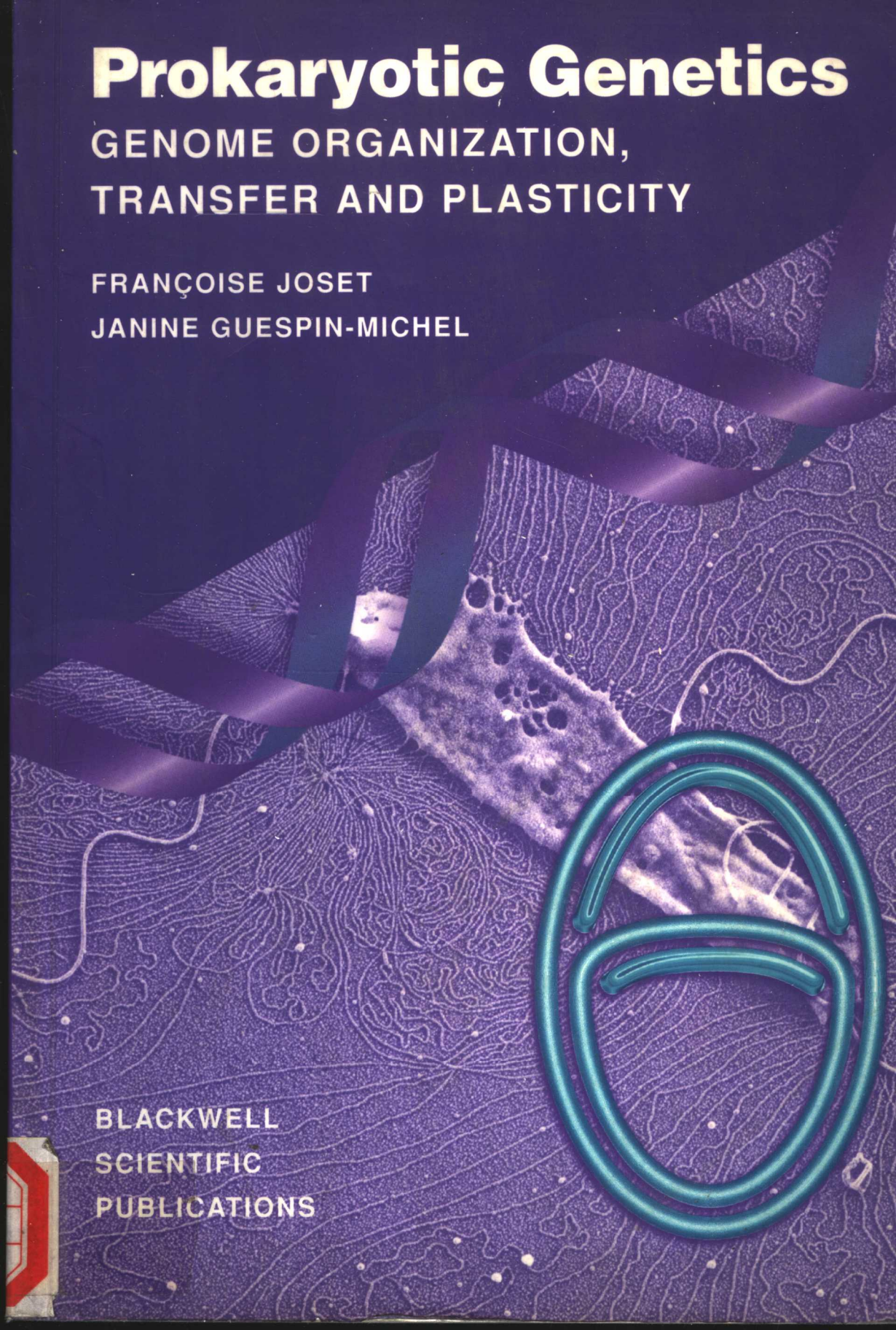
Prokaryotic Genetics

GENOME ORGANIZATION, TRANSFER AND PLASTICITY

FRANÇOISE JOSET
JANINE GUESPIN-MICHEL



BLACKWELL
SCIENTIFIC
PUBLICATIONS

STUDIES IN MICROBIOLOGY

Prokaryotic Genetics

GENOME ORGANIZATION, TRANSFER AND PLASTICITY

FRANÇOISE JOSET
Professor, CNRS and Université Aix-Marseille II

JANINE GUESPIN-MICHEL
Professor, CNRS and Université de Rouen

WITH THE COLLABORATION OF
LESLIE O. BUTLER
Professor, St George's Hospital Medical School, London

OXFORD
Blackwell Scientific Publications
LONDON EDINBURGH BOSTON
MELBOURNE PARIS BERLIN VIENNA

Preface

Prokaryotic genetics? While the meaning of 'genetics' is nowadays clear, the word 'prokaryote' may raise some questions. It is used here in place of the better-known word 'bacterium'. Its usage emphasizes the central fact that the component organisms characteristically lack a distinct nucleus and nuclear membrane, in sharp contrast with all other living cells. Prokaryotic cell division does not involve the degree of structural elaboration associated with eukaryotic cell division. The essential difference between these two groups is not that of unicellular versus multicellular structures but lies in the organization of the cells and genomes, and consequently of the genetic exchange processes involved.

Although viruses may at one level be considered as a borderline between living organisms and macromolecular complexes, at the genetic level they share a fundamental characteristic of all living organisms in having an autonomously reproducing genome. Furthermore, the understanding of their genetics and reproduction processes has played an important role in the unravelling of genetics, especially that of bacteria. For this reason, the viruses which infect bacteria, the bacteriophages, lie within the scope of this book, to the extent that they have contributed to the study of bacterial genetics.

Bacteria are usually assumed to have been the earliest living organisms on earth. Although they had already been observed in the seventeenth century, their importance in everyday life as pathogenic agents or as activators in chemical transformations, in particular in industrial processes, was recognized only during the nineteenth century, after Pasteur's work. The existence of bacterial viruses, although postulated for some time, was established almost a century later in the works by Twort and d'Hérelle. The history of work with prokaryotes has been one of identifying and separating an ever-increasing number of individual species, the classification of which has been based on our evolving knowledge of morphology, physiology, biochemistry and latterly molecular sequences, the more recent stages of this process having been increasingly underpinned by genetic knowledge. The first report of a study on bacterial genetics as such was in 1942, with the first recognition that DNA was the agent of inheritance. Successive

landmarks in bacterial genetics which opened the way to molecular genetics of both prokaryotes and eukaryotes, were revealed over a period of about 20 years, during the 1940s to 1960s. These included: the establishment of the role of DNA as carrier of hereditary information, the deciphering of the chemical structure of DNA and the elucidation of the molecular organization of the DNA double helix and of the genetic code. Since then, bacterial genetics has developed at such an accelerating rate that it has, in fact, brought us most of our present knowledge of molecular genetics. The accompanying creation of *in vivo* and *in vitro* genetic tools has proved powerful not only for the improvement of our understanding of fundamental biological processes, but also for the utilization of genetic capabilities in biotechnology.

The diversity of prokaryotic species present in today's world is a measure of the varied genetic potentialities which yield stable, viable, individuals. These genetic characters are organized in coded structures, or genetic units, to form the bacterial chromosomes, which must, in a manner not yet understood, have been the precursors of those of eukaryotes. The complex fine structure of the bacterial chromosomes had long been assumed to be a token of their stability until recent discoveries indicated a degree of flexibility, a feature which may have had an important evolutionary role.

This text on prokaryotic genetics adopts the following sequence. Part 1 describes the genetic features by which individual organisms are determined and identified (i.e. the characteristics of the elements forming their genetic material), and discusses the perpetuation of the strains and species thus determined (i.e. the transmission and protection of their genetic content). The second part discusses the available means or opportunities of the strains or species to modify their genetic information, while maintaining limits compatible with their survival. The applications of these genetic capabilities as tools for further understanding of bacterial metabolic processes and for molecular genetic technology are discussed in Part 3.

Most of our knowledge in prokaryotic genetics has been acquired from only a small number of bacterial species and, even more limiting, from only particular strains of these species. The ease of manipulation of these 'model' strains in the laboratory, especially *Escherichia coli* K12, is responsible for this situation. Interest in the genetics of a greater range of species has arisen mostly since microbiology has revealed the large variety of resources existing in these organisms and the possibility of their exploitation by molecular genetics. Such techniques have allowed the jump from descriptive accounts of whole cells to precise molecular analysis of particular genetic functions, often with much basic genetic information remaining unknown.

While it is obviously impossible to describe the genetics of 'non-model' prokaryotes without reference to 'model' bacteria, we will attempt to emphasize the similarities and diversities of genetic processes in a range of species. Such comparisons, while illustrating the variety of biological processes, should encourage readers to apply existing knowledge to a greater range of organisms.

The first book on bacterial genetics was written by William Hayes in 1964 (*The Genetics of Bacteria and their Viruses*, published by Blackwell Scientific Publications) and was unique in presenting the totality of available knowledge in this field. It was a 'Bible' for both students and researchers thanks to its clarity, enthusiasm and thorough presentation of the existing knowledge, and it had a crucial role in converting microbiologists into geneticists. The difficulty at that time, remarkably overcome by Hayes, probably lay in the lack of general models linking together diverse information. The difficulty at the present time, in contrast, is that information and models are now so abundant that difficult choices must be made as to what to present.

This book is intended for readers with some knowledge of biochemistry, general genetics and cellular physiology and metabolism. It is aimed at a main readership of undergraduate students in the biological sciences, but it should also be of value to technicians and research workers in biology or biotechnology.

The authors would like to offer particular thanks to Professor Leslie Butler for his kindness in reading the entire manuscript and for helping with the index. We would also like to express our appreciation to Dr Pam McAthey of the Polytechnic of North London for contributing Chapter 8 on DNA repair, to acknowledge the helpful discussions with many colleagues and to thank those who have provided or given permission for the reproduction of figures.

F.J., J.G.-M.

Contents

PART 1
STRUCTURE OF THE GENETIC MATERIAL

1: The Prokaryotes: a Brief Introduction

1 Identity and structure: a review

Prokaryotes (a synonym of Bacteria) are the simplest autonomous organisms known, i.e. organisms which possess all the functions (and hence the genetic information) allowing them to reproduce, although some need various forms of association, such as syntrophy, symbiosis and parasitism, either with other prokaryotes or with eukaryotes, to do so permanently or in certain conditions. They are also the smallest cells described, although their sizes vary widely, the largest ones reaching dimensions similar to those of some eukaryotic protists. Their morphological types are limited, but various external structures (such as flagella, pili, fimbriae, sheath, capsule, etc.) with different functions are often present. Most prokaryotes exist as unicellular organisms, although multicellular associations of cells within the same strain, endowed with particular physiological functions, are encountered; among the more complex, some form truly multicellular networks of branching hyphae or pluricellular fruiting bodies.

All prokaryotes share a number of common traits, the most specific of which is their lack of internal cellular organization into physically separated and functionally specialized compartments as found in eukaryotes, which are thought to have developed later. Present trends based on ribosomal RNA sequence comparisons divide living beings into three domains or kingdoms, the Bacteria, the Archaea and the Eukarya. The relation between these three domains is still debated, although a recent proposal by Woese (1987, 1990) accepts that a common ancestor, a progenote, led first to two branches, the Bacteria on one side and on the other a second branch which later divided into the Archaea and the Eukarya (Fig. 1.1). Bacteria (currently referred to as Eubacteria) are composed of Gram-positive and Gram-negative families, reflecting the presence of either a thick peptidoglycan cell wall (the former) or a thin peptidoglycan cell wall surrounded by an outer membrane (the latter), external to the cytoplasmic membrane and protecting the cell. The cell walls of Archaea show no structural uniformity and have completely different compositions from those of Eubacteria. Many other features, particularly at the molecular level, distinguish organisms from these two domains.

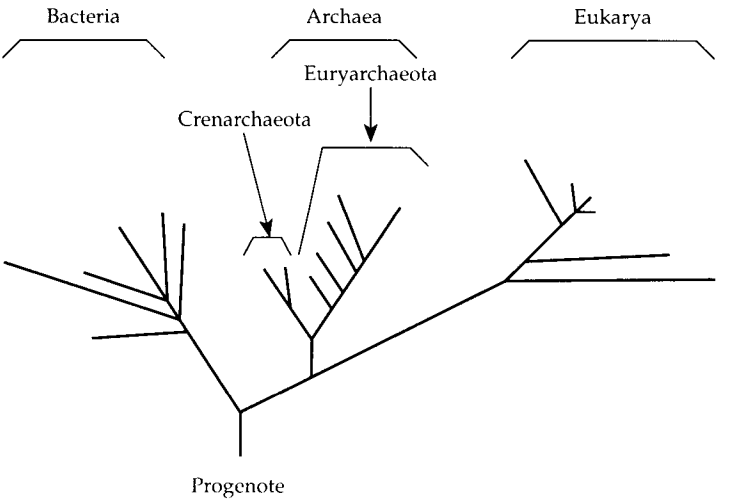


Fig. 1.1 Proposed phylogenetic tree of living organisms. This tree, showing three domains as presently known, is based particularly on ribosomal RNA comparisons. A putative original Progenote cell would have evolved into two branches, one generating the Bacteria and the other later dividing into two other domains, the Archaea and the Eukarya. The lengths of the different branches are proportional to the 'evolutionary distance' separating the various groups indicated, as known today. Each domain, or kingdom, subdivides into several 'groups'.

One feature important for the geneticist is the particular state of the genetic material of prokaryotes (Section 5). As might be expected, they contain less genetic information than even the least elaborate eukaryotes.

Although they have a fairly minimal structural complexity, prokaryotes display an amazing variety of adaptative capacities allowing them to populate a very large range of environmental conditions. These capacities are reflected both in the variety of 'nutritional' groups that can be defined, and in the intrinsic capacities of a given species or strain to adapt to varying environments. These features, together with morphological and other biochemical characteristics, have been the initial criteria used to define taxonomical groups. Accumulation of molecular data has led to modifications or confirmation of the original classifications. The nutritional criteria have, however, proved in most instances convenient and fruitful.

These extremely varied metabolic capacities, the frequent synthesis of industrially interesting secondary metabolites (Section 4) and the ease of growing large populations (Section 3) have promoted the use of bacteria as biological tools for applied purposes. This field of biotechnology has developed rapidly in the past decade or so, mostly

because of the parallel increase of our knowledge in bacterial genetics. Molecular genetics and genetic engineering make use of most of the genetic processes described (Part 3).

2 Nutrition and nutritional modes

The life of a unicellular organism such as most prokaryotes, apart from possible differentiation processes, consists mainly in nutritional reactions leading to cell doubling and division, the result of the co-ordinated functioning of energetic pathways (transformation of external energy sources into energy-rich metabolites such as ATP) and assimilatory pathways (the making up of cellular material). These biosynthetic pathways, in general very similar among prokaryotes and eukaryotes, will not be described here. The complex molecular composition of the bacterium *E. coli* (Table 1.1) reflects the intense metabolic activity that must take place in such cells, considering that in this particular case generation times range from around 20 min to 2 h.

Bacteria can fulfil their needs for energy and for the basic constituent carbon through the use of several processes, characterized by the nature of the energy and carbon sources and that of the electron source and acceptor.

Table 1.1 Average molecular composition of an *E. coli* cell.

Constituents	Number of copies/cell	Proportion in % of dry weight
Cell wall and outer layers including exopolysaccharides	1	≃10
Membrane including lipids	1	10
DNA		
chromosome	1 to 3–4	2 to 6–8
plasmids	0 to 8–n	
RNA		
mRNA	>3000	2
tRNA	>150000	3
rRNA	20000	21
Proteins (soluble and ribosomal)	$>10^6$	51
Small organic molecules	$6–7 \times 10^6$	1
Inorganic molecules	?	Negligible
Reserves (carbonic, nitrogenous, etc.)		Dependent on physiological state

E. coli is a rod-shaped bacterium, approximately 1.2 μ long × 0.5 μ diameter. This represents a volume of approximately 2.5×10^{-7} cm^3, or a weight of 2.5×10^{-7} g/cell (supposing an average density of 1), of which 70% is water, the rest comprising the constituents listed in the table.

Table 1.2 Prokaryotes grouped according to their energy source.

Class	Energy source	Electron		Energetic pathway
		Donor	Acceptor	
Chemoorganotrophs	Reduced carbon molecules		O_2 NO_3^-, SO_4^{2-} Organic substances	Aerobic respiration (oxidative phosphorylation) Anaerobic respiration Fermentation (substrate phosphorylation)
Photolithotrophs	Light	H_2O, H_2S, S	NAD, NADP	Photosynthesis (photophosphorylation)
Photoorganotrophs	Light	Organic substances (e.g. acids, sugars)	NAD, NADP	Anoxygenic photosynthesis (photophosphorylation)
Chemolithotrophs	Mineral reduced substances (e.g. H_2, NH_4, H_2S, S, Fe^{2+}, Cu^{2+})		O_2, NO_3^-, CO_2, H_2SO_4, etc.	Sulphatoreducing, methanogenic, nitrifying, etc.

NAD, nicotinamide adenine dinucleotide; NADP, nicotinamide adenine dinucleotide phosphate.

Table 1.3 Nutritional groups.

Groups	Carbon source	Energetic groups
Heterotrophs	Organic molecules such as sugars, amino acids, phenols, lignins, etc.	Chemoorganotrophs Some photo- or chemolithotrophs
Autotrophs	CO_2, $-CH_3$	Photoorganotrophs Photolithotrophs Chemolithotrophs
Methylotrophs	CO_2	Chemoorganotrophs

Energy can be derived from three main sources: light, oxidation of minerals or oxidation of organic compounds. Most bacteria can use one or two of these energy sources, e.g. light through photosynthesis or a sugar through aerobic or anaerobic respiration. The transformation of these energy sources into an assimilable form of energy is coupled to oxidation–reduction reactions, the transfer of electrons and protons involving various electron donors and acceptors (Table 1.2). The names given to the different classes refer to both energy sources and electron donor criteria.

The nature (organic/mineral) of the carbon source used provides a helpful classification (Table 1.3). Chemoorganotrophs are heterotrophic organisms which usually use the same molecule as both energy and carbon sources. For instance, glucose is respired, yielding ATP, CO_2, electrons and H^+, and assimilated into other organic compounds via the so-called intermediary metabolism.

Members of the three other energetic groups, photoorganotrophs, photolithotrophs and chemolithotrophs are most often autotrophs, i.e. use a mineral carbon source, CO_2, although some may show heterotrophic capacities.

One group, the methylotrophs, at the frontier between organo- and lithotrophs, utilizes $-CH_3$ as electron donor and CO_2 as carbon source.

3 Cell division and population growth

3.1 Cell division

With few exceptions, prokaryotes multiply by binary fission, one cell yielding two daughter cells, equivalent and genetically identical to the parental one and derived from a grossly equal partition of all constituents. It is thus normally impossible to distinguish the parent from the progeny in the two cells of the following generation. These cells will, in time, give rise to two new progeny cells each, through the

same process. Thus the material of the original parent will pass to all successive descendants (of course, accidents may lead to the death of a cell, and thus to the loss of this original material). Except for some species displaying complex multicellular structures and certain levels of differentiation, no ageing or death of the original organism, as defined for multicellular beings, thus occurs.

A prokaryotic cell includes two categories of constituents (Table 1.1). Firstly, there are those present in numerous copies, such as ribosomes, metabolites, etc., of which more can always be made provided the genetic information is conserved. The second category covers only two types of complexes, each present in only one copy per cell, the envelope (taken as the sum of all the external layers) and the genetic material. While an approximate distribution of constituents of the first category in the two daughter cells is sufficient, it is of vital importance that the genetic material be copied and shared equally, and that the envelope be continuously synthesized so as to ensure cell integrity upon division.

Molecules and macromolecules of the first category are synthesized permanently. So is the synthesis of membrane and cell wall materials, although a particular process of synthesis starting along a diameter of the cell, but different for Gram-positive and Gram-negative bacteria, takes place at the initiation of division. At cell division, the newly made layers proceed towards the inside of the cell, making a septum which separates the two daughter cells.

Reproduction of the chromosome, which may require the whole span of time separating two divisions, is initiated only once per cell cycle, under usual growth conditions. It is coordinated in both space and time with the formation of the septum, so as to ensure the presence of one copy in each progeny cell, except in very fast-growing cultures for which chromosome reproduction will take longer than the division time, resulting in the necessity for simultaneous chromosome replication cycles (Chapter 2).

3.2 Cell and population growth cycles

Bacteria do not display a cell cycle in the strict sense, with a succession of well-defined phases leading to duplication and equal partitioning of the unique cell constituents, as takes place during mitosis in eukaryotic cells. Under theoretical conditions in which no limitation or modification of the available nutrients occurred, cell growth and divisions could continue indefinitely. In practice, however, be it in natural habitats or under artificial laboratory conditions, the amount of nutrients is limited. This, together with secondary effects (changes in pH, excretion of toxins, etc.), always sets a limit to proliferation. Thus, in a given set

of physically and nutritionally determined conditions, a population growth cycle can be defined.

3.3 A typical growth curve

A characteristic curve, schematically illustrated in Fig. 1.2, represents the population growth cycle as obtained under strictly defined conditions in laboratory assays. Several phases are apparent:

1 The latent phase is not, strictly speaking, part of the cycle since it is not a compulsory phase. It is present only when the organisms need to adapt to different, less favourable, conditions, or when starved cells are used as inoculum.

2 The exponential phase is the actual growth phase. It derives its name from the binary mode of division, which yields an exponential increase in cell number as divisions take place. The mean time necessary for one such division (a doubling of the number of cells per unit volume of culture medium), called the generation time (usually denoted G), can easily be determined from the semilogarithmic representation of the growth curve. It is a characteristic parameter of the growth of a given strain under a given set of conditions (physical and chemical). The mathematical equation of exponential growth utilizes another term, μ, the growth rate, which is strictly related to G:

$$\mu(h^{-1}) = \ln 2/G,$$

and is determined as $dX/dt = \mu X$, where X is the biomass at time t.

3 The stationary phase indicates the end of growth. It takes place

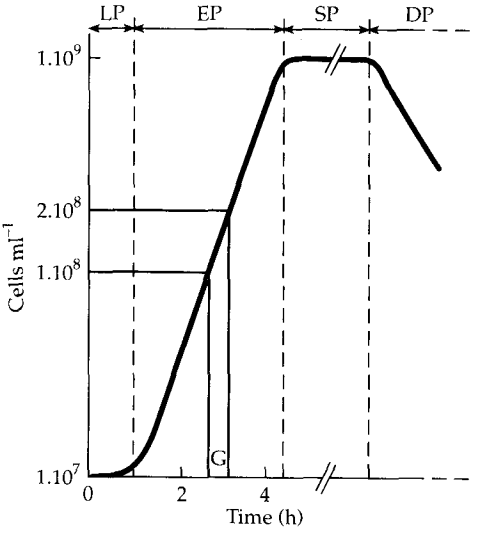


Fig. 1.2 Standard schematic growth cycle. Schematic growth cycle of *E. coli* at 37 °C in a rich medium under aeration. DP, death phase; EP, exponential phase; LP, latent phase; SP, stationary phase.

when an unfavourable factor prevents continuation of growth, e.g. exhaustion of one (or several) nutrient(s) (referred to as limiting factor(s)), or an unacceptable modification of the physical conditions (change in pH for instance). This may represent an equilibrium between the number of cells dividing and dying.

Since growth takes place at the expense of nutrients, a yield can be defined for each nutrient concentration, N:

$$Y_F(X/N) = dX/dN.$$

The general equation for the determination of yield can be written:

$$Y_F(X/N_1) = \frac{X_{max} - X_0}{N_0 - N_{max}}$$

where X_{max} = maximal biomass obtained at the stationary phase, X_0 = biomass of the inoculum (at t_0), N_0 = concentration of nutrient N at time t_0, and N_{max} = concentration of nutrient N at stationary phase. When growth ceases because of exhaustion of a nutrient, i.e. this nutrient is limiting, the yield for this nutrient can easily be evaluated, since the equation becomes

$$Y_F(X/N_1) = \frac{X_{max} - X_0}{N_0}.$$

Note that Y is usually larger than 1, except if referring to the energy source.

The yield constitutes the second parameter, with the growth rate, which characterizes the growth of organisms of a given strain under given conditions.

4 The last phase depicted in Fig. 1.2, the death phase, although a normal outcome of the end of growth, can follow very variable kinetics depending on the conditions and the organism. As noted above, death takes place statistically in the whole population and not in particular cells sorted out through a physiological process of ageing. The kinetics of death are also exponential. It follows that, unless unfavourable conditions last long enough, a fraction of the population will usually survive. This is important to remember when one is setting up sterilization conditions or mutant selection procedures.

4 Secondary metabolism

The functions grouped under the designation secondary metabolism are very heterogeneous, their common characteristic being their apparent dispensability for the producing cells. The molecules produced through secondary metabolic processes belong to a wide spectrum of chemical families, e.g. alkaloids, peptides, epoxides, pyrrols, hydro-