CHAPTER 7

DIFFERENTIATION, DEVELOPMENT, AND PROGRAMMED CELL DEATH

M. CRISTINA CARDOSO and HEINRICH LEONHARDT
Franz Volhard Clinic/Max Delbrueck Center for Molecular Medicine,
Humboldt University, 13122 Berlin, Germany

INTRODUCTION

A central issue in understanding the mechanisms of development is to elucidate how cellular division is controlled. In animal development, most structures are generated by a combination of cell division, cell migration, and cell death. Interestingly, in plants, cell migration plays no role, and cell death, though used for many purposes, does not contribute to organogenesis. During plant morphogenesis, controlled patterns and numbers of cell divisions constitute the main strategy (Jurgens, 1995; Meyerowitz, 1997). In this chapter, we focus on control of cell division and death in animal development, including epigenetic regulation.

CELL CYCLE REGULATION AND DIFFERENTIATION

In higher eukaryotes, one can distinguish three types of cells in terms of their proliferative properties: the ones that are always dividing throughout the lifetime of the organism (e.g., epithelial cells or hematopoietic stem cells); the ones that remain in a quiescent state (designated G0) but can be stimulated to proliferate (e.g., T and B lymphocytes or fibroblasts); and the ones that are terminally differentiated and cannot start proliferating and do not divide throughout the life of the organism. The latter cell type, which includes skeletal myotubes, cardiac myocytes, adipocytes, and neurons, can grow (hypertrophy) in response to particular external stimuli but cannot divide (hyperplasia). They differ from quiescent cells in this additional block. Hypertrophy is actually the only mode of coping with increasing functional demands that is accessible to

The Molecular Basis of Cell Cycle and Growth Control, Edited by G.S. Stein,
R. Baserga, A. Giordano, and D.T. Denhardt
several vital tissues such as myocardium, skeletal muscle, and nerves, which lose their proliferative capacity during ontogeny. This growth strategy ensues that such tissues have little to no regenerative capacity and cannot compensate for malfunction or loss of their structural units (reviewed in Goss, 1966). That raises the question of whether the terminally differentiated cells in these tissues irreversibly lose essential components of the replicative apparatus or whether this block to proliferation is an actively maintained state. Elucidating how terminally differentiated cells can grow but, at a certain stage of development or differentiation, loose their ability to divide, should indicate ways to reverse this state to obtain tissue regeneration and, in the opposite direction, to suppress malignancy in these tissues. In fact, uncontrolled cell proliferation leading to cancer results from an impediment in the execution of a particular differentiation program, and, consequently, cancers can be as distinct as the cell types from which they originate. This is not so to say that incomplete differentiation will yield malignant tumors or that tumors cannot also contain differentiated cells. Cell fusion experiments between normal and malignant cells showed that the hybrids, in which malignancy was abrogated, executed the normal cell differentiation program (reviewed in Harris, 1990).

Cell cycle studies have been traditionally performed using quiescent cell systems such as fibroblast cells or primary lymphocyte cultures. The vast majority of the cells that form our most essential organs permanently lose their ability to proliferate during development and, therefore, belong not to the quiescent type but to the terminally differentiated cell type. Initially the work concentrated on how cell cycle progression is stimulated and how cellular transformation is achieved. Over a decade ago, researchers working in once separate fields, such as DNA tumor viruses, growth factors including their respective signaling pathways, oncogenes, and tumor suppressors, recognized that their target factors were the same. Also, work on very different systems spanning from yeast to human cells, including marine invertebrates, Drosophila, and Xenopus, established many basic principles that are remarkably conserved throughout evolution. Hence, cyclins and their respective cyclin-dependent kinases (see Chapter 2, this volume) were recognized as key positive regulators of cell cycle functioning even across species. The pursuit of factors that could bring about cell cycle arrest led then to a search for negative regulators, which included tumor suppressors such as the retinoblastoma-related proteins and the more recently isolated cyclin-dependent kinase inhibitors (see Chapter 2, this volume).

Recent observations from gene expression studies indicated that many cell cycle regulators exhibit differences in their tissue-specific abundance. The latter suggests, on the one hand, that members of the same family may function in a nonredundant manner in vivo and, on the other hand, that they might have differentiation-specific functions either in the establishment or in the maintenance of a particular differentiated tissue. This hypothesis has been recently supported by the tissue-specific effects of mice with null mutations in genes coding for cell cycle factors (e.g., cyclin
D1−/− mice show reduced body size, a hypoplastic retina, and failure to undergo steroid-induced proliferation of mammary epithelium during pregnancy [Sicinski et al., 1995], whereas cyclin D2−/− mice have hypoplastic ovaries and testes [Sicinski et al., 1996]). Coming from another research area, studies on the characterization of terminal differentiation had, more than three decades ago, clearly established that, upon appropriate environmental cues, determined progenitor cells stop proliferating and differentiate into the corresponding lineage (Konigsberg et al., 1960; Stockdale and Holtzer, 1961). In fact, cell proliferation and terminal differentiation are mutually exclusive phenomena, and both occur at specific times and places during development. In the past years, as several of the basic molecules that regulate cell cycle progression were uncovered as well as antibodies and probes for both activating and inhibiting regulators of cell cycle became available, much attention turned to how terminally differentiated cells establish and maintain their permanent cell cycle withdrawal at a specific time in development. Indeed, the answer to these questions is a pivotal issue in elucidating the mechanisms of development.

Among the different terminal differentiation systems, skeletal muscle stands as a paradigm, and, though much progress has recently been made with neuronal and cardiac cultures, the lack of appropriate cell lines has slowed down progress in these areas. Myogenesis became a model system in which to dissect the molecular basis of the terminal differentiation program and how it can be disrupted, mostly because of the availability of several in vitro culture systems that can reproduce all the basic features of myogenic differentiation. These systems include primary cultures of fetal myoblasts or newborn satellite cells, as well as established mammalian myogenic cell lines, and have been used to study terminal differentiation for over three decades. By now, an enormous wealth of information has accumulated about many different aspects of the differentiation process in this cell type. In particular, in the last few years a combination of molecular probes together with cell lines and animals with mutations in myogenic determination genes or in cell cycle regulators allowed direct testing of the key molecular determinants of cell cycle control during terminal differentiation in this cell type. Hypotheses mostly derived from work on skeletal muscle cells have inspired experiments to test whether similar mechanisms operate in other differentiated systems, including cardiogenesis and neurogenesis. Though among these cell types there are differences in the cell cycle regulators present, their roles are likely to be at least partially overlapping. We will, therefore, use the skeletal muscle system to present the current view of cell cycle regulation during differentiation.

About 30 years ago, different myogenic cell culture systems from avian and rodent origins were first generated and the basic phenomenology of the differentiation process established. Myoblast cells can terminally differentiate in vitro and form syncytial myotube cultures, in this manner recapitulating the developmental program of muscle formation. In vivo as in vitro, conversion of undifferentiated myoblasts into terminally
Figure 7.1. Establishment, maintenance, and reversal of terminal differentiation. Conversion of undifferentiated myoblasts into terminally differentiated myotubes is induced under low-mitogen conditions and entails a sequence of cellular events. These include induction of an array of muscle-specific genes (biochemical differentiation), fusion to form multinucleated myotubes (morphological differentiation) that in vivo will ultimately form functional myofibers, and permanent withdrawal from the cell cycle. Differentiated myotubes respond to mitogen stimulation by the expression of immediate-early genes and progression through S and G1, but can neither express late G1 markers nor re-enter S phase. This inability can be reversed by expression of viral oncoproteins in the myotubes, indicating that it is not an irreversible state but rather an actively maintained one. Upon viral oncoprotein expression, S phase is induced and muscle gene expression repressed.

differentiated myotubes entails a sequence of cellular events. These include induction of an array of muscle-specific genes (biochemical differentiation), fusion to form multinucleated myotubes (morphological differentiation) that in vivo will ultimately form functional myofibers, and permanent withdrawal from the cell cycle (Nadal-Ginard, 1978; and references therein) (Fig. 7.1).

By taking advantage of the availability of myogenic cell lines and of the molecular cloning and characterization of muscle-specific regulatory sequences, it was possible to identify two major classes of muscle transcription factors in vertebrates, the MyoD family and the myocyte-enhancer factor 2 (MEF2) family. The former includes four members: MyoD, myogenin, Myf-5, and Mrf-4.
MyoD (Davis et al., 1987), Myf5 (Braun et al., 1989), myogenin (Edmondson and Olsson, 1989; Wright et al., 1989), and MRF4 (Rhodes and Koniecny, 1989; Braun et al., 1990; Miner and Wold, 1990). These myogenic basic helix-loop-helix (bHLH) proteins bind to a DNA sequence termed the E box as heterodimers with the ubiquitously expressed bHLH factors of the E-protein family (reviewed in Weintreub, 1993; Olson and Klein, 1994). The MEF2 family also encompasses four members, MEF2A–D (Pollock and Treisman, 1991; Yu et al., 1992; Breitbart et al., 1993; Martin et al., 1993, 1994; McDermott et al., 1995), which generate multiple isoforms through alternative splicing and bind an A/T rich DNA sequence termed the MEF2 site (reviewed in Olson et al., 1995). A combination of work in different vertebrate systems, including gene knockouts in mice, led to the following picture: Early in development, expression of MyoD and Myf5 in proliferating precursor cells (normally in different subsets of cells) is necessary for the determination of these cells to the myogenic lineage and to the subsequent initiation of a cascade of events leading to the expression of the MEF2 factors; these activate expression of myogenin, which, in addition to stimulating transcription of muscle structural proteins also stimulates MEF2 transcription, creating a positive feedback loop resulting in high levels of MEF2 and myogenin and, later in differentiation, of MRF4. Both of these transcription factor families synergistically activate muscle-specific promoters containing either or both DNA-binding sites. In addition to directing muscle-specific transcription and in a manner independent of its transcriptional activity, at least MyoD has been shown to be involved in cell cycle withdrawal during terminal differentiation (Crescenzi et al., 1990; Sorrentino et al., 1990; Thorburn et al., 1993), thereby creating a mechanism for crosstalk between differentiation and cell cycle machinery. Indeed, muscle differentiation is intimately linked to the cell cycle in that muscle gene transcription can only be initiated in myoblasts in G0/G1 phase. This suggests that factors expressed or functional only in this phase should be necessary for differentiation or that at other cell cycle phases inhibitors of myogenesis are present. Elucidating how the activity of these myogenic factors is regulated during the cell cycle should help to clarify this issue.

Cessation of proliferation is a prerequisite for myogenic cells (and other cell types) to switch on the terminal differentiation program, and indeed several viral and cellular oncogenes and growth factors were shown to block myogenic differentiation (Fig. 7.1) (reviewed in A’lemu and Tato, 1994; Olson, 1992). Peptide growth factors (see Chapter 6, this volume) have been shown to act as positive or as negative regulators of muscle differentiation. Fibroblast growth factor (FGF) (Clegg et al., 1987) and transforming growth factor type-β (TGF-β) (Massague et al., 1986; Olson et al., 1986) were reported to act as potent inhibitors of myogenic differentiation, whereas insulin-like growth factors (IGF) were shown to stimulate muscle differentiation (Ewton and Florini, 1981). Oncogenes from all functional groups, when expressed in myoblast cells, induce cellular transformation and block differentiation. Inhibition of
Differentiation was correlated with interference, at different levels, with myogenic transcription factor function. Proliferating myoblasts express at least one myogenic transcription factor (MyoD or Myf5) before differentiation, which under mitogen-deprivation conditions initiates a series of events including the induction of a repertoire of muscle-specific genes and the down regulation of proliferation-associated genes, ultimately leading to the formation of multinucleated myotubes. Several levels of regulation have been proposed to negatively control the activities of these factors in proliferating myoblasts. One level is the inhibition of the accumulation of myogenic factor transcripts by fos, jun, src, ras, and msx1 (Massague et al., 1986; Olson et al., 1986; Konieczny et al., 1989; Lassar et al., 1989; Falcone et al., 1991; Beagal et al., 1992; Song et al., 1992; Woloshin et al., 1995). Id, a dominant negative HLH protein, has been shown to competitively inhibit transcriptionally active heterodimer formation between myogenic factors and E2 proteins (Benezra et al., 1990; Jen et al., 1992). Another level of regulation is protein kinase C-mediated phosphorylation of the basic domain of myogenic transcription factors, which has been shown to interfere with their DNA binding in some cases, but it does not seem to do so in others (Li et al., 1992; Hardy et al., 1993). Recently, inhibition of MyoD and myogenin muscle gene activation by cyclin D1-dependent kinase was reported (Rao et al., 1994; Rao and Kohtz, 1995; Skapek et al., 1995, 1996; Gao and Walsh, 1997), providing another important piece in the molecular crosstalk between differentiation and cell cycle machinery (Fig. 7.2). Interestingly, the C-terminal acidic region of cyclin D1 was shown to be required for inhibition of myogenic factor transcriptional activity, but the N-terminal domain, which binds the retinoblastoma protein (pRb), was dispensable (Rao et al., 1994). Consistent with this finding, phosphorylation of pRb by cyclin D1-dependent kinase was not necessary for inhibition of MyoD transactivation of muscle transcription by this G1 cyclin (Skapek et al., 1996).

The postmitotic state of myotubes differs significantly from the reversible G1 cell cycle arrest of quiescent fibroblasts in that mitogen stimulation does not lead to proliferation (Konigsberg et al., 1960; Stockdale and Holtzer, 1961). Terminally differentiated myotubes can still respond to serum stimulation by expressing immediate-early genes and progressing until a mid G1 state (Fig. 7.1) (Endo and Nadal-Ginard, 1986; Tsai and Olson, 1996a). Therefore, myotubes have not permanently lost essential components of the signal transduction pathway that links cell membrane with the nucleus, though cell surface receptors are in general downregulated (Olwin and Hauschka, 1988; Hu and Olson, 1990).

After fusion, mitogen stimulation can lead to hypertrophy, but nuclear DNA synthesis and cell division are permanently inhibited (Fig. 7.1). This inability could be due to the loss or irreversible alteration of the basic cellular machinery necessary for DNA replication and mitosis during differentiation. This fundamental question was first addressed by infecting or microinjecting muscle cultures with transforming DNA and
RNA viruses such as polyoma virus, simian virus 40 (SV40), and Rous sarcoma virus (RSV) (Fogel and Defendi, 1967; Yaffe and Gershon, 1967; Lee et al., 1968; Graessmann et al., 1973) because these viruses had been shown to induce host DNA synthesis and cell division in cells arrested in the cell cycle either by contact inhibition or by X-irradiation (Dubecco et al., 1965). DNA synthesis as measured by tritiated thymidine incorporation could be detected at the time of fusion of the infected myoblasts into myotubes, but preformed myotubes were refractory to viral infection (Fogel and Defendi, 1967). Other reports, though, presented evidence of cellular DNA synthesis and mitotic figures in fully differentiated myotubes (Yaffe and Gershon, 1967; Lee et al., 1968; Graessmann et al., 1973). These results were questioned by other groups, which proposed that DNA synthesis and mitosis can be induced in the nuclei of cells that have just fused but not in the nuclei of mature myotubes (Holtzer et al., 1975; Falcone et al., 1984; Hu and Olson, 1990). It was only in 1989, with the generation and characterization of a myoblast cell line carrying a thermosensitive mutant form of SV40 T antigen under the control of an inducible promoter that this question was finally answered. In the course of studies on the molecular nature of the terminal differentiated state, the C2C12 mouse myogenic cell line
(Yaffe and Saxel, 1977; Blau et al., 1985) was stably transfected with a heavy metal–inducible, thermolabile SV40 T antigen (C2SVTts cells) (Endo and Nacal-Ginard, 1989; Cardoso et al., 1993; Gu et al., 1993). This cell line can differentiate into multinucleated muscle fibers under nonpermissive conditions for T antigen expression and activity, but, upon mitogen stimulation and shift to permissive conditions, it is able to re-enter the cell cycle, as opposed to control “wild-type” muscle cultures (Fig. 7.1). The patterns of DNA replication foci observed in myotube nuclei induced to retrodifferentiate are indistinguishable from those of cycling myoblasts and fibroblasts (Cardoso et al., 1993). The visualization of typical early, middle, and late S phase replication patterns indicates that myotubes retain the potential to undergo a complete and well-coordinated S phase, which can be unlocked simply by expression of the viral SV40 T antigen (Cardoso et al., 1993) but not by a mutant T antigen form (K1 mutation), which is unable to bind pRb (Gu et al., 1993). Yet, continuous expression of large T is not required, suggesting a role in overriding a block and/or passing a restriction point, but not to actually drive DNA replication as in the viral infection (Cardoso et al., 1993), and supporting a view of terminal differentiation as a product of active and continuous regulation (Blau, 1992) rather than an irreversible default state.

In addition to SV40 large T antigen, adenovirus E1A 12S protein has been recently shown to induce cell cycle entry in terminally differentiated muscle cells (Crescenzi et al., 1995; Tiainen et al., 1996b). SV40 T antigen and adenovirus E1A are to date the only known proteins able to reactivate the cell cycle in terminally differentiated cells. Both of these DNA tumor virus oncogenes have no cellular counterparts. This differs from retroviral oncogenes whose expression can in many cases inhibit establishment of differentiation but whose cellular homologues remain inducible upon mitogenic stimulation in myotubes and cannot mediate retrodifferentiation.

A potential target of large T antigen and E1A during reversal of terminal differentiation is pRb. Elevated Rb mRNA levels, either due to transcription rate variation or to increased stability, as well as the hypophosphorylation of the pRb and tethering to the nucleus, have been correlated with cellular differentiation (Chen et al., 1989; Gu et al., 1993; Thorburn et al., 1993). Also, the hypophosphorylated pRb forms a complex with SV40 large T antigen (Hill et al., 1989), which correlates with the transforming properties of this viral oncogene (Cherington et al., 1988; DeCaprio et al., 1989). Rb phosphorylation as a consequence of large T antigen expression could lead to an increase in the free E2F-DP transcription factor complexes (see Chapter 5, this volume). The latter is presumably the transcriptionally active form (Hiebert et al., 1992; Weintraub et al., 1992) and could then stimulate transcription from promoters containing E2F sites, among which are several genes encoding proliferation-associated proteins (Fig. 7.2) (reviewed in Nevins, 1992; Muller, 1995). In agreement with this hypothesis, E2F and DP transcription factors have been found associated with pRb and
p130 in extracts of differentiated muscle cultures, indicating that either of these tumor suppressor proteins or both may regulate E2F in muscle (Corbel et al., 1995; Halevy et al., 1995; Kiess et al., 1995; Shin et al., 1995).

To directly test whether pRb inactivation by T antigen is necessary for cell cycle re-entry in terminally differentiated myotubes, a myogenic cell line derived from teratomas obtained from Rb<sup>−/−</sup> embryonic stem cells was generated (the CC42 cell line). The characterization of this cell line showed that the pRb-related protein p107 can substitute for pRb during establishment of muscle differentiation. Contrary to pRb, though, p107 is downregulated upon serum addition and cannot, thereafter, maintain the postmitotic state of myotubes (Schneider et al., 1994). Similar results have been independently obtained and extended to p130-deficient muscle cells. The latter was also shown to be unable to prevent cell cycle re-entry of myotubes (Novitch et al., 1996). Furthermore, studies of muscle differentiation in a mouse model expressing low levels of the pRb protein clearly demonstrated that during myogenesis in vivo pRb is essential for myoblast differentiation and survival (Zackenhaus et al., 1996). Another possible target of E1A and SV40 T antigen is the nuclear protein p300, and, in fact, the ability of E1A to prevent skeletal muscle differentiation was shown to require sequences that bind p300 (Mymryk et al., 1992; Caruso et al., 1993). Recently, p300 was shown to potenti ate MyoD and myogenin-dependent transcription activation and to be required for MyoD-dependent cell cycle arrest (Pyri et al., 1997). Nevertheless, p300 is dispensable for the maintenance of the postmitotic state of myotubes and cannot substitute for pRb.

In addition to its role in the establishment and maintenance of permanent cell cycle withdrawal, pRb was shown to be required for activation of the muscle-specific transcription program in cooperation with MyoD (Gu et al., 1993). More specifically, pRb-deficient myogenic cells can induce the expression of early differentiation markers such as myogenin and the cyclin-dependent kinase inhibitor p21 but fail to elicit the expression of late differentiation markers such as myosin heavy chain (Novitch et al., 1996). At present, it is unclear how pRb cooperates with the myogenic factors to bring about the expression of the full differentiation program. Although ectopic expression of pRb stimulates MyoD-mediated transactivation of muscle promoters, pRb is not present in E-box DNA-binding complexes, which contain heterodimers of myogenic factors and E proteins (Gu et al., 1993). On the other hand, pRb and p130 are found in E2F-DP DNA-binding complexes in muscle extracts, but these complexes lack any associated myogenic factor. Thus, it is feasible that direct in vivo binding does not occur, and, alternatively, an indirect regulation takes place. Interestingly, in transient assays the Rb gene promoter has been shown to be transactivated by coexpressed MyoD by a mechanism independent of direct DNA binding (Martelli et al., 1994), suggesting an indirect mechanism for cross-regulation.

In addition to Rb, the transcription of the general cyclin-dependent
kinase inhibitor p21 is also upregulated during myogenic differentiation in a manner that is independent of p53 but can be induced by MyoD (Fig. 7.2) (Halevy et al., 1995). This upregulation has also been observed in vivo during embryonic muscle differentiation (Parker et al., 1995). Even though p21 and p27 (as well as other cyclin-dependent kinase inhibitors such as p18) (Franklin and Xiong, 1996) are expressed at high levels in differentiated myotubes, they cannot maintain the postmitotic state of myotubes in the absence of functional pRb (Novitch et al., 1996). Furthermore, in spite of high p21 levels, pRb phosphorylation and re-entry into S phase occurred upon expression of E1A in pRb wild-type myotubes (Tlaien et al., 1996b). A primary function of cyclin-dependent kinase inhibitors in myotubes may be to maintain the hypophosphorylated, activated state of pRb, but they are not sufficient to maintain the block to cell cycle progression and prevent pRb phosphorylation as a consequence of SV40 T antigen or adenovirus E1A protein activity. An exception occurs in myotubes cultured from newts. In these amphibians, myotubes can dedifferentiate and yield proliferating mononucleated blastema cells in the regenerating limb (Lo et al., 1993). Although newt myotubes in culture have hypophosphorylated pRb, upon serum stimulation they undergo S phase. pRb is thought to be the physiological substrate of cyclin D/CDK4 complexes, which carry out the first inactivating phosphorylation steps of pRb. Cell cycle re-entry in newt myotubes also requires phosphorylation of pRb via a cyclin D/CDK4/6-dependent mechanism, but can be blocked by overexpression of cyclin-dependent kinase inhibitors (Tanaka et al., 1997). The difference in the regulation of pRb between newt and mammalian muscle might explain the amazing regenerative potential of these animals.

Our understanding of the molecular basis of cell cycle control during terminal differentiation is still limited, but several of the essential components and links have been identified, especially in the skeletal muscle system. As depicted in Figure 7.2, in mitogen-rich conditions, high levels of cyclin D1 mediate repression of (1) pRb growth suppressive activity and (2) MyoD transactivation of the general cyclin-dependent kinase inhibitor p21 and muscle-specific genes, thus keeping myoblasts in a proliferative state. Mitogen depletion (or other inductive cues) induces the downregulation of cyclin D1. This allows accumulation of underphosphorylated pRb, which represses via E2F-DP binding the transcription of proliferation-associated genes. In addition, release of cyclin D1 inhibition of MyoD transactivation leads to the induction of the muscle differentiation program and stimulation of p21 accumulation and cell cycle withdrawal. It is now possible to test multiple hypotheses derived from this model using skeletal muscle as well as other terminal differentiation systems.

In myogenic differentiation, as in other differentiation systems, the fine balance between opposing cellular signals determines whether the cell will divide or differentiate. When the correct control of the mutual exclusion between proliferative and differentiated states fails, the outcome is either neoplasia or apoptotic cell death.
APOPTOSIS DURING DEVELOPMENT

In the previous section, we focus on molecular mechanisms that are the basis for the decision to divide or to differentiate. In animal development, another important process regulating the appropriate numbers of each cell type is cell death by a controlled process termed apoptosis. Cell death during mammalian embryogenesis starts at the blastula stage and continues throughout the entire life as a means of keeping tissue homeostasis.

As a cell depends on extracellular signals to proliferate or to differentiate, it also does so to survive, and, in their absence, it activates an intrinsic program leading to death. This dependence on environmental signals for division and differentiation ensures that a particular cell type will divide or differentiate only when and where it is needed. The abrogation of this dependence on extracellular factors in the absence of cell death results in uncontrolled proliferation and ensuing neoplasia. To the decision whether to turn on the proliferation or the differentiation programs is added the decision to switch on the death program, which, from a certain point on at least, is an irreversible process. The balance among these three phenomena results in the development of specialized structures that execute particular functions in the organism.

A clear distinction between the processes underlying the death of cells during embryogenesis and tissue kinetics (and, also, some pathological states) and the one occurring in the center of acute conditions such as ischemia or hypoxia was established by Kerr et al. (1972). Morphological evidence indicated that in the latter cases nuclei become pyknotic (condensed, hyperchromatic nuclei), and cells and organelles swell and rupture, leaking their contents and provoking an inflammatory response. This cell death process is called necrosis, and it is characterized by general cell lysis due to physical or chemical insult to the cell. Necrosis is a pathological process, and it occurs as a consequence of impaired membrane transport with uncontrolled movement of ions and fluids. The cell death occurring during embryogenesis, metamorphosis, normal tissue turnover, and endocrine-dependent tissue atrophy called apoptosis entails, by contrast, cell shrinkage and membrane blebbing. Moreover, the plasma membrane and the membranes of organelles do not rupture; rather, they undergo changes in their composition, the chromatin condenses, the nuclear DNA becomes fragmented, and the nucleus collapses. One of the early events in the apoptotic program is the activation of an inside–outside phosphatidylserine (PS) translocase, which catalyzes the externalization of PS in the outer leaflet of the plasma membrane (Martin et al., 1995). This early change in the composition of the plasma membrane signals to the neighboring cells via their PS receptors that apoptosis is in progress, and, before there is leakage of cellular contents, the apoptotic cell is phagocytosed by its neighbors or by macrophages in the absence of inflammation.

Several methods, in addition to morphological observation, have been developed that allow detection of apoptotic cells. Two methods are based
CELL CYCLE, DIFFERENTIATION, AND CELL DEATH

on the enzymatic or electrophoretic detection of DNA fragmentation, which is a hallmark of genetic death in apoptosis. In one, the cellular DNA is isolated and analyzed by agarose gel electrophoresis, giving the characteristic “ladder” of 200–300 bp fragments resulting from cleavage of the chromatin between the nucleosomes during apoptosis (oligonucleosomal ladder). The second method utilizes the enzyme terminal deoxynucleotidyltransferase (TdT) to catalyze the addition of labeled nucleotides to the 3'-OH ends of the fragmented cellular DNA, forming random heteropolymers of incorporated nucleotides, which can then be detected in situ or by flow cytometry. Two other very recent protocols rely on early changes in the execution of the apoptosis program, before morphological changes are visualized. These include the above-described externalization of PS in the plasma membrane and the activation of the ICE family of proteases (see below). The former takes advantage of the nonenergy-dependent specific binding of annexin V to PS and can be performed with living or fixed cells. The latter does not require intact cells and assays for the activity of an ICE-like protease (CPP32 or caspase 3) with a substrate conjugate that upon cleavage releases free fluorescent or chromogenic conjugate that can then be detected and quantified. All these methods exploit events that are common to the death program in most mammalian cells studied thus far.

Cell death was first described during amphibian metamorphosis more than 150 years ago and was later found to occur in most developing tissues of multicellular organisms (reviewed in Clarke and Clarke, 1996). It was only about a decade ago that the first genes required for apoptosis during nematode development were identified in genetic studies (Ellis and Horvitz, 1986). Subsequent studies showed that these genes had mammalian counterparts (e.g., the Caenorhabditis elegans Ced-9 and Ced-3 proteins are homologous to the mammalian Bcl-2 and ICE, respectively) (Yuan et al., 1993; Hengartner and Horvitz, 1994) and that the human genes could function in the nematode (Vaux et al., 1992; Hengartner and Horvitz, 1994), indicating that the death program is evolutionarily conserved, as it is the case for the cell cycle machinery.

In the past few years, studies from different research areas, including viral oncogenes, the genetics of C. elegans development, amphibian metamorphosis, thymocyte maturation, and cancer and the p53 protein, have culminated in a number of very important findings on the molecular basis of apoptosis. Genetic and biochemical experiments indicated that the apoptotic process is controlled by the Bcl-2 family of proteins, which either function as agonists (Bax or Bak) or as antagonists (Bcl-2 or Bcl-XL) of apoptosis. At least 16 different genes of the Ced-9/Bcl-2 family have been identified to date (reviewed in Korsmeyer, 1995), and the various members can dimerize with each other. The ratio of apoptosis-promoting and -inhibiting factors in the cell determines whether the apoptotic program will be turned on (Fig. 7.3). A hint toward the mechanism of action of the Bcl-2 family members comes from structural studies suggesting that their three-dimensional structure is similar to bacterial colicins and that they may function, as the bacterial proteins, as pore-
forming transmembrane proteins (Muchmore et al., 1996). The various Bcl-2 family members have been found associated with the outer membrane of the mitochondria, the nuclear envelope, and the endoplasmic reticulum, and this association is mediated by a protein domain located at their C-terminal 22 amino acids. Though their exact mechanism of action is not known, genetic disruption of the Ced-9 in the nematode (Hengartner et al., 1992) and of Bcl-2 or Bcl-X in mice (Veis et al., 1993; Motoyama et al., 1995) clearly established their role in preventing cell death during development. Conversely, disruption of Bax in mice demonstrates its role in promoting neuronal cell death in development (Deckwerth et al., 1996).

Bcl-2 family members control the activation of a family of cysteine
proteases, designated ICE proteases (see Fig. 7.3). These are related to the interleukin-1β-converting enzyme (ICE), which was identified as being homologous to the C. elegans protein Ced-3 that is required for all the cell deaths occurring during worm development (Yuan et al., 1993). Multiple members of this family have since been identified and in some cases implicated in apoptosis (reviewed in Nagata, 1997). They are all cysteine proteases that cleave their substrates after specific aspartate residues (therefore, they are designated caspases, for cysteine aspar- tate proteases) (Alnemri et al., 1996) and they are themselves activated by cleavage at specific aspartic acid residues. By now, there are at least 10 ICE members identified in humans, which can be subdivided into three groups (ICE-like, CPP32-like, and Ich1-like) based on their sequence homology (Alnemri et al., 1996). Specific recognition sequences have been identified for three caspases (1, 3, and 6) as well as some of their respective substrates. The latter include procaspases, poly(ADP) ribose polymerase (PARP), DNA-PK, rho-GDI, actin, and lamin A. The observation that, at least in vitro, caspases can activate themselves and also other caspases suggests that they function in a proteolytic cascade much as the cellular cascades of protein kinases that transduce proliferative or stress-activated signals. Because, in all animal cells studied, specific peptide or protein caspase inhibitors can block the cell death program, this family of proteins seems to play a central role in apoptosis by cleaving intracellular proteins from different subcellular compartments (nucleus, nuclear lamina, cytoskeleton, cytoplasm, and endoplasmic reticulum). A recent example of their fundamental role in a characteristic morphological change in cells undergoing apoptosis is the nuclear collapse. When mutant lamin A, carrying an amino acid substitution in a conserved position that prevents its cleavage by caspases, is overexpressed in mammalian cells and apoptosis induced, nuclear envelope breakdown and cell death is delayed by several hours, though other nuclear changes, such as nuclease activation, still take place at the normal pace (Rao et al., 1996).

Although the role of several of the cell deaths in development is not fully understood, the analysis of mutant organisms (from invertebrates to humans) begins to clarify the functions of the individual components either in tissue-specific or in generalized developmental defects. Cell death in development serves multiple purposes (reviewed in Sanders and Wride, 1995), among them sculpturing structures or deleting unneeded ones. Examples of the former are elimination of interdigital material during digit formation in mammals; creation of the preamniotic cavity in preimplantation mouse embryos; development of the presumptive neural tissue during gastrulation and formation of the neural plate during neurulation; and heart morphogenesis. In the latter case, one study reported 31 zones in the developing chick heart where apoptosis occurs (Pexieder, 1975), which clearly indicates an important role of cell death in the remodeling that occurs in this vital organ. During embryogenesis, some structures are created that are only required by one sex and are deleted at a certain time in the other sex by apoptosis. Other structures are remnants required by ancestral species or structures
needed only at one stage of development, which are eliminated by apoptosis. As an example, during amphibian metamorphosis, there is massive muscle apoptotic cell death in the regressing tadpole tail and also in the remodeling of the dorsal body muscle during larval-to-adult maturation (Nishikawa and Hayashi, 1995).

Another function of apoptosis is the elimination of overproduced cells or of dangerous and nonfunctional cells. An example of the former happens in the differentiation of the nervous system, where half of the neurons and oligodeadocytes produced die (reviewed in Cowan et al., 1984). Overproduction followed by massive cell death may be a mechanism to match the number of innervating neurons to the target cells and to match the number of glial cells to the axons they myeline. Neuronal survival is largely determined by the limited amount of neurotrophic factors produced by the target cells and by the competition for those factors with other neurons. Interestingly, inhibition of apoptosis by disrupting one of the ICE-like proteases (CPP32/caspase 3) yields mice that die around birth with a vast excess of cells in the central nervous system (Kuida et al., 1996). A very well-studied system that illustrates the role of apoptosis in the elimination of dangerous cells is the maturation of T and B lymphocytes, during which more than 95% of the cells produced die by apoptosis. The diversity of lymphocyte receptors generated via undirected genetic rearrangements during differentiation inevitably causes potentially harmful self-reactive receptors. These are eliminated by apoptosis in the thymus, and this maturation process requires the activity of death factor–receptor pairs. To date, three death factors (TNF, FasL, TRAIL) and four death factor receptors (TNFR1, Fas, DR3/Wst-1, CAR1) have been isolated. Mice strains with mutations in FasL (gld) (Takataishi et al., 1994) and Fas (lpr) (Watanabe-Fukunaga et al., 1992; Adachi et al., 1995) clearly demonstrated the importance of this factor–receptor pair in the death of lymphocytes. Binding of the ligand FasL is thought to induce oligomerization of the receptor with transduction of the death signal by recruiting a protein called FADD/Mort 1, which in turn binds to caspase 8, initiating the protease cascade of the apoptotic program (reviewed in Nagata, 1997).

Abnormal or nonfunctional cells resulting from different cellular stresses (ultraviolet or ionizing radiation, heat shock, oxidative stress) are also removed by apoptosis. Though the pathway is not well understood, stress stimuli seem to involve initially a cascade of kinases belonging to the stress-activated protein kinases (SAPK) and the p38/RK family. It is not known how this signaling pathway interacts with the Bcl-2 family and/or with the ICE protease cascade, because most of the known targets of these kinases are transcription factors (reviewed in Cosulich and Clarke, 1996). In the case of DNA damage, the tumor suppressor p53 is often involved in the response process, and it can either cause cell cycle arrest while the damage is being repaired via the upregulation of cyclin-dependent kinase inhibitor p21 (El Deiry et al., 1993), or, possibly when the damage is too severe, it can induce apoptosis via upregulation of Bax (Miyashita and Reed, 1995). The upregulation of
these two factors is mediated by p53 at the transcriptional level (reviewed in Levine, 1997).

Finally, cells that either do not complete or attempt to reverse terminal differentiation, most often activate the apoptosis program. The ones that do not die will likely give rise to tumors (see previous section). Two types of experimental approaches have consolidated this proposal. One has been the analysis of the phenotype of mice with homozygous deletion in cell cycle regulators and in transcription factors. The other has resulted from attempts to reinude cell cycle progression in terminally differentiated cells such as cardiocytes and myotubes. The best example of the former case has been the analysis of pRb-deficient mice. These mice show extensive cell death in the tissues that normally express high levels of the pRb protein, which include liver, nervous system, lens, and skeletal muscle (Clarke et al., 1992; Jacks et al., 1992; Lee et al., 1992; Zacksenhaus et al., 1996).

These observations suggest that pRb is required to prevent apoptosis in the cell population committed to differentiate in addition to playing a role in the differentiation process itself. In accordance with this, ectopic expression of pRb has been reported to inhibit apoptosis induced by ionizing radiation (Haas-Kogan et al., 1995). One possibility is that apoptosis is activated upon inappropriate cell cycle entry, and the corollary of it would be that pRb prevents apoptosis by keeping cells out of the cell cycle. In support of this proposal, many genes involved in cell cycle progression are also associated with cell death. Examples of these are myc (Evan et al., 1992), cyclin D1 (Freeman et al., 1994), E1A (Debbas and White, 1993), and E7 (Pan and Griep, 1994). All of these proteins have been shown to bind and antagonize the function of pRb in cell cycle arrest. In addition, there is evidence suggesting that early steps in cell cycle progression also take place in the cell death pathway. Epithelial cells from the prostate were shown to synthesize DNA and induce proliferation markers such as proliferative cell nuclear antigen (PCNA) before undergoing apoptotic DNA fragmentation (Colombel et al., 1992). Another line of evidence comes from terminally differentiated cells induced to re-enter the cell cycle. Over 20 years ago, morphological observations of myotubes transformed with a temperature-sensitive mutant of the oncogenic Rous sarcoma virus and differentiated under nonpermissive conditions showed that, when these cultures were shifted back to the permissive temperature for viral function, myotubes vacuolated and degenerated completely within 72 hours (Holtzer et al., 1975). More recently, retrodifferentiation of myotubes by inducible expression of temperature-sensitive SV40 T antigen mutant not only achieved re-entering the cell cycle but also a certain percentage of the myotubes showed morphological features of apoptosis (Endo and Nadal-Ginard, 1989). A separate study also using the heat-labile SV40 T antigen mutant, reported that at nonpermissive conditions the cells underwent apoptosis. These cells expressed at permissive conditions high levels of wild-type p53 complexed with large T antigen, which after the temperature shift and T antigen degradation was not destabilized and was no
longer in a complex with T antigen. Because overexpression of p53 was shown to induce apoptosis, the authors suggested that this accumulation of p53 could be responsible for the apoptosis in these cells (Yanai and Obinata, 1994). Whether this is also the cause of the cell death observed in part of the cells undergoing retrodifferentiation needs further investigation. SV40 T antigen expression in a different type of cells, the cerebellar Purkinje cells from transgenic mice, did not bring about tumorigenesis; rather, it correlated once again with cell ablation (Feddersen et al., 1992). In another terminally differentiated muscle cell type, cardiac myocytes, expression of E1A (Kirshenbaum and Schneider, 1995; Liu and Kitis, 1996) or of E2F1 (Kirshenbaum et al., 1996) induced S phase re-entry and apoptosis. The kinetics of both processes suggested that the DNA synthesis preceded the apoptosis and that the cells did not complete the cell cycle (Liu and Kitis, 1996). Expression of the adenoviral E1B protein together with E1A or with E2F1 avoids apoptosis and allows the cell cycle to progress (Kirshenbaum and Schneider, 1995; Kirshenbaum et al., 1996), most likely via inhibition of p53 activation of apoptosis (Debbaas and White, 1993).

Altogether, these results reinforce the idea that the processes of cell cycle, cell death, and cell differentiation are coupled, and this is at least in part achieved by "sharing" some of the same regulatory molecules like pRb.

**EPIGENETIC PATHWAYS IN DEVELOPMENT AND GROWTH CONTROL**

One of the most fundamental discoveries in molecular biology was the deciphering of the genetic code. According to standard textbooks, the genetic alphabet is made out of four letters (nucleotides or bases): A (adenine), C (cytosine), G (guanine), and T (thymine). However, if we put this hypothesis to a simple test and analyze some of our own DNA, we find that there is in fact a fifth letter. About 1% of all bases in our DNA is 5-methylcytosine (\(^{\text{5}}\)C; Fig. 7.4). Now, where does this fifth nucleotide come from, how does it change our view of genetics, and what does it have to do with cell cycle and growth control? The purpose of this section is to address these questions. We first give a brief introduction to the basic knowledge and methods and then outline the role and regulation of DNA methylation.

This fifth, minor base, 5-methylcytosine, is generated by a postreplicative modification called *DNA methylation*, whereby a DNA methyltransferase (DNA MTase) transfers a methyl group to the target cytosine at the palindromic DNA sequence 5'-'CG-3' (CpG sites) (see Fig. 7.4). However, not all of these CpG sites are methylated. On average, only 60%–70% are methylated in mammalian cells, and the pattern of methylated and unmethylated CpG sites changes during development and disease.

The following brief overview of current knowledge in this rapidly
progressing field shows that DNA methylation is anything but just a minor modification. By now, it is clear that DNA methylation is absolutely required for mammalian development and plays a critical role in the regulation of gene expression. This means, on the other hand, that errors in the methylation pattern can have far-reaching effects. DNA methylation can cause the erroneous activation or inactivation of genes and thus offset the balance of growth promoting and growth suppressing factors, just like changes in the DNA sequence itself. This potential of DNA methylation has long been overlooked or underestimated, but it finally caught broad attention and was recently found in many types of tumors. In addition to this direct effect on gene expression, DNA methylation can also indirectly affect the balance of growth control, because 5-methylcytosine residues are also hot spots for mutations. Altogether, these observations warrant a closer examination of the role of this fifth base in the regulation and deregulation of growth control in mammalian cells.

**Detection and Regulation of DNA Methylation**

A variety of techniques are available with which to analyze the methylation status of DNA (Grigg and Clark, 1994). The easiest and traditional method is based on digestion with methylation-sensitive restriction endonucleases. Most often used is the enzyme HpaII, which recognizes and cleaves the DNA sequence CCGG but not C^mCGG. For comparison the DNA is cut with its isoschizomere MspI, which does not discriminate between methylated and unmethylated sites and thus cleaves both. This assay can only be performed on the original genomic DNA because the methylation gets lost after subcloning and propagation in *Escherichia coli*. Therefore, the products of the HpaII and MspI digests are analyzed by Southern blotting techniques. Because one blot can be hybridized multiple times with different probes, the methylation levels of several
regions of interest can be determined from one sample. The disadvantage of this approach is that only 1 out of 16 possible methylation sites (C<sup>n</sup>G but not A<sup>n</sup>CGG, and so forth) can be tested with this method.

Because methylation at a few single sites can often play a critical role in the regulation of gene expression, other methods are needed. The most comprehensive but also technically demanding one is the bisulfite method for selective base conversion. The basic principle is that cytosine residues in single-stranded DNA are efficiently deaminated to uracil in the presence of sodium bisulfite. Under these conditions, 5-methylcytosine residues do not react and remain unchanged (Frommer et al., 1992). The methylation status can then be determined with standard sequencing techniques because uracil is being read as a T (thymine) in place of the original cytosine. Moreover, the degree of methylation at any cytosine residue can be quantified by automated fluorescence-based genomic sequencing (Paul and Clark, 1996). This is also extremely sensitive because, after the selective base conversion in the first step, the template can be amplified by polymerase chain reaction and theoretically allows analysis of single cells.

Using any of the above methods to follow the DNA methylation pattern, one can find that the pattern of methylated and unmethylated sites is precisely maintained over many cell division cycles (Wigler et al., 1981). This process is often referred to as maintenance methylation. However, during development and differentiation dramatic changes, i.e., de novo methylation as well as loss of methylation, can be observed. (These processes are schematically outlined below in Fig. 7.6) Given the far-reaching effects of DNA methylation, it becomes clear that a precise maintenance as well as the controlled change of methylation pattern is crucial for the viability of mammals. The basic questions are (1) how, on the one hand, the methylation pattern is precisely maintained over many rounds of DNA replication and, on the other hand, changes are introduced at specific developmental stages; and (2) what goes wrong in some diseases.

We are still far from completely understanding the regulation of DNA methylation, but an important step in this direction was the cloning of the mouse and later the human DNA MTase (Bestor et al., 1988; Yen et al., 1992). This recombinant DNA MTase shows a strong preference for hemimethylated sites, which are generated by semiconservative DNA replication, and specifically transfers the methyl group to the newly synthesized strand at these sites. However, unmethylated sites are not recognized and remain unchanged. These enzymatic properties can in principle account for the maintenance of a given methylation pattern after DNA replication (Fig. 7.5). Still, an important question was how is this precision achieved in vivo; after all, there are on the order of 10<sup>6</sup> hemimethylated sites to be recognized and methylated during each cell cycle. A first clue came from studies on the subcellular localization of the DNA MTase. It has previously been shown that DNA replication occurs in a discrete pattern of nuclear foci that undergo characteristic changes throughout S phase (reviewed in Leonhardt and Cardoso, 1995).
**Figure 7.5.** Regulation of DNA methylation patterns. The basic processes are outlined with the example of a specific DNA fragment containing five CpG sites of which three are methylated. In quantitative terms, the predominant process is the maintenance of a given methylation pattern after replication of the genome. The enzyme responsible for this reaction, DNA MTase (see also Fig. 7.6), has a strong preference for hemimethylated sites, i.e., sites that carry a methyl group in the "old" strand and are unmethylated in the newly synthesized strand. Completely unmethylated sites are not modified, and thus the methylation pattern is precisely maintained. As discussed in the text, changes occur at certain stages of development and disease. At this moment, it is not yet known which enzymes are involved in the de novo methylation. On the other side, demethylation can occur by two different processes, either by active removal of 5-methylcytosine from the DNA or by a passive mechanism of failing to fill in a methyl group in the newly synthesized strand.

Several replication enzymes (PCNA, DNA polymerase α, and replication protein A) were found to be tightly associated with these foci (Bravo and Maccanakl-Bravo, 1985; Cardoso et al., 1993; Hozak et al., 1993). It came as a surprise to find that DNA MTase was also localized at these foci, just like a replication enzyme. Further experiments identified a targeting sequence that is necessary and sufficient for localization at replication foci. This sequence is located in the N-terminal regulatory domain of DNA MTase and by itself can direct unrelated proteins like the β-galactosidase from *Escherichia coli* to these replication foci (Fig. 7.6). These results suggested an assembly-like organization of DNA replication and methylation, which might explain the high fidelity of the
Figure 7.6. Structure of the mammalian DNA MTase. The known DNA MTase has a catalytic domain of about 500 amino acids that is very similar to the group of prokaryotic C6 DNA MTase and in fact has all the conserved peptide motifs described for this group (Leonhardt and Bestor, 1993). The regulatory domain contains the targeting sequence that is necessary and sufficient to direct DNA MTase to nuclear sites of DNA replication (Leonhardt et al., 1992). At this moment the exact size and structure of the regulatory domain is unclear, because new upstream sequences were discovered that could potentially extend the open reading frame by up to 171 amino acids and include also a testis-specific exon (Tucker et al., 1996). The role of these tissue-specific forms in the maintenance and change of DNA methylation patterns remains to be elucidated.

The maintenance of DNA methylation in mammalian cells (Leonhardt et al., 1992; Leonhardt and Cardoso, 1995).

The role of this maintenance MTase was clearly demonstrated by site-directed gene disruption. Homozygous mice (Dnmt-/-) carrying a deletion in both alleles die at midgestation (E. Li et al., 1992). Interestingly, Dnmt-/- ES cells contain residual amounts of 5-methylcytosine and retain some de novo methyltransferase activity (Lei et al., 1996). These results suggest that, in addition to the known maintenance MTase (see Fig. 7.6) there are also one or more de novo MTases, which might introduce changes in the methylation pattern at specific stages, as was postulated more than 20 years ago (Holliday and Pugh, 1975). The postulated de novo MTase remained undetected despite intensive searches, and only recently were ESTs (expressed sequence tags) published that are homologous but not identical to the already known maintenance MTase, suggesting that another MTase exists in mammalian cells. The intriguingly complex structure of the known MTase suggests that it may be more than “just” a maintenance enzyme and that it might also participate in the change of methylation pattern. A first hint in support of this hypothesis comes from the recent identification of tissue-specific MTase mRNA forms (Tucker et al., 1996b). It was furthermore shown that partial in vitro proteolysis can convert this DNA MTase into an enzyme with de novo specificity (Bestor, 1992).

Another way of changing the methylation pattern is a loss of methylation. As outlined in Figure 7.5, there are two possible pathways, an active demethylation and a passive loss. The latter requires two rounds of DNA replication, in which a particular site is not methylated. This loss could be caused by transcription factors either blocking the site or preventing DNA MTase from accessing the site, e.g., by preventing DNA MTase from binding to the replication foci or by retaining it in the cytoplasm.
Interestingly, DNA MTase was found in the cytoplasm during preimplantation development, when the total concentration of 5-methylcytosine in the genome is decreasing (Carlson et al., 1992).

Direct evidence exists for active demethylation, i.e., the loss of methylation at specific sites in the absence of DNA replication. Well-documented examples of active demethylation include the chicken vitellogenin gene (Wilks et al., 1982, 1984), the rat α-actin gene (Paroush et al., 1990), and the chicken δ-crystallin (Sullivan and Grzinger, 1987). Thus far, two different mechanisms of demethylation have been described. The vitellogenin gene becomes demethylated by an excision repair-like process (Jost, 1995). Demethylation of the α-actin gene was shown in vitro to be protease resistant and to require a catalytic RNA (Weiss et al., 1996). In summary, great progress has been made over the past years, but we are still far from understanding the regulation of DNA methylation. From the first results it is likely to be as complex as the regulation of gene expression itself.

**DNA Methylation and the Control of Gene Expression**

DNA methylation sites (CpG dinucleotides) are not evenly distributed throughout the genome, and their methylation status is not random. CpG dinucleotides are clearly underrepresented in the genome and occur only at about one-fifth of the statistically expected frequency (Russell et al., 1976). In the course of evolution, CpG sites were most likely destroyed by C to T transitions (see below and Fig. 7.4). The few remaining CpG sites within the coding part of genes are highly methylated, while unmethylated CpG sites are clustered at the 5' end of all housekeeping genes and about 40% of all tissue-specific genes (Bird, 1986; Larsen et al., 1992). Due to the high density of CpG sites in the promoter region, these clusters are generally referred to as CpG islands.

There are literally hundreds of studies showing an inverse relationship between DNA methylation in the promoter region and transcriptional activity and the respective changes throughout development and differentiation. Just to mention one out of many studies in the early 1980s, it was shown that in the process of lymphocyte maturation the transcribed immunoglobulin genes were always hypomethylated and that the methylation in these loci decreased as maturation progressed (Storb and Arp, 1983). This type of observation raised the question of whether it was just a correlation or whether DNA methylation was indeed the cause of gene inactivation or maybe only a consequence of gene inactivation that was brought about by other means. The effect of DNA methylation on gene expression was directly analyzed by introducing unmethylated and in vitro methylated reporter constructs into mammalian cells. Using, for example, the γ-globin gene, it was shown that methylation in the promoter region, but not in the coding part of the gene, prevents transcription (Busslinger et al., 1983).

In principle, two different models involving a direct or an indirect mechanism can be envisaged to explain the effect of DNA methylation...
on gene transcription. Studies on the expression control of the tyrosine aminotransferase gene showed that all promoter-binding factors were present in both expressing and nonexpressing cells. However, the promoter region was methylated in nonexpressing cells, and that prevented binding of these factors (Becker et al., 1987). Similar results were obtained with an adenovirus major late promoter (Watt and Molloy, 1988) and the cAMP-responsive element (Iguchi-Ariga and Schaffner, 1989). Interestingly, methylation of the SP1-binding site, which is typical for housekeeping promoters, did not prevent transcription factor binding (Höller et al., 1988). These experiments clearly show that DNA methylation can specifically interfere with protein–DNA interactions and thus influence gene expression.

Even though the methyl group is strategically well positioned in the major groove of the DNA double helix, where it can directly interfere with the binding of regulatory factors, in most cases the effects on transcription control are probably exerted by nonspecific means. In that regard, it was shown that DNA methylation prevents the formation of active chromatin and renders integrated DNA sequences inactive and DNase I insensitive (Keshet et al., 1986). The formation of inactive chromatin is mediated by nuclear factor(s) that specifically bind to methylated CpG sites (Boyce and Bird, 1991). Thus far, two different methyl-CpG–binding proteins (MeCP1 and MeCP2) have been identified (Meehan et al., 1989; Lewis et al., 1992). MeCP2 is, like DNA MTase itself, dispensable in embryonic stem cells but is required for embryonic development (Tate et al., 1996).

In summary, by now there is no question that DNA methylation can do control gene expression. It is also clear that DNA methylation preserves the inactive state of genes, but it remains to be elucidated how the methylation of CpG islands is regulated. Whatever the final picture may look like, DNA methylation is undoubtedly an essential element in the control of gene expression during mammalian development. Most of the last paragraphs dealt with its effect on the expression of endogenous genes, but also the expression of foreign DNA is affected. Mammalian cells seem to be able to recognize foreign DNA and inactivate it by DNA methylation. Integrated viral sequences, e.g., are highly methylated, and infectivity as well as viral gene expression are inversely correlated with methylation (Stuhlmann et al., 1981; Jähner et al., 1982). In fact, viral sequences are scattered throughout our genome, but—fortunately for us—they are highly methylated and transcriptionally inactive. What appears like a blessing in the case of viral infection constitutes a serious problem for gene therapy trials, where transgene expression decreases with time.

**DNA Methylation During Differentiation, Development, and Aging**

As discussed throughout this section, DNA methylation often correlates with gene activity and changes as genes are being turned on or off during
development and differentiation. The regulation of these developmental changes in the methylation pattern is still far from being understood, but DNA methylation is clearly more than an accompanying molecular ornamentation. DNA methylation by itself can indeed induce entire differentiation programs. This is best illustrated by some early experiments treating undifferentiated fibroblasts (10T1/2 and 3T3 cells) with the demethylating agent 5-azacytidine. Within several days after addition of the drug clear phenotypical conversions took place, and several types of differentiated cells were observed, including mostly contractile striated muscle cells, adipocytes, and chondrocytes (Taylor and Jones, 1979). Also other differentiation programs like smooth muscle cell differentiation (Iehara et al., 1996) and neuronal differentiation (Bartolucci et al., 1989) can be induced by forced hypomethylation. By now, much progress has been made in studying the induction of myogenic differentiation. One factor that is directly and/or indirectly activated by hypomethylation and can convert fibroblasts to myoblasts is the myogenic transcription factor MyoD (Davis et al., 1987). The MyoD promoter is partially methylated in nonmethylated skeletal muscle (Zingg et al., 1994). A distal enhancer element is also methylated in nonmuscle tissues, and demethylation precedes transcription of MyoD (Brun et al., 1996). The role of DNA methylation is further supported by experiments showing myogenic differentiation upon overexpression of a DNA MTase antisense mRNA (Szyf et al., 1992). Curiously, overexpression of the DNA MTase itself in myoblasts was also reported to accelerate myogenic differentiation and myotube formation (Takagi et al., 1995). Altogether, these results clearly show a regulatory role of DNA methylation during myogenic differentiation; however, the mechanisms causing these changes in the methylation pattern during differentiation remain to be elucidated.

In the first part of this chapter, it is discussed that transformation/proliferation and differentiation are mutually exclusive, which requires that differentiation factors be inactivated in the process of cellular transformation. In support of this hypothesis, hypermethylation of the MyoD promoter region was observed during immortalization and transformation of fibroblast cell lines (Jones et al., 1990), as well as in the process of oncogenic transfection (Rideout et al., 1994). The reverse also seems possible and, for example, forced hypomethylation induces neuronal differentiation in neuroblastoma cells (Bartolucci et al., 1989).

Among these developmental changes in the methylation pattern, two phenomena, chromosomal imprinting and X-inactivation, stand out in view of their particular regulation and significance in biology. Early pronuclei transplantation experiments on one-cell-stage embryos indicated that the genomes inherited from the father and the mother are not equivalent. Embryos with either two maternally or two paternally derived genomes could not complete embryogenesis (McGrath and Solter, 1984). Because both genomes should be identical at the DNA sequence level, there had to be something else that distinguished both
genomes and rendered them nonequivalent. This distinctive mark also had to be of a transient nature because, after passage through the germ-line, an initially maternally derived genome might then be inherited from a father. In brief, this distinctive mark is a difference in the methylation pattern between maternal and paternal chromosomes. As a consequence of these different patterns, different sets of genes are expressed from the paternal and the maternal genomes. Genes exhibiting this parent-specific expression pattern are called imprinted genes (for recent reviews on this rapidly progressing area, see Barlow, 1995; Latham et al., 1995). There are altogether 17 imprinted genes known, including insulin-like growth factor-II (IGF-II) which is discussed in more detail below. The most recent entry to this list of imprinted genes is the human potassium channel gene KVLT1 (Lee et al., 1997; Neyroud et al., 1997), and more imprinted genes are likely to be discovered soon.

Whatever the intricate regulatory mechanisms may be that lead to the transcriptional silencing of imprinted genes, the disruption of the mouse DNA MTase gene (Dnmt) leading to genome-wide hypomethylation clearly demonstrated that DNA methylation is required to maintain this inactive status (Li et al., 1993). The parent-specific imprint, i.e., methylation pattern, is acquired during oogenesis and spermatogenesis. Once this imprint is erased by genetic manipulations it can only be restored by renewed passage through the germ-line (Tucker et al., 1996a).

A special case of developmentally controlled monoallelic expression is the X-chromosome inactivation. In this case, an entire chromosome (one of the two X chromosomes in female somatic cells) is transcriptionally silenced and methylated. Methylation does not appear to be the cause, however, but rather a secondary phenomenon because, in the case of the X-linked Hpht gene, methylation clearly occurs after inactivation and transcriptional silencing (Lock et al., 1987). Instead, inactivation seems to be initiated by the cis-acting regulatory RNA Xist, which is expressed before X-inactivation (Kay et al., 1993). The Xist gene itself is one of the above-mentioned imprinted genes and is expressed from the inactive X-chromosome. DNA hypomethylation causes the ectopic expression of the Xist gene on the active X chromosome, which then causes the inactivation of this chromosome (Panning and Jaenisch, 1996).

Finally, the DNA methylation pattern also changes in the process of aging. In some protooncogenes, e.g., c-fos and c-myc, age-dependent changes in the DNA methylation pattern were observed (Ono et al., 1989; Uehara et al., 1989), suggesting that altered gene expression caused by DNA methylation may affect cellular vitality in the senescent phase. In some cases, similar changes were observed in the process of aging and tumor formation. In the aging colorectal mucosa, for example, increasing methylation of the estrogen receptor was observed, which might constitute an age-dependent predisposition to sporadic colon tumors (Issa et al., 1994). Moreover, the age-dependent changes in the methylation pattern of the c-fos gene can also be found in the early stages of hepatocellular}
carcinogenesis (Choi et al., 1996). Although it is still too early for general conclusions, the changes in the methylation pattern occurring in the aging organism might affect the overall fitness of cells and also represent a predisposition for neoplastic transformations.

**DNA Methylation and Cancer**

DNA methylation can affect cell cycle and growth control by two different mechanisms. First, ectopic changes in the methylation pattern can cause either transcriptional silencing of growth suppressing factors or activation of growth promoting factors. Second, 5-methylcytosine is also a hot spot for mutations, which can cause inactivation of growth suppressors. Both mechanisms can unbalance the regulation of growth control and thus cause tumor formation (Fig. 7.7). For a better understanding of these different roles of DNA methylation during neoplastic transformation and tumor outgrowth, one has to take into account the distribution of methylated and unmethylated CpG sites in the genome (Fig. 7.8).

Methylation of the CpG island in the promoter region causes transcriptional silencing, and deamination of methylated sites within the coding sequence causes mutations.

It was early observed that in the process of tumor formation the overall level of 5-methylcytosine decreases (Feinberg and Vogelstein,

---

**Figure 7.7.** Balance of growth regulators. The successful control of cell proliferation in normal tissues depends on a balance of growth promoting and growth suppressing factors, as well as apoptosis. This balance can be disturbed by a variety of different mechanisms potentially leading to uncontrolled cell proliferation and neoplasia. Two of these mechanisms, transcriptional silencing, and C to T transition mutations, involve DNA methylation (see Figs. 7.4 and 7.8).
gene inactivation by

de novo methylation  point mutation

promoter coding sequence

Figure 3.8. Inactivation of genes by two methylation-dependent mechanisms. Unlike all other dinucleotides, CpG sites are not evenly distributed throughout the genome. They tend to be clustered in the promoter region, where they form the so-called CpG island (see text), and they are clearly underrepresented within the coding sequence of genes. Moreover, cytosine residues in the promoter region are usually unmethylated, whereas cytosine residues within the coding sequence are mostly methylated. Genes can be inactivated by either de novo methylation in the promoter region or by C to T transition mutations within the coding sequence. Transcriptional silencing by de novo methylation is a normal and integral mechanism of transcription control during development, but errors may cause accidental inactivation of genes with dramatic consequences for the organism.

1983; Gama-Sosa et al., 1983; Feinberg et al., 1988), whereas selected CpG islands are methylated (Bayin et al., 1987). Practically nothing is known about how these changes in the methylation pattern are caused. Several studies showed an elevated level of DNA methyltransferase activity in tumor cells, which seems to increase during the progression of neoplasms (Kauttainen and Jones, 1986; El Deiry et al., 1991). However, it remains obscure how this increased activity can be compatible with the observed general hypomethylation and the specific methylation of CpG islands. Moreover, tumors are characterized by a higher proliferation, which is accompanied by a general upregulation of proliferation-associated genes, and higher DNA methyltransferase levels may simply be a reflection of this overall change (Lee et al., 1996).

The pRb was the first identified tumor suppressor protein and served as a paradigm for tumor formation. In most cases, inactivation of the Rb gene is caused by a mutation (deletion and point mutations), but in about 10%–20% of unilateral retinoblastomas hypermethylation was found in the 5' region of the Rb gene (Greger et al., 1989; 1994; Sakai et al., 1991). In some sporadic retinoblastomas (16% of all unilateral tumors), no mutations were detected, and it was found that inactivation was caused by hypermethylation, which prevented binding of transcription factors (Ohtani-Fujita et al., 1993). This type of analysis was later
extended to the newly discovered family of cyclin-dependent kinase
inhibitors, and hypermethylation with transcriptional silencing was found
for p15 (Herman et al., 1996) and p16 (Gonzalez-Zulueta et al., 1995;
Herman et al., 1995; Merlo et al., 1995; Otserson et al., 1995).

In more than half of human breast cancers, hypermethylation and
absence of mRNA of the mammary-derived growth inhibitor were ob-
served (Huynh et al., 1996). In addition to these suppressors, other
growth regulators like the estrogen receptor (Issa et al., 1994) and even
a DNA repair enzyme (Harris et al., 1996) were found to be hypermethyl-
ated and transcriptionally silenced during transformation. Finally, differen-
tiation factors like the MyoD transcription factor were also found to
become hypermethylated during oncogenic transformation (Rideout et
al., 1994).

Many of these studies show a correlation between tumor formation,
transcriptional silencing, and hypermethylation, but in some cases a
direct causal relationship could be established. In some renal carcinoma
cell lines, re-expression of the von Hippel-Lindau gene was achieved by
treatment with the demethylating agent 5-azacytosine (Herman et al.,
1994). The promoter of the cell adhesion molecule E-cadherin was also
found to be hypermethylated in breast and prostate carcinomas. These
cells were shown to contain the necessary transcription factors, and 5-
azacytosine treatment indeed caused re-expression of E-cadherin as well
as concomitant reversal to epithelial cell morphology and adhesive prop-
erties (Graff et al., 1995; Yoshiura et al., 1995).

This growing evidence that hypermethylation of tumor suppressor
genes may play an important role during carcinogenesis also inspired
the reverse approach, i.e., to screen for hypermethylated sites in tumors.
Using this approach, a new potential tumor suppressor gene (HIC-1,
hypemethylated in cancer) was identified (Wales et al., 1995). In light
of these results, the action of known carcinogens was also revisited and,
e.g., nickel seems to cause chromatin condensation and DNA methyla-
tion (Lee et al., 1995).

The above discussion concerns the inactivation of tumor suppressors
in a broad sense. However, the activation of growth factors can also be
caused by changes in the DNA methylation pattern and lead to cancer.
As discussed above, most genes are expressed from both alleles, but
some genes are expressed exclusively from either the paternal or the
maternal chromosome, a process termed imprinting. Transcription of
the imprinted genes is blocked by DNA methylation. Imprinted genes
generally include factors involved in cell cycle and growth control, and
normal development depends on the balanced expression of these fac-
tors. Relaxation of imprinted genes was first observed in Wilms' tumor
patients. Normally, IGF-II is expressed from the paternal chromosome,
but about two-thirds of Wilms' tumors showed biallelic IGF-II expression
(Ogawa et al., 1993; Rainier et al., 1993). Loss of imprinting of the IGF-
II gene was also discovered in a number of other human tumors.

DNA methylation is also a hot spot for mutations, and it can affect
cell cycle and growth control. From early studies in E. coli, it is clear
that methylated cytosines are hot spots for C to T transitions. Analyse of mutations causing human genetic diseases have shown that this is also true in humans, and it was estimated that about 35% of all point mutations occur at CpG sites (Cooper and Youssoufian, 1988). This mutational hot spot is also involved in the inactivation of cell cycle regulators. In the case of the p53 tumor suppressor gene, it was calculated that on average about 25% of all mutations reported in all human tumors occur at the CpG dinucleotide (Greenblatt et al., 1994). These data clearly indicate that "normal" DNA methylation by itself can pose a risk to human health. In general, it seems that less methylation is better, and in fact forced hypomethylation appears to decrease the incidence of intestinal neoplasia in some tumor-prone mice strains (Laird et al., 1995).

The molecular basis for this high mutation rate at CpG sites lies in the similarity between 5-methylcytosine and thymine (note that both have a methyl group at the C-5 position; see Fig. 7.4). Hydrolytic deamination of 5-methylcytosine, but not cytosine, causes a direct conversion from C to T. It was shown that this conversion can be enhanced by DNA MTase itself under conditions of limiting S-adenosylmethionine (SAM); (Shen et al., 1992). These results concur with an earlier observation that a methyl-deficient diet lowers the intracellular concentration of SAM and enhances liver tumor formation in rats (Mikol et al., 1983). Admittedly, the conditions used for these experiments were rather extreme and are highly unlikely to occur in human nutrition, but less dramatic changes in the intracellular SAM levels might be effective in the long run. The latter idea was indeed supported by two large health studies in the United States showing that a low dietary intake of methionine and folate and a high consumption of alcohol, which seems to influence the methyl group availability, increases the risk for colorectal neoplasia (Giovannucci et al., 1993). Thus, the old adage to "eat fresh fruit and vegetables and not drink too much alcohol" makes sense from the point of view of DNA methylation and cancer. It is still too early for definite conclusions, but these results may link DNA methylation, dietary habits, and carcinogenesis.

PERSPECTIVES

Most of the above section deals with all of the problems that can be caused by DNA methylation, raising the question of what is it good for. The results with DNA MTase–deficient mice, which die at midgestation, indicate that we cannot live without it. After millions of years with DNA methylation, it became an indispensable regulatory factor in the coordinated and developmentally controlled expression of genes.

At first, the demethylating agents like 5-azacytidine and 5-aza-2'-deoxyctydidine appeared to be miracle drugs. They reactivate silenced tumor suppressors, reverse the process of oncogenic transformation, lower the risks of mutation, and so forth, but—as so often in life—there are side effects involved, and more specific approaches are needed. For
a long time little was known about the regulation of DNA methylation. Now, since the cloning of mammalian DNA MTase and the new possibilities of genetic manipulations in mice, rapid progress is being made and will lead to a better understanding of the regulation and, most importantly, the deregulation of DNA methylation in mammals. Future studies should elucidate genetic, dietary, and environmental factors affecting DNA methylation. With this knowledge, we might some day be able to positively influence DNA methylation in the prevention of cancer and in the process of aging. We may also be able to use it for antiviral defenses and bypass it in gene therapy approaches.

REFERENCES


Blau HM, Pavlath GK, Hardenan EC, Chiu CP, Silberstein L, Webster SG,


CELL CYCLE, DIFFERENTIATION, AND CELL DEATH


