acids

| | | | | | |
|---|---|---|---|---|---|
| alanine | Ala | A | A | Ala | alanine |
| arginine | Arg | R | C | Cys | cysteine |
| asparagine | Asn | N | D | Asp | aspartic acid |
| aspartic acid | Asp | D | E | Glu | glutamic acid |
| cysteine | Cys | C | F | Phe | phenylalanine |
| glutamic acid | Glu | E | G | Gly | glycine |
| glutamine | Gln | Q | H | His | histidine |
| glycine | Gly | G | I | Ile | isoleucine |
| histidine | His | H | K | Lys | lysine |
| isoleucine | Ile | I | L | Leu | leucine |
| leucine | Leu | L | M | Met | methionine |
| lysine | Lys | K | N | Asn | asparagine |
| methionine | Met | M | P | Pro | proline |
| phenylalanine | Phe | F | Q | Gln | glutamine |
| proline | Pro | P | R | Arg | arginine |
| serine | Ser | S | S | Ser | serine |
| threonine | Thr | T | T | Thr | threonine |
| tryptophan | Trp | W | V | Val | valine |
| tyrosine | Tyr | Y | W | Trp | tryptophan |
| unknown | Xaa | X | X | Xaa | unknown |
| valine | Val | V | Y | Tyr | tyrosine |

## Prefixes used with SI (Système International) units

| (prefix, | value, | symbol) | (value, | prefix, | symbol) |
|---|---|---|---|---|---|
| | | **FRACTIONS** | | | |
| atto | $10^{-18}$ | a | $10^{-1}$ | deci | d |
| centi | $10^{-2}$ | c | $10^{-2}$ | centi | c |
| deci | $10^{-1}$ | d | $10^{-3}$ | milli | m |
| femto | $10^{-15}$ | f | $10^{-6}$ | micro | μ |
| micro | $10^{-6}$ | μ | $10^{-9}$ | nano | n |
| milli | $10^{-3}$ | m | $10^{-12}$ | pico | p |
| nano | $10^{-9}$ | n | $10^{-15}$ | femto | f |
| pico | $10^{-12}$ | p | $10^{-18}$ | atto | a |
| yocto | $10^{-24}$ | y | $10^{-21}$ | zepto | z |
| zepto | $10^{-21}$ | z | $10^{-24}$ | yocto | y |
| | | **MULTIPLES** | | | |
| deca | 10 | da | 10 | deca | da |
| exa | $10^{18}$ | E | $10^{2}$ | hecto | h |
| giga | $10^{9}$ | G | $10^{3}$ | kilo | k |
| hecto | $10^{2}$ | h | $10^{6}$ | mega | M |
| kilo | $10^{3}$ | k | $10^{9}$ | giga | G |
| mega | $10^{6}$ | M | $10^{12}$ | tera | T |
| peta | $10^{15}$ | P | $10^{15}$ | peta | P |
| tera | $10^{12}$ | T | $10^{18}$ | exa | E |
| yotta | $10^{24}$ | Y | $10^{21}$ | zetta | Z |
| zetta | $10^{21}$ | Z | $10^{24}$ | yotta | Y |

## Micro-measurements

1 Å (Ångström unit) = $10^{-1}$ nm = $10^{-4}$ μm = $10^{-10}$ m
1 nm (nanometer) = $10^{-3}$ μm = $10^{-6}$ mm = $10^{-9}$ m
1 μm (micrometer, formerly micron) = $10^{-3}$ mm = $10^{-6}$ m
1 mm = $10^{-1}$ cm = $10^{-3}$ m = $10^{-6}$ km

# Contents

# A

**A** (1) Adenine (as a base, or the corresponding nucleoside or nucleotide).
(2) L-Alanine (alternative to Ala).

**Å** Ångström unit, $10^{-10}$ m; a unit of length used e.g. to indicate intermolecular distances.

**$A_{260}$** See the entry ULTRAVIOLET ABSORBANCE.

**A box** The adenine riboswitch aptamer (see RIBOSWITCH).

**A-DNA** One of the conformations adopted by dsDNA: a right-handed helix with ~11 base-pairs per turn.
(Note that aDNA is used to refer to ANCIENT DNA.)
(cf. B-DNA and Z-DNA.)

**A family** (of DNA polymerases) A group of DNA-DEPENDENT DNA POLYMERASES that include prokaryotic, eukaryotic and viral enzymes. Members of the A family include some phage polymerases (although not those from phages φ29 or T4) and the *Escherichia coli* pol I (involved e.g. in the maturation of Okazaki fragments and in BASE EXCISION REPAIR).
Also included in this family is POLQ (= pol θ; pol theta), an enzyme found in human and other eukaryotic cells. POLQ is able to carry out translesion synthesis of DNA. It may participate in base excision repair, a suggestion supported by the *in vitro* demonstration of its 5′-deoxyribose phosphate lyase activity, a role apparently involved in single-nucleotide base excision repair [Nucleic Acids Res (2009) 37(6):1868–1877].
(See also B FAMILY, X FAMILY and Y FAMILY.)

**A site** (of a ribosome) The aminoacyl or 'acceptor' site at which tRNA molecules carrying the second and subsequent amino acids bind during translation. (cf. P SITE.)

**A-tract** In genomic DNA: a nucleotide motif that is reported to be associated with regions of the most pronounced curvature of the molecule; an A-tract is a poly(A) (i.e. poly-adenosine) sequence. In the genome of *Escherichia coli*, A-tracts were reported to be distributed 'quasi-regularly', in both coding and non-coding sequences; the A-tracts occur in clusters ~100 bp long, with consecutive A-tracts exhibiting a periodicity of 10 to 12 bp. It was suggested that the clusters of A-tracts may constitute a form of 'structural code' for DNA compaction in the NUCLEOID [Nucleic Acids Res (2005) 33:3907–3918].
Studies on the mechanics and dynamics of DNA suggested a rationale, incorporating A-tracts, for the stable bending of DNA [Nucleic Acids Res (2008) 3:2268–2283].
Studies on *eukaryotic* genomes have reported that A-tracts are absent specifically in those coding sequences (exons) that correspond to the locations of nucleosomes. It was concluded that the pattern of absence/presence of A-tracts in the genome constitutes a code for the presence/absence – respectively – of nucleosome locations. [The coexistence of the nucleosome positioning code with the GENETIC CODE on eukaryotic genomes: Nucleic Acids Res (2009) doi: 10.1093/nar/gkp689.]
(cf. CLASS A FLEXIBLE PATTERNS.)

**AAA ATPases** 'ATPases associated with diverse cellular activities' – ATPases which are found in various locations, such as proteasomes and peroxisomes. They have been categorized as AAA+ PROTEINS.

**AAA+ proteins** A family of NTPases whose members include proteins with diverse functions; AAA ATPASES are examples of this group.
[Review of AAA+ proteins: Genome Biol (2008) 9(4): 216.]

**AAAVs** Avian adeno-associated viruses (see the entry AAVS).

**AAS** Aminoalkylsilane (3-aminopropyltriethoxysilane; APES): a reagent used e.g. to bind tissue sections to glass (for *in situ* hybridization etc.).
[Uses (e.g.): Am J Pathol (2006) 169(1):258–267; Nucleic Acids Res (2008) 36(16):5335–5349.]

***aat* gene** In *Escherichia coli*: a gene encoding the enzyme that catalyzes addition of a leucine or phenylalanine residue to the N-terminal of proteins that are synthesized with either an N-terminal arginine or a lysine residue; such addition facilitates degradation of the protein.
(See also N-END RULE.)

**AatII** A RESTRICTION ENDONUCLEASE from *Acetobacter aceti*. Recognition sequence/cutting site: GACGT↓C.

**AAUAAA** In a pre-mRNA: a polyadenylation signal upstream of the site at which the molecule is cut and polyadenylated; the polyadenylation sequence is similar in various organisms, although there are variations.
Other *cis*-acting elements may have roles in regulating the polyadenylation of human mRNAs – including upstream U-rich sequences similar to those which have been identified in yeast and plants.
As well as acting as a polyadenylation signal, this sequence was reported to affect the *rate* of transcription [RNA (2006) 12(8):1534–1544].

**AAV** Adeno-associated virus: see the entry AAVS.

**AAV Helper-Free System** A commercial gene-delivery system (Stratagene, La Jolla CA) in which the genes in two plasmids provide functions necessary for production of infective AAV virions (see AAVS) without the need for a helper virus; these virions are used to deliver genes to target cells within which viral DNA – containing the gene of interest – integrates in the host cell's DNA.
Essentially, the gene/fragment of interest is first cloned in a plasmid cloning vector in which the insert is bracketed by a pair of inverted terminal repeats (ITRs) which are necessary for subsequent viral packaging. This plasmid is then used to transfect PACKAGING CELLS – which are *co-transfected* with two other plasmids: (i) a plasmid containing the genes that encode viral capsid and replication functions, (ii) a plasmid containing genes that encode the lytic phase of AAV. The resulting infective (*but still replication-deficient*) virions that are produced in the packaging cells can then be used to infect the required target cells (in which the gene of interest can be expressed).
This GENE-DELIVERY SYSTEM has been used e.g. to express

siRNAs [Mitochondrion (2007) 7(4):253–259]; to deliver an anti-angiogenic gene (for investigating age-related macular degeneration) [Mol Vision (2008) 14:471–480]; and to study some features of food/energy metabolism [J Neurosci (2009) 29(1):179–190].

(See also: VIRAPORT RETROVIRAL GENE EXPRESSION SYSTEM and VIRAPOWER LENTIVIRAL EXPRESSION SYSTEM.)

**AAVs** Adeno-associated viruses (also known as: adeno-satellite viruses): defective viruses that are able to replicate only when certain functions are provided by a co-infecting *helper virus* (adenovirus or herpesvirus) – or, in certain *in vitro* systems, when these functions are provided by plasmid-borne genes (as e.g. in the AAV HELPER-FREE SYSTEM).

Functions provided by adenovirus type 5 (for AAV type 5) include both positive and negative effects. For example, the E4Orf6 function (involved in replication of AAV5 genomic DNA) – together with E1b – degrades AAV5 capsid proteins and Rep52 [J Virol (2007) 81(5):2205–2212]. The functions provided by herpes simplex virus type 1 (for the early stages of AAV replication) were reported to involve nine proteins from the helper virus [PLoS Pathog (2009) 5(3):e1000340].

The AAVs are parvoviruses in which the genome is linear ssDNA. Positive and negative strands of the viral DNA are encapsidated in separate virions.

The AAVs infect a wide range of vertebrates. Initial stages of infection, including internalization of DNA, occur without a helper virus. [Cloning an *avian* AAV (an AAAV) and the generation of recombinant AAAVs: J Virol (2003) 77:6799–6810.]

AAVs are used, for example, in GENE THERAPY. Efforts are being made to increase the efficacy of AAV vectors in gene therapy by designing the CAPSID on the basis of e.g. information obtained from studies on the naturally occurring capsid variants of AAVs in mammals [see: Gene Therapy (2009) 16: 311–319]. (See also KU70.)

An inducible and highly efficient system was reported for the production of recombinant AAV vectors in insect (Sf9) cells [Proc Natl Acad Sci USA (2009) 106(13):5059–5064].

AAV vectors, encoding genes of the α and the β subunits of hexosaminidase, were inoculated, *intracranially*, into mice in order to assess the potential of gene therapy for treatment of the human GM2 gangliosidoses such as Tay–Sachs disease and Sandhoff disease [Proc Natl Acad Sci USA (2006) 103 (27):10373–10378]. A simpler method for delivering genes to brain cells was reported later (see below).

AAV9 has been used, in mice, for gene delivery to cells of the central nervous system (brain and spinal cord) *by intravenous injection*. It was thought that this approach may allow the development of gene therapy for e.g. some human neurodegenerative diseases [Nature Biotechnol (2008) 27:59–65].

AAV vectors were also used for the genetic manipulation of cultured neurons [Brain Res (2008) 1190:15–22].

It was reported earlier that, in human cells, AAV DNA (in the absence of helper virus) integrates in the genome with an apparent preference for CPG ISLANDS. More recently, AAVs have been reported to integrate, site-specifically, into a locus on chromosome 19, and the occurrence of such integration is apparently influenced by the TRP-185 protein [J Virol (2007) 81(4):1990–2001]. Palindromes of length greater than about 40 bp are reported to be significant targets for the integration of recombinant AAV vectors [J Virol (2007) 81(20):11290–11303].

The site of insertion of AAVs within chromosome 19 was reported to contain a 347-bp sequence capable of enhancing the promoter and transcriptional functions of AAV vectors in liver cells; inclusion of this small fragment in AAV vectors may thus facilitate their use for the delivery and expression of transgenes in liver cells [Gene Therapy (2009) 16:43–51].

**AB1380** A strain of the yeast *Saccharomyces cerevisiae* (see the entry SACCHAROMYCES for some details).

(See also YEAST ARTIFICIAL CHROMOSOME.)

**abacavir** A NUCLEOSIDE REVERSE TRANSCRIPTASE INHIBITOR used e.g in antiretroviral therapy; CSF–plasma ratios indicate that it may reach therapeutic levels in the cerebrospinal fluid (CSF).

A trial that compared abacavir with nevirapine (as part of a combined therapy) reported that abacavir tended to produce a lower rate of serious adverse effects, suggesting a wider use of this drug in resource-limited settings [Trop Med Int Health (2008) 13(1):6–16].

**abasic site** *Syn.* AP SITE.

**abasic-site mimic** See the entry RPA.

**ABC excinuclease** See UVRABC-MEDIATED REPAIR.

**Abelson murine leukemia virus** See the entry ABL.

**aberrant RNA** (aRNA) See the entry ARNA (sense 2).

***abl*** (*ABL*) An ONCOGENE first identified in the Abelson murine leukemia virus. The v-*abl* product has TYROSINE KINASE activity. The human homolog of v-*abl*, c-*abl*, is usually present on chromosome 9; however, in the majority of patients with CHRONIC MYELOGENOUS LEUKEMIA it has been translocated to chromosome 22, forming a chimeric gene, known as *bcr-abl*, that encodes a tumor-specific tyrosine kinase (designated P210). Chromosome 22 containing the chimeric *bcr-abl* gene is called the Philadelphia chromosome (also called $Ph^1$).

Subcellular localization of c-Abl protein at an early stage in myogenic differentiation was reported to be influenced by its acetylation [EMBO Rep (2006) 7(7):727–733].

**abortive transduction** TRANSDUCTION in which the transduced DNA persists in a recipient cell as a stable, extrachromosomal but non-replicating molecule; when the recipient divides only one daughter cell receives the DNA fragment.

**absorbance** (ultraviolet) See ULTRAVIOLET ABSORBANCE.

**abzyme** *Syn.* CATALYTIC ANTIBODY.

**Abzyme®** A reagent kit (Abbott Laboratories) used for detecting antibodies in the context of hepatitis B.

**acceptor splice site** (acceptor splice junction) In a pre-mRNA: the splice site (consensus AG) at the 3′ end of an intron.

(cf. DONOR SPLICE SITE.)

**accession number** A number which refers to a unique database entry for a given sequence or gene. Some examples include:

(i) GenBank® accession number X17012, referring to data on the gene for rat insulin-like growth factor II (IGF II); (ii) GenBank® accession number AY024353, referring to data on the *ftsZ* gene of the bacterium *Sodalis glossinidius;* (iii) GenBank® accession number AM160602, referring to data on mRNA of the gene for cinnamyl alcohol dehydrogenase in a species of oak (*Quercus ilex*).

(See also ANNOTATION.)

**AccuPrime™ GC-rich DNA polymerase** A DNA polymerase (Invitrogen, Carlsbad CA) optimized for DNA synthesis on 'difficult-to-amplify' templates, including those with a GC content >65%. Targets up to 5 kb may be amplified with this polymerase.

[Uses (e.g.): J Bacteriol (2008) 190(24):8096–8105; J Exp Clin Cancer Res (2008) 27:54; FEMS Microbiol Lett (2009) 294(1)32–36.]

**AccuProbe®** A family of PROBES (Gen-Probe, San Diego CA) used for identifying certain medically important bacteria by detecting specific sequences of nucleotides from lysed cells. The method involves a *hybridization protection assay.* In this assay, an added reagent cleaves the acridinium ester label on all *unbound* probes. Labels on the *bound* probes (which are protected from cleavage by virtue of their position in the probe–target duplex) react with a second reagent, producing a chemiluminescent (light) signal. The light produced by this reaction is measured in RLUs (i.e. relative light units). The threshold value (in RLUs) for a positive result must be carefully examined [see for example: J Clin Microbiol (2005) 43:3474–3478].

[Use for *Staphylococcus aureus*: J Clin Microbiol (2008) 46(6):1989–1995. Use for *Streptococcus pneumoniae* (as a reference): J Clin Microbiol (2008) 46(7):2184–2188. Use for *Mycobacterium avium*: J Clin Microbiol (2008) 46 (8):2790–2793. Use for identifying *Mycobacterium* spp: Emerg Infect Dis (2009) 15(1):53–55, and Emerg Infect Dis (2009) 15(2):242–249.]

(See also PACE 2C and TMA.)

**acetosyringone** A phenolic substance which promotes activity of the *vir* operon in species of the plant-pathogenic bacterium *Agrobacterium* (see CROWN GALL).

(See also AGROINFILTRATION.)

Acetosyringone has been used e.g. for studies on terpenoid metabolism in the tomato plant [Plant Physiol (2009) 149(1): 499–514], and studies on the transformation of wheat [Plant Cell Rep (2009) 28(6):903–913].

*Agrobacterium* can also transfer T-DNA to other types of cell, including e.g. human and fungal cells; acetosyringone was used to promote transfer of T-DNA from *Agrobacterium* to the fungus *Aspergillus fumigatus* for (random) insertional mutagenesis [PLoS ONE (2009) 4(1):e4224].

***N*-acetyl-L-cysteine** See MUCOLYTIC AGENT.

**acetylation** (of histones) HISTONE acetylation is regulated e.g. by the opposing effects of histone acetyltransferases (HATs) and histone deacetylases (see HDAC); the (de)acetylation of histones can affect CHROMATIN structure, and may therefore alter the accessibility of DNA for events such as transcription and repair.

The acetylation of histones can be studied/manipulated e.g. by using HDAC *inhibitors* (e.g. TRICHOSTATIN A).

[Genome-wide analysis of histone acetylation and its effect on gene expression in (the protozoan) *Entamoeba histolytica*: BMC Genomics (2007) 8:216.]

A general perception is that transcription of genes requires – as a pre-condition – an 'open' form of CHROMATIN (the so-called *euchromatin*) in the vicinity of the given genes; acetylation of vicinal histone(s) is usually regarded as an important factor associated with the presence of euchromatin. (In some types of chemotherapy – EPIGENETIC THERAPY – an inhibitor of HDACs is sometimes included in order to promote 'open' chromatin in the vicinity of specific gene(s) with the object of contributing to de-repression of the genes.) However, histone acetylation is only one factor that regulates gene expression; for example, it was reported that the drug-induced formation of 'open' chromatin (involving hyper-acetylation of vicinal histone(s)) was not, on its own, sufficient to de-repress lytic-cycle genes in the Epstein–Barr virus [J Virol (2008) 82(10): 4706–4719]. Nevertheless, the complexity of this issue may be indicated by a study in which histone acetylation – but not DNA demethylation – *was* found to be sufficient to break the latency of gamma-herpesvirus 68 in a mouse cell line [PLoS ONE (2009) 4 (2):e4556].

In a genomewide study of HDACs in *Schizosaccharomyces pombe* (a fission yeast), the patterns of histone acetylation, HDAC binding and nucleosome density were compared with gene expression profiles; it was found that different HDACs may have different roles in repression and activation of genes [EMBO J (2005) 24(16):2906–2918]. Following *damage* to DNA in *S. pombe*, the restoration of chromatin structure was reported to involve deacetylation of histone H3 by Hst4 (a putative HDAC) [Eukaryotic Cell (2008) 7:800–813], while recovery from DNA damage was reported to involve Mst1 (a histone acetyltransferase) [Genetics (2008) 179(2):757–771].

In (human) nucleosomes, the acetylation of certain lysine residues depends primarily on HATs, but the effect of these enzymes appears to be promoted by binding protein HMGN1 [EMBO J (2005) 24(17):3038–3048].

Acetylation of the histone *chaperone* NUCLEOPHOSMIN, as well as histone acetylation, apparently promotes transcription [Mol Cell Biol (2005) 25(17):7534–7545], while chaperone-stimulated, histone-acetylation-independent transcription has also been reported [Nucleic Acids Res (2007) 35:705–715].

HDACs in (human) development and physiology have been reviewed within the context of implications for disease and therapy [Nature Rev Genetics (2009) 10:32–42].

The c-Abl protein (see ABL) was reported to be a substrate for the p300 and other histone acetyltransferases.

***N*-acetylmuramidase** See LYSOZYME.

***N*-acetylneuraminic acid** (NANA) See NEURAMINIDASE.

**ACF** APOBEC-1 complementation factor: see RNA EDITING.

**Achilles' heel technique** A technique in which a RESTRICTION ENDONUCLEASE is targeted to one *particular* recognition site when multiple copies of that site are freely available. In one method, a triplex-forming oligonucleotide (see TRIPLEX DNA) is used to mask the required cleavage site. While this site is masked, the remaining sites are methylated in order to inhibit subsequent cleavage; the triplex is then removed and specific cleavage can be carried out.

(See also PROGRAMABLE ENDONUCLEASE.)

**aciclovir** Alternative spelling for ACYCLOVIR.

**acid-fast bacilli** Those bacilli (i.e. rod-shaped bacteria) which, when stained with the Ziehl–Neelsen (or similar) stain, resist decolorization with mineral acid or an acid–alcohol mixture. This kind of staining method is used for screening respiratory specimens, e.g. samples of sputum, and for examining other types of specimen, for *Mycobacterium tuberculosis* (an acid-fast species).

**Ac*M*NPV** *Autographa californica* NPV: see NUCLEAR POLYHEDROSIS VIRUSES.

**AcNPV** *Syn.* Ac*M*NPV – see entry NUCLEAR POLYHEDROSIS VIRUSES.

**acridines** Heterocyclic, fluorescent compounds which bind to dsDNA (primarily as an INTERCALATING AGENT) and also to single-stranded nucleic acids (and to the backbone chains of double-stranded nucleic acids). Acridines have antimicrobial activity and they are mutagenic; they are also used as stains for nucleic acids and can be used for CURING plasmids.

**acridinium ester label** (on probes) See ACCUPROBE.

**acrocentric** Refers to a CHROMOSOME in which the CENTROMERE is located close to one end.

**acrydite hybridization assay** An assay in which molecules of labeled ssDNA or ssRNA, passing through a polyacrylamide gel by electrophoresis, are captured (bound) by complementary oligonucleotides immobilized in a (central) 'capture zone' within the gel; all the molecules of nucleic acid that are *not* complementary to the capture oligos pass through the central capture zone and continue their migration to the end of the gel strip. The complementary oligos are synthesized with a 5′ terminal acrydite group which binds them to the polyacrylamide matrix so that they are immobilized in the gel. (Note that the central region of the gel strip is prepared separately.)

**acrylamide** A toxic, water-soluble agent ($CH_2$=CH—$CONH_2$) which can be polymerized to POLYACRYLAMIDE by catalysts such as *N*,*N*′-methylene-*bis*-acrylamide ('Bis') which promote cross-linking.

**actinomycin $C_1$** *Syn.* ACTINOMYCIN D.

**actinomycin D** An antibiotic (a substituted phenoxazone linked to two pentapeptide lactone rings) produced by some species of *Streptomyces;* it acts as an INTERCALATING AGENT, binding to DNA and inhibiting DNA-dependent RNA polymerase. The drug has low affinity for AT-rich promoter regions; hence, *initiation* of transcription from such promoters may be little affected by the antibiotic.

**activation domain** (AD) See YEAST TWO-HYBRID SYSTEM.

**activation-induced cytidine deaminase** (AID) An enzyme that occurs in germinal center B lymphocytes (B cells) and which is an absolute requirement for affinity maturation and class switching in normal development of antibodies.

The (autosomal recessive) form of HYPER-IGM SYNDROME has been linked to a deficiency of AID (see the table in entry GENETIC DISEASE).

AID also inhibits retrotransposition of L1 – suggesting that it has function(s) in addition to its role of creating antibody diversity [Nucleic Acids Res (2009) 37(6):1854–1867].

(See also CYTIDINE DEAMINASE AND RNA EDITING.)

**activation/regulation of genes** (*DNA technol.*) See e.g. entry CONDITIONAL GENE ACTIVATION/REGULATION.

**activity-based probe** A type of probe used for the real-time study of APOPTOSIS (q.v.).

**acyclonucleotide** Any analog of a deoxyribonucleotide or ribonucleotide in which a non-cyclic moiety carries the base. One example is a monomer of glycerol nucleic acid (see the entry GNA). Polymerization of certain acyclonucleotides on a DNA template has been achieved with THERMINATOR DNA POLYMERASE.

**acyclovir** (alternative spelling: aciclovir) 9-(2-hydroxyethoxymethyl)guanine: an antiviral agent which is active against a number of herpesviruses, including herpes simplex. In cells, acyclovir is phosphorylated to the monophosphate by (viral) thymidine kinase; subsequently it is converted to the (active) triphosphate form via host-encoded enzymes. The active drug inhibits *viral* DNA polymerase; the host cell's polymerase is much less sensitive.

In cells which are not virally infected, acyclovir appears not to be significantly phosphorylated.

Acyclovir has been used topically and systemically.

***N*-acyl-homocysteine thiolactone** See QUORUM SENSING.

***N*-acyl-L-homoserine lactone** (AHL) See QUORUM SENSING.

**acylneuraminyl hydrolase** *Syn.* NEURAMINIDASE.

**AD primer** (arbitrary degenerate primer) See TAIL-PCR.

**Ada protein** (in *Escherichia coli*) See DNA REPAIR.

**adaptamer** See ORFMER SETS.

**adaptive response** (to alkylating agents) See DNA REPAIR.

**adaptor** A short, synthetic, double-stranded fragment of DNA which is similar, in principle, to a LINKER but which generally offers more flexibility. Thus, for example, the two ends of a given adaptor may consist of dissimilar STICKY ENDS – one end able to bind to a (complementary) sticky end on a DNA fragment and the other able to bind to a different sticky end on a vector molecule, facilitating the integration of fragments and vectors which were cut by different restriction enzymes. An adaptor may also include one or more internal restriction sites, offering the chance to select alternative sticky ends at a later stage in the work, and there may also be primer-binding sites and/or a promoter sequence etc.

(See also NOTI.)

**ADAR1** A dsRNA ADENOSINE DEAMINASE which is involved e.g. in RNA EDITING.

(See also Zα in the entry Z-DNA.)

**AdEasy™ XL adenoviral vector system** A GENE-DELIVERY