Targeting and Association of Proteins with Functional Domains in the Nucleus: The Insoluble Solution

Heinrich Leonhardt and M. Cristina Cardoso
Humboldt Universität Berlin, Franz-Volhard-Klinik am Max-Delbrück-Centrum für Molekulare Medizin, Department of Nephrology, Hypertension, and Genetics, 13122 Berlin, Germany

The mammalian nucleus is highly organized into distinct functional domains separating different biochemical processes such as transcription, RNA processing, DNA synthesis, and ribosome assembly. A number of proteins known to participate in these processes were found to be specifically localized at their corresponding functional domains. A distinct targeting sequence, necessary and sufficient for the localization to DNA replication foci, was identified in the N-terminal, regulatory domain of DNA methyltransferase and DNA ligase I and might play a role in the coordination of DNA replication and DNA methylation. The fact that the targeting sequence is absent in lower eukaryotic and prokaryotic DNA ligase I homologs suggests that “targeting” is a rather recent development in evolution. Finally, targeting sequences have also been identified in some splicing factors and in viral proteins, which are responsible for their localization to the speckled compartment and to the nucleolus, respectively. These higher levels of organization are likely to contribute to the regulation and coordination of the complex and interdependent biochemical processes in the mammalian nucleus.

KEY WORDS: DNA methyltransferase, DNA ligase I, Cyclin A, cdk2, Replication protein A, Targeting sequence, RS domain, Nucleolus, Coiled body, DNA replication, DNA methylation, RNA splicing, Cell cycle.

I. Introduction

The goal of this chapter is to outline the functional organization of mammalian nuclei with emphasis on protein targeting. The term “protein targeting” in a broad sense refers to mechanisms that cause proteins to end up in
specific cellular compartments. The term was initially coined in the field of intracellular protein traffic and was used to refer to the controlled and energy-dependent transport of proteins across membranes and the transport between intracellular compartments in vesicles. Starting with the synthesis at the ribosomes in the cytoplasm, a new protein has several options regarding the cellular localization. Short signal or presequences, usually at the N-terminus, determine whether a protein is directly translocated into the endoplasmic reticulum, imported into mitochondria or into the nucleus, or whether it only remains in the cytoplasm. All these transport or targeting processes are energy dependent and, except for nuclear transport, the signal or presequences are cleaved off during the transport. The nuclear import is special because the nucleus breaks down during each cell cycle and most proteins must be reimported after the nucleus is rebuilt (for a detailed discussion on the nuclear import process and the function of the nuclear pore see chapters by Agutter; Panté and Aebi; and Bastos and Burke).

For a long time the nucleus was considered to be the final destination of protein targeting. This is true as there are no organizing or dividing membranes within the nucleus, but nonetheless the nucleus is anything but a homogeneous cellular compartment with freely diffusible components. Even by regular light microscopy the nucleolus can be identified as a distinct subnuclear structure. Electron micrography clearly showed that there are no membranes around the nucleolus, but still a number of proteins were found to be specifically localized in the nucleolus while other nuclear proteins were excluded. A detailed analysis led to the identification of nucleolar targeting sequences, which are necessary and sufficient for the nucleolar localization just like the above nuclear localization sequences are required for nuclear uptake (see Section II). In addition to the nucleolus there are several other subnuclear structures, which are not visible by regular light microscopy. Three distinct RNA-containing structures including interchromatin granules (IGs), perichromatin fibrils (PFs), and coiled bodies (CBs) were identified by immunofluorescence and electron microscopy. The targeting of proteins to these “compartments” and their role in transcription and RNA processing is discussed in Section III. Finally, DNA replication also occurs in distinct nuclear foci, and proteins directly or indirectly involved in DNA replication are localized and targeted to these foci (see Section IV).

The specific enrichment of proteins in distinct nuclear structures or compartments where the respective nuclear processes occur is often referred to as “functional organization of the nucleus” and obviously raises a number of interesting questions. First of all, what forces are responsible for this organization into functional domains? In case of the nuclear localization the basic principle seems to be simple; there is a separating membrane and only proteins with a nuclear localization sequence (NLS) are recognized

and imported through the nuclear pore complex. Other proteins lacking the NLS are either small enough to passively diffuse through the nuclear pore or remain outside unless they are associated with other NLS-containing factors. As mentioned above, there are no organizing membranes in the nucleus and—apart from nuclear components like chromosomes, proteins, and RNA limiting the free diffusion path—imported proteins seem to be per se freely diffusible. Hence the specific enrichment of proteins at, e.g., replication foci must be based on association and retention at these structures. In several cases targeting sequences were identified, which are necessary and sufficient for localization at these nuclear structures. Examples for targeting to nucleoli, speckled compartment, coiled bodies, and replication foci are reviewed throughout this chapter.

The functional organization of the nucleus also has consequences for the regulation of biological processes. The concentration of factors, which are involved in different steps of a complex process, in functional nuclear foci and the organization of the process in an assembly line fashion is expected to greatly increase the overall efficiency, fidelity, and specificity. Over the past few years progress has been made toward the “appreciation” of these higher order structures. After a decade of dispute over potential artificial experimental conditions, several different experimental approaches are now in place to examine the architecture of these structures and to analyze their contribution to the regulation and coordination of biological processes throughout the cell cycle.

II. Targeting of Proteins to the Nucleolus

In the interphase eukaryotic nucleus, the most conspicuous morphological structure seen in the light microscope is the nucleolus. Even though it has been a preferred target of cytologists for centuries, it was only in the 1960s that findings derived by electron microscopic autoradiography of [3H]uridine incorporation established its major function in rRNA synthesis and processing and ribosome biogenesis (Granboulan and Granboulan, 1965).

The electron microscope unveils three major ultrastructural components within the nucleolus: pale-staining fibrillar centers (FC) surrounded by dense fibrillar components (DFC), which are themselves embedded in a granular component (GC). The size of the nucleolus changes according to the metabolic state of the cell and it is the contraction or expansion of the granular component that is mostly responsible for these differences. The nucleolus is usually surrounded by a layer of condensed chromatin that is continuous with the intranucleolar chromatin in the FCs. The major components and bio-
chemical functions of the different intranucleolar compartments were essentially laid out in the mid-1960s but the details of where in the nucleolus ribosomal genes are transcribed remained highly controversial. Recently, the discrepancies in the mapping of the transcriptional sites within the nucleolus have been reevaluated by the development of a nonisotopic method based on the incorporation of 5-bromouridine-5'-triphosphate (BrUTP) into RNA in streptolysin O-permeabilized mammalian cells (Dundr and Raska, 1993). This new technique provided decisive evidence that the site of nucleolar transcription is the DFC including its boundary to the FC (Hozák et al., 1994; Hozák, 1995). The FCs contain DNA from the nucleolar organizer region (NOR) of a chromosome (i.e., the region of the chromosome containing tandem copies of rRNA genes separated by nontranscribed spacer DNA) and proteins such as RNA polymerase I and others needed for rRNA synthesis (e.g., DNA topoisomerase I and the transcription factor UBF). The DFC surrounding the FC contains the pre-rRNA transcripts with their 5' tails excised by a protein-rich granule and U3 snRNA (Mougety et al., 1993). The latter is the most abundant nucleolar-specific snRNA and it has been implicated in the early pre-rRNA processing steps (Kass et al., 1990; Savino and Gerbi, 1990; Hughes and Ares, 1991). Upon hypotonic treatment, each of these actively transcribing units shows a “Christmas tree”-like appearance, with the tip of the tree representing the point on the DNA where transcription begins. These transcripts are then processed and converted into preribosomal particles forming the GC. It is not yet clear whether these various processing and assembly steps occur at particular sites within the nucleolus or whether each pre-rRNA carries its private processing machinery, and maturation, therefore, takes place independently of the location within the nucleolus. In any event, after these steps, ribosomal precursors are released from the nucleolus into the nucleoplasm and are transported to the cytoplasm where the final steps of ribosome maturation take place. Detailed information on nuclear transcription and structure can be found in the chapters by Jackson and Cook, van Driel et al., and Moreno Diaz de la Espina.

This vectorial process involves an unusual degree of nucleocytoplasmic interaction, with precursor rRNA being synthesized in the nucleolus and being soon associated with proteins synthesized in the cytoplasm. These include most ribosomal proteins and, in addition, nonribosomal nucleolar proteins as well as small RNA molecules (nucleolar snRNAs) that are involved in pre-rRNA processing and preribosome assembly. The two last are recycled in the nucleolus when the preribosomal particles are exported to the cytoplasm. Among the nonribosomal nucleolar proteins identified so far, some may represent structural proteins forming a nucleolar skeleton (Franke et al., 1981), others are transiently associated with preribosomes and are involved in processing of rRNA (e.g., fibrillarin; Tollery et al., 1991), and others shuttle between the nucleolus and the cytoplasm (e.g., C23/nucleolin and B23/NO38, Borer et al., 1989; Nopp140, Meier and Blobel, 1992; NAP57, Meier and Blobel, 1994) (see Table I).

To establish and maintain this functional nonmembranous intranuclear compartment where rDNA transcription, pre-rRNA processing, and preribosome assembly take place, a very intricate set of protein-protein and protein-nucleic acid interactions are required at each mitotic cycle. During prophase of mitosis, the nucleolus disintegrates. In the telophase nucleus precnucleolar bodies (PNB) appear and, when rRNA transcription resumes, they fuse at the NORs forming the interphase nucleolus (Scheer and Benavente, 1990; Scheer et al., 1993). This complex process of nucleolar reformation is called nucleogenesis. The nucleolus can be used as a paradigm of how function relates to spatial and temporal organization, and indeed it uncovers, at its own scale, the underlying principles of the functional organization of the nucleus within which it is formed. Although it has no separating membranes, specific proteins have been localized to this highly specialized region of the nucleus, and some have even been assigned to certain ultra-

<table>
<thead>
<tr>
<th>Protein</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>RNA polymerase I</td>
<td>Scher and Rose (1984); Raska et al. (1989)</td>
</tr>
<tr>
<td>UBF transcription factor</td>
<td>Rendon et al. (1992); Roussel et al. (1993)</td>
</tr>
<tr>
<td>Topoisomerase I</td>
<td>Zini et al. (1994)</td>
</tr>
<tr>
<td>Topoisomerase II α and β</td>
<td>Zini et al. (1994)</td>
</tr>
<tr>
<td>NOP1/fibrillarin</td>
<td>Ochs et al. (1985)</td>
</tr>
<tr>
<td>GAR1</td>
<td>Girard et al. (1992)</td>
</tr>
<tr>
<td>B23/NO38/nucleophosmin</td>
<td>Ochs et al. (1985)</td>
</tr>
<tr>
<td>C23/nucleolin</td>
<td>Ochs et al. (1985); Pfeife et al. (1986)</td>
</tr>
<tr>
<td>Nopp140</td>
<td>Meier and Blobel (1992)</td>
</tr>
<tr>
<td>NSR1</td>
<td>Yan and Mélieé (1993)</td>
</tr>
<tr>
<td>NAP57</td>
<td>Meier and Blobel (1994)</td>
</tr>
<tr>
<td>HSP70</td>
<td>Pelham (1984)</td>
</tr>
<tr>
<td>PCNA</td>
<td>Waseem and Lane (1990)</td>
</tr>
<tr>
<td>Annexin V</td>
<td>Sun et al. (1992)</td>
</tr>
<tr>
<td>Cyclin B-cdc2 complex</td>
<td>Ooka inaka et al. (1992); Gallagher et al. (1993)</td>
</tr>
<tr>
<td>Angiogenin</td>
<td>Moroianu and Riordan (1994)</td>
</tr>
</tbody>
</table>

* Ribosomal proteins not included.
transcription factor UBF indicated that a wide region, including the HMG-box 1 (crucial for rDNA binding) and the COOH-terminal acidic tail, is required for nucleolar accumulation (Maeda et al., 1992). Similarly, deletion of 24-amino acids near the carboxy-terminus of nucleolar shuttling protein NO38 abrogates nucleolar localization, but fusion of this domain to a reporter protein did not target the hybrid protein to the nucleolus (Peculis and Gall, 1992). Extensive mutational analysis of nucleolin, the most abundant nucleolar protein in vertebrate cells, established the requirement of both the RNA-binding domains and the GAR domain for nucleolar localization (Créancier et al., 1993; Messmer and Dreyer, 1993; Schmidt-Zachmann and Nigg, 1993). Similar experiments using NSR1, a yeast nucleolar protein, demonstrated that either the amino terminus or the RNA-binding domains are sufficient to target a heterologous protein to the nucleolus, and the GAR domain, even though not sufficient, facilitates nucleolar accumulation (Yan and Mélese, 1993). More recently, a central domain (aa 21–124) was identified in the yeast nucleolar protein GAR1, which is sufficient for both nuclear uptake and nucleolar localization of a reporter protein. This protein, unlike the previous ones, does not contain a functional classical NLS nor a bipartite NLS. The authors propose that the protein could be transported to the nucleolus passively by binding to another nucleolar component (for instance, as a part of a ribonucleoprotein complex) or, alternatively, GAR1 nucleolar transport could occur via a pathway different from the other nucleolar proteins (Girard et al., 1994).

Indeed, the NOS signals of Rev, Rex and Tat also share little homology, apart from a stretch of basic amino acids. Clusters of basic amino acids may suggest binding to nucleic acids (e.g., rRNA) or to acidic proteins but, per se, are not sufficient for nucleolar accumulation since other proteins share this characteristic (e.g., histones) but do not localize to the nucleolus. The fact that there is no common nucleolar targeting signal suggests that
nucleolar recruitment is the result of functional interactions with other molecules already present at the nucleolus. In fact, the NOS domain of HTLV-1 Rex specifically interacts with the nucleolar shuttling protein B23, which is thought to play a role in ribosome assembly (Adachi et al., 1993). Moreover, B23 and other nucleolar shuttling proteins such as nucleolin (Borer et al., 1989), Nopp140 (Meier and Blobel, 1992), and NAP57 (Meier and Blobel, 1994) possibly play a dual role, in ribosome assembly (consistent with their ultrastructural localization) and in the nucleolocyttoplasmic transport of ribosomal proteins into the nucleolus and preribosomal particles to the cytoplasm.

Altogether, these experiments illustrate multiple ways by which proteins (and, possibly also nucleolar snRNAs and SS rRNA) are recruited and retained in the nucleolus dependent upon specific functional interactions with different nucleolar components. These multiple interactions establish the foundations of this biochemical and structural entity where ribosomes are formed. How is this structure reformed at each mitotic cell cycle? The nucleating sites at telophase are the rDNA loci (NORs) and these, upon transcriptional activation, recruit fibrillar structures (PNBs) that provide essential proteins and RNAs to assemble a functional nucleolus. PNBs are formed from material surrounding chromosomes that is either segregated to the daughter cells or synthesized very early after anaphase (Scheer and Benavente, 1990; Scheer et al., 1993; see also Jiménez-García et al., 1994). Certain aspects of this molecular assembly process can benefit from the availability of cell-free systems for its study, one being the assembly of PNBs in Xenopus egg extracts (Bell et al., 1992; Wu et al., 1993), and another, the observation that solubilized amoeba nucleolar components can reassociate in vitro into nucleolus-like bodies with similar protein composition and ultrastructure (Trimbur and Walsh, 1993). A detailed analysis of the macromolecular interactions leading to this coordinated assembly should elucidate how the functional compartmentalization of the nucleolus is achieved and how this architecture enhances the efficiency of the process of ribosome biosynthesis.

III. Sites and Organization of RNA Processing

Over the past two decades a vast amount of information was gathered about the transcriptional activation of individual genes and the mechanism of the subsequent RNA processing, but our understanding of the overall coordination of genome activity and the spatial organization of transcription, splicing, and mRNA export is still at an early stage. In addition to the nucleolus (see Section II) there are at least three other distinct RNA-containing nuclear structures identified by electron microscopy, which are generally referred to as perichromatin fibrils, interchromatin granules, and coiled bodies (Fakan and Puvion, 1980).

The characterization of these structures was greatly facilitated by the identification of antisera from autoimmune disease patients that react with splicing factors or small nuclear ribonucleoprotein particles (snRNPs). These antisera as well as a variety of other antibodies and in situ hybridization probes against snRNP components give a speckled staining pattern, also referred to as “speckled compartment,” which typically includes 20 to 50 bright spots (interchromatin granules) that seem to be connected by a less bright reticulum (perichromatin fibrils) (Spector, 1993). Furthermore, one to five nuclear foci are stained, which correspond to coiled bodies (Raska et al., 1991; Carro-Fonseca et al., 1992; Huang and Spector, 1992). In addition to snRNP components, non-snRNP splicing factors including SC35 and SF2/ASF also were identified at these nuclear speckles (Fu and Mainfatis, 1990; Spector et al., 1991).

These three structures—interchromatin granules, perichromatin fibrils, and coiled bodies—are morphologically distinct and also differ in their composition. While snRNP components are mostly concentrated at interchromatin granules and coiled bodies, hnRNP proteins are mostly associated with nascent RNA transcripts at perichromatin fibrils (Pino-Roma and Dreyfuss, 1993). Interestingly, the essential splicing factor U2AF was detected at coiled bodies but not at the other nuclear speckles (Carro-Fonseca et al., 1991; Zamore and Green, 1991). Coiled bodies can be specifically visualized by immunofluorescence or immunoelectron microscopy using antibodies against p80-collin (Andrade et al., 1991; Raska et al., 1991). Furthermore, they contain fibrillarin, which is otherwise found only in the nucleolus (Raska et al., 1990, 1991; Andrade et al., 1991).

Since coiled bodies do not contain DNA, hnRNP, nor the SC35 and SF2/ASF splicing factors, it was argued that they are most likely not directly involved in transcription and pre-mRNA splicing (Lamond and Carro-Fonseca, 1993).

Ever since these snRNP-containing nuclear structures were identified, their role in transcription and splicing has been discussed. The analysis of the spatial organization of RNA processing is hampered by the lack of methods to directly visualize the process of RNA splicing itself. In situ autoradiography of [3H]uridine incorporation studies showed rapid labeling at perichromatin fibrils but little or no labeling at interchromatin granules or coiled bodies (Fakan et al., 1976, 1984; Fakan and Nobis, 1978; Fakan and Puvion, 1980; Spector, 1990; Fakan, 1994). Moreover, recent studies using DNA, intron, and splice-junction probes, which specifically recognize pre-mRNA but not spliced mRNA, indicated that splicing occurs cotranscriptionally at the location of the gene and neither spliced mRNA nor introns were detected at interchromatin granules or coiled bodies (Zhang
splicing regulators with the ones of the above described nuclear matrix antigens. Considering the two lines of evidence it seems reasonable to speculate that SR proteins like tra and su(*w*) might exert their regulatory role in part by coupling splicing complexes to the nuclear matrix.

Cell division poses several interesting problems—not only the chromosomes but also the nuclear matrix and its associated structures have to be duplicated and distributed to the two daughter cells—and involves a transient and at least partial disassembly of these otherwise insoluble structures during mitosis. It has been shown before that the speckled compartment breaks up during mitosis and splicing factors are dispersed throughout the cytoplasm (Spector and Smith, 1986). Recently, a serine kinase (SPRK1) was identified, which phosphorylates serine residues of serine/arginine-rich domain of SR proteins and regulates their subnuclear localization.

Phosphorylation causes disassembly of the speckled compartment upon entry into mitosis and inhibits splicing in vitro indicating an important role of SRPK1 in the regulation of RNA processing and in the control of the dynamic organization of nuclear structures (Gui et al., 1994a,b; Spector, 1994). Another example of subnuclear protein targeting is p80-collin. The first 102 amino acids of p80-collin are necessary and sufficient for targeting to specific organelles, which appear to be the amphibian homolog of mammalian coiled bodies (Wu et al., 1994).

Finally, there also seems to be a potential role for pRb at snRNP-containing nuclear structures. Several independent studies established that the early G1, hypophosphorylated form of the retinoblastoma protein (pRb) is tightly bound to the insoluble nuclear framework (Mittnacht and Weinberg, 1991; Templeton et al., 1991; Templeton, 1992; Mancini et al., 1994; Mittnacht et al., 1994). Hypophosphorylated pRb binds viral oncogenes and, for this reason, is thought to be the form active in growth suppression (Decaprio et al., 1988; Ludlow et al., 1989; Imai et al., 1991; Templeton et al., 1991). In concert with these results, microinjection of the serine/threonine protein phosphatases types 1 (PP1) and 2A (PP2A) leads to increased nuclear affinity of pRb and to cell cycle arrest (Alberts et al., 1993). Many cellular proteins have been identified that interact with pRb (Riley et al., 1994; Wang et al., 1994), but they are easily extractable soluble proteins and therefore not likely to play a role in “docking” to the nuclear matrix. Recently, in a screen for pRb-interacting cellular proteins using the yeast two-hybrid system and the N-terminal 300 amino acids of pRb as a bait, one gene was isolated encoding a novel nuclear matrix protein, p84, which localizes to subnuclear regions associated with RNA processing (Durfee et al., 1994). Also, p84 interacts via its carboxy-terminal half with the hypophosphorylated form of pRb that is present during G1 and, by immunofluorescence and confocal microscopy of preextracted cells, p84 intranuclear
speckles colocalize extensively (even though not totally) with pRb, Sm antigens, and another recently identified nuclear matrix protein B1C8 (Durfée et al., 1994). These results suggest a connection between the nuclear matrix (including B1C8), pRb, and RNA metabolism. Genetic evidence pointing to a functional role of the N-terminus of pRb in tumor suppression has recently been obtained from a mutation isolated from a retinoblastoma tumor, which is predicted to remove only exon 4, coding for amino acids 126-166 (Dryja et al., 1993; Hogg et al., 1993). The association of pRb with both matrix-bound and soluble components of the nucleus might represent a novel mechanism by which pRb exerts its growth-suppressive activity under physiological conditions and it would be interesting to examine whether pRb or its associated protein p84 have a direct role in the organization and regulation of RNA metabolism. In summary, all of these recently discovered interactions surrounding the nuclear matrix and the splicing machinery are still at a very speculative stage but they open exciting new avenues of research.

IV. DNA Replication Foci in Mammalian Nuclei

A. Visualization of DNA Replication Foci and Changes during S-Phase

The successful reproduction of viral replication in the test tube with isolated soluble components (Li and Kelly, 1984; Stillman and Gluzman, 1985; Wobbe et al., 1985) suggested that DNA replication would occur evenly distributed throughout the nucleus. However, over the last two decades several methods for the visualization of cellular sites of DNA synthesis have been developed that clearly show that viral and cellular DNA replication occurs in discrete foci. The first methods were based on the incorporation of radioactively labeled [H]thymidine, which can be detected by autoradiography (Huberman et al., 1973; Fakan and Hancock, 1974). More recently, a number of nonradioactive labeling techniques combined with immunofluorescence detection were developed that utilize thymidine analogs like 5-bromo-2'-deoxyuridine (BrdU) (Gratzner, 1982; Nakamura et al., 1986; van Dierendonck et al., 1989; Mazzotti et al., 1990; Fox et al., 1991) or iododeoxyuridine (IdU) and chlorodeoxyuridine (CldU) (Atem et al., 1992; Manders et al., 1992). Alternatively, biotin-11-dUTP labeling was used with permeabilized cells (Banfalvi et al., 1989; Mills et al., 1989; Nakayasu and Berezney, 1989). Most recently, fluorescently labeled nucleotides also were used for direct visualization of DNA synthesis (Hassan and Cook, 1993). It is reassuring that all these different approaches lead to the same basic conclusion that DNA replication occurs in discrete nuclear foci and the pattern of these replication foci changes in a characteristic manner throughout S-phase (for an extensive review on this subject see chapter by Berezney et al.).

The fact that the pattern of replication foci changes throughout S-phase can be exploited to address basic questions concerning the regulation of DNA replication. Yanishevsky and Prescott (1978) could show in cell fusion experiments that late S-phase cells induce early S-phase DNA labeling patterns in G1 nuclei. These results indicate a preexisting regulatory mechanism in G1 nuclei that determines the spatial and temporal order of DNA replication and is prematurely triggered by a diffusible factor from S-phase cells. Using a retrodifferentiation cellular system, Cardoso et al. (1993) could show that transient expression of SV40 large T antigen induces terminally differentiated myotubes to reenter S-phase with the exact same patterns of replication foci as, e.g., cycling myoblasts, indicating that this regulatory mechanism also persists in terminally differentiated cells that are permanently withdrawn from the cell cycle. Moreover, in these induced myotubes, neighboring nuclei, which share the same cytoplasm and should, therefore, have equal access to cytoplasmic factors, show very different replication patterns ranging from typical early to late S-phase patterns and include even G1 nuclei with no detectable DNA synthesis. These results suggest that the nucleus represents an autonomous unit and may at least in part control the timing of S-phase (Cardoso et al., 1993).

B. Proteins Present at Replication Foci

The first step toward an understanding of molecular events involved in the exact duplication of DNA molecules was done in the late 1950s, when the group of Arthur Kornberg isolated from extracts of Escherichia coli an enzyme that could synthesize DNA and was accordingly named DNA polymerase. This first enzyme had, however, only a miniscule activity, which was only little over background levels and was clearly orders of magnitude below rates observed in vivo. During the following decades other forms of DNA polymerase and, most importantly, other subunits and accessory proteins were isolated that improved the rate and fidelity of DNA replication in vitro and gradually helped to narrow the gap between replication in vitro and in vivo. This development is also reflected in the terminology, starting with DNA replication enzymes over replication complexes to replication machinery. This inflation in the terminology recently culminated in the term "replication factory" coined by the group of Peter Cook (Hozák et al., 1993).
While research in proarytes has progressed to a point where an understanding of the basic process and its control mechanisms appears to be within reach, there is no complete in vitro system for cellular DNA replication in mammals available and it is still debated what an origin of chromosomal DNA replication is and whether they exist in mammals to begin with. Even with all the proteins identified and characterized by genetic and biochemical approaches it remains impossible to imagine how they can work together to achieve as precise and well-coordinated a round of DNA replication as observed in mammalian cells. It remains to be explained how \( \times 10^9 \) base pairs per human cell get replicated with extremely high fidelity in a matter of hours involving up to 60,000 origins of DNA replication, which are activated at specific times during S-phase and yet each segment of the 46 chromosomes is replicated once and only once per cell cycle. A first clue about how to close the gap between in vitro replication results and observations in mammalian cells came from early experiments showing an association of newly synthesized DNA with the nuclear matrix (Berezney and Coffey, 1975). Later biochemical fractionation experiments showed that DNA polymerase activity was specifically associated with the insoluble fraction of the nucleus (Smith and Berezney, 1980; Noguchi et al., 1983; Jackson and Cook, 1986; Tubo and Berezney, 1987). Over the past few years it has become clear that the association of DNA replication with insoluble nuclear structures and its organization into megacomplexes of up to a micrometer in size may provide the higher order structure, which make the precision, efficiency, and complex coordination of mammalian DNA replication possible. The role of the nuclear matrix in the organization and regulation of DNA replication is discussed in the chapter by Berezney et al. This section focuses on the view of cellular DNA replication through the microscope. As discussed in Section IV.A, DNA replication occurs at discrete foci in mammalian nuclei and can easily be visualized by pulse labeling with either biotinylated dUTP or BrdU. It is therefore possible with standard immunofluorescence microscopy to determine whether a given protein is present at sites of DNA replication in vivo (Cardoso and Leonhardt, 1995). The first protein to be identified at replication foci was PCNA, which shows dispersed nucleoplasmic distribution throughout the cell cycle and only in S-phase relocates to replication foci and then shows the same characteristic and changing replication patterns (Bravo and MacDonald-Bravo, 1985, 1987; Celis and Celis, 1985, 1986). Since then a number of other proteins directly or indirectly involved in DNA replication were found to be specifically localized at these replication foci (see Table III).

In addition to PCNA DNA polymerase α (Hozák et al., 1993) and RPA70 (Cardoso et al., 1993) also were shown to be localized at sites of cellular DNA replication. RPA70 is the 70-kDa subunit of the single-stranded DNA

<table>
<thead>
<tr>
<th>Protein</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCNA (proliferating cell nuclear antigen)</td>
<td>Bravo and MacDonald-Bravo (1987)</td>
</tr>
<tr>
<td>DNA methyltransferase</td>
<td>Leonhardt et al. (1992)</td>
</tr>
<tr>
<td>DNA polymerase α</td>
<td>Hozák et al. (1993)</td>
</tr>
<tr>
<td>Replication factor A, 70-kDa subunit (RPA70)</td>
<td>Cardoso et al. (1993)</td>
</tr>
<tr>
<td>cdk2 (cyclin-dependent kinase)</td>
<td>Cardoso et al. (1993)</td>
</tr>
<tr>
<td>Cyclin A</td>
<td>Cardoso et al. (1993)</td>
</tr>
<tr>
<td>DNA ligase I</td>
<td>Cardoso et al. (1995)</td>
</tr>
</tbody>
</table>

binding replication factor A. Interestingly, the 34-kDa subunit, RPA34, was not detected at replication foci but instead appeared to be associated with condensed chromosomes during mitosis. These results may indicate a transient interaction of RPA34 or might be due to a specific distortion or masking of the epitope at replication foci, since a monoclonal antibody was used (Cardoso et al., 1993). Most recently, DNA ligase I, which is involved in joining Okazaki fragments during lagging strand synthesis, was also shown to be localized at nuclear replication foci (Cardoso et al., 1995). In addition to these replication factors several other proteins not directly involved in DNA replication like DNA methyltransferase also were found at replication foci (Leonhardt et al., 1992). DNA methyltransferase adds methyl groups to the newly synthesized DNA strand at hemimethylated CpG sites after DNA replication and thus maintains the cellular DNA methylation pattern. In the course of experiments on the control of DNA replication the cell cycle regulators cyclin A and cdk2 also were identified at sites of cellular DNA replication (Cardoso et al., 1993; for a discussion see Section IV.D). This localization of cyclin A during S-phase was also shown in HeLa cells by means of immunoelectron microscopy (Sobczak-Thepot et al., 1993). Finally, even lamin B was shown to be associated with sites of DNA replication during mid to late S-phase, suggesting a potential role in the organization of replicating chromatin (Moir et al., 1994). These analyses screening for proteins specifically localized at replication foci have already provided several interesting clues about the organization and regulation of nuclear replication foci and further insights are to be expected for the coming years as more proteins are screened.

Equally interesting are questions concerning the formation of these replication foci, meaning which proteins are present, in what phase of the cell cycle, and in what order do they bind. First evidence for the possible existence of such prerelation foci or centers was provided by microinjec-
tion experiments showing PCNA containing foci about 15 min before DNA synthesis was detected (Kill et al., 1991). Similar conclusions were also derived from in vitro experiments with Xenopus egg extracts showing RPA70 containing foci during G1-phase (Adachi and Laemmli, 1992). The dynamics of the nuclear organization also becomes evident upon viral infection, which not only forces reprogramming of the transcriptional and translational machinery, but also causes an extensive reorganization of nuclear structures (Quinlan et al., 1984; de Bruyn Kops and Knipe, 1988; Walton et al., 1989; for a detailed discussion of “nuclear matrix and virus function” see chapter by Deppert). A number of cellular proteins were reported to be recruited to sites of viral DNA replication including replication protein A (RPA), PCNA, retinoblastoma protein, p53, DNA ligase I, and DNA polymerase α (Wilcock and Lane, 1991). Interestingly, the retinoblastoma protein was not detected at sites of cellular DNA replication (H. Leonhardt and M. C. Cardoso, unpublished results). It will be interesting to compare the structure of viral and cellular replication foci. The changes induced by viral infection may reveal some of the organizing principles of mammalian nuclei, much like it has been used to elucidate cell cycle regulation.

C. Targeting of Proteins to Replication Foci

1. DNA Methyltransferase

In Section IV.B a number of proteins were described that are specifically localized at nuclear replication foci (see Table III). In the absence of dividing membranes, which could explain this compartmentalization, it seems reasonable to assume that this localization is caused by binding to some sort of insoluble structures. This hypothesis was first put to a test in the case of the mammalian DNA methyltransferase (DNA MTase), which is localized at sites of DNA replication (Leonhardt et al., 1992). Mammalian DNA is methylated in both strands at about 50–80% of all CpG dinucleotides. After DNA replication, the newly synthesized DNA strand is at first unmethylated resulting in hemimethylated CpG sites, which are then specifically recognized and methylated by DNA MTase, precisely propagating the previous DNA methylation pattern. Already partial inactivation of DNA MTase causes embryonic lethality demonstrating the importance of the precise maintenance of methylation patterns for differentiation and mammalian development in general (Li et al., 1992, 1993).

The mammalian DNA MTase has a C-terminal, catalytic domain of about 500 amino acids and an N-terminal regulatory domain of about 1000 amino acids (Bestor et al., 1988; Leonhardt and Bestor, 1993). The enzyme has a strong preference for hemimethylated CpG sites, which is consistent with a role in the maintenance of methylation patterns (Gruenbaum et al., 1982; Bestor and Ingram, 1983). The specific localization of DNA MTase at replication foci could, therefore, be caused by high substrate concentration, since hemimethylated sites are generated by DNA replication and should be most abundant at these foci. Alternatively, the specific localization could be achieved by a targeting sequence that concentrates the enzyme at the replication foci by some sort of docking process. To test this alternative, various parts of the DNA MTase gene were fused to the totally unrelated β-galactosidase gene from E. coli and expressed in mouse fibroblast cells. The distribution of the fusion protein can be determined with antibodies recognizing an epitope in the β-galactosidase part and can be compared to the distribution of the endogenous enzyme or—for that matter—any other cellular components that can be visualized by fluorescent microscopy techniques. With this experimental approach a targeting sequence was identified, which is necessary and sufficient for targeting to replication foci and by itself causes the localization of a protein as unrelated as the bacterial β-galactosidase to replication foci (Fig. 1A–1C, color plate 12). For comparison, targeting to the speckled compartment using an RS domain fusion is shown in Fig. 1D–1F, color plate 12). The targeting sequence of the DNA MTase is located in the N-terminal, regulatory domain and is dispensable for enzyme activity in vitro (Leonhardt et al., 1992; Fig. 2).

The DNA MTase targeting sequence is surprisingly hydrophobic and large (about 250 amino acids) suggesting that most of it has a function in the globular folding of the enzyme placing short and discontinuous stretches of amino acids into the right three-dimensional configuration to form the binding interface. Searches of protein and DNA databases failed to identify other proteins with sequence similarity suggesting that the DNA MTase targeting sequence is rather unique and not a member of a general new type of protein motif. The definition of this targeting sequence should now make it possible to identify the component(s) of the replication foci that DNA MTase binds to and should contribute to our understanding of the architecture of replication factories. It is tempting to speculate that the targeting sequence plays an important role in the coordination of DNA replication and methylation in vivo by placing the DNA MTase at the right position in the replication factory. The efficient recognition and methylation of hemimethylated sites would then not simply rely on diffusion and random encounters but would be achieved by one processive replication-methylation machinery or assembly line (see model, Fig. 3).

2. DNA Ligase I

The identification of a targeting sequence in case of the mammalian DNA MTase raised the question whether other proteins also use a similar princi-
ple for localization at replication foci. A good candidate was the DNA ligase I, which is involved in joining Okazaki fragments during lagging-strand synthesis and in excision-repair (Barnes et al., 1992). The human DNA ligase I has 919 amino acids and the catalytic center is around the active site lysine residue at position 568 (Barnes et al., 1990). The enzyme has, like DNA MTase, a protease-sensitive N-terminal domain of 249 amino acids, which has a negative regulatory function and is dispensable for enzyme activity in vitro and for complementation of yeast and bacterial ligase mutants (Kodama et al., 1991; see Fig. 2). DNA ligase is activated at the onset of S-phase by phosphorylation of the N-terminal regulatory domain (Prigent et al., 1992). Biochemical fractionation experiments showed that DNA ligase I is part of a large protein complex of replication factors (Malkas et al., 1990) and is localized at replication foci in vivo (Cardoso et al., 1995). To study the basis of this subnuclear localization pattern, the same approach was used that had led to the identification of the DNA MTase targeting sequence (Leonhardt et al., 1992; see Section IV,C,1). The analysis of a set of epitope-tagged deletion mutants identified a bipartite targeting sequence within the N-terminal, regulatory domain of the human DNA ligase I (Cardoso et al., 1995; see also Fig. 2). This targeting sequence is in functional terms very similar to that of the DNA MTase, as it is dispensable for enzyme activity in vitro, can by itself direct unrelated proteins like β-galactosidase to replication foci, and is necessary and sufficient for targeting. A direct comparison of both sequences, however, failed to detect any similarities and in fact they are quite different. The DNA ligase I targeting sequence is very hydrophilic, while the DNA MTase sequence is on the contrary very hydrophobic. These results suggest that the two enzymes bind to different components of the replication factory, which makes sense in light of their different enzymatic activities. A comparison with other DNA ligases showed that the targeting sequence is highly conserved in mammals (mouse and human) but is absent in lower eukaryotes like yeast, suggesting that targeting is a rather recent but important development in evolution. It is noteworthy that both enzymes are involved in cellular processes where efficiency is critical, DNA MTase has to methylate
in the order of 10^6 hemimethylated sites and DNA ligase has to join about 10^8-10^10 Okazaki fragments per cell cycle (assuming that Okazaki fragments are 200 to 1000 bp in size). In both cases it is crucial for the viability of the cell that these processes are totally completed; any omissions might have serious consequences. The targeting of DNA ligase I to replication foci might be a mechanism used by mammalian cells to ensure this efficiency (see model, Fig. 3).

3. Other Candidates

In addition to proteins already identified at replication foci (see Table III), there are many more proteins involved in SV40 replication that are, therefore, also expected to be involved in cellular DNA replication. Complete SV40 replication can be obtained with SV40 large T antigen, replication protein A, DNA topoisomerase I and II, DNA polymerase α-primease, replication factor C, PCNA, DNA polymerase δ, maturation factor I, and DNA ligase I (Waga et al., 1994). Most—if not all—of these proteins together with still unidentified factors are expected to be localized at replication foci. It should, therefore, be interesting to determine what other factors are localized at replication foci and whether distinct targeting sequences can be identified using the same approach as described above. Obviously, not all of these proteins are expected to have distinct targeting sequence. Especially smaller proteins, e.g., cdk2, which are altogether as big (or small) as the targeting sequence of the DNA MTase, are not likely to have a separate targeting sequence. Nonetheless, the analysis of epitope-tagged deletion mutants should identify regions that are necessary for localization at replication foci, even if they are part of the catalytic domain. It would then be a semantic question whether it can still be called targeting sequence. In any event, the identification of such sequences can be used to screen for interacting proteins, which should lead to the identification of other known and unknown proteins. These results would then outline a network of interactions and should definitely help to elucidate the architecture of replication foci in vivo.

Finally, an interesting candidate for having a targeting sequence is DNA topoisomerase I. This enzyme has, like DNA ligase I and DNA MTase, a protease-sensitive N-terminal domain, which is dispensable for enzyme activity in vitro (Bjornst and Wang, 1987; D'Arpa et al., 1988; Alsnor et al., 1992) and is required for complete SV40 DNA replication (Yang et al., 1987).

D. A Link between Cell Cycle Regulation and DNA Replication

Our understanding of cell cycle regulation is continuously challenged by the ever-increasing number of new cyclins, cyclin-dependent kinases, and other regulatory proteins, such as the recently identified family of cyclin-cdk inhibitors. The most critical step in cell cycle progression is the control of S-phase entry. Biochemical analyses showed that the phosphorylation of RPA34 is a key event at the initiation of DNA replication (Din et al., 1990). The problem is that, as is the case with other putative substrates for phosphorylation by cyclin–cdk complexes, RPA34 can be phosphorylated in vitro by several distinct complexes including cyclins A, B, or E and cdc2 or cdk2 (Fotedar and Roberts, 1991; Dutta and Stillman, 1992; Elledge et al., 1992). Do these different cyclin–cdk complexes have overlapping substrate specificity in vivo? Or, alternatively, do they carry out specific tasks in vivo that are nonredundant? If so, how can the physiologically meaningful complex be sorted out?

Overexpression and inhibition experiments suggest that different cyclins and cdk5 target different substrates and act at distinct times during the cell cycle. Immunofluorescence analysis in mammalian cells induced to reenter S-phase showed specific colocalization of cyclin A and cdk2, but not of cyclin B and cdc2, with subnuclear sites of ongoing DNA synthesis (Fig. 4; color plate 13) (Cardoso et al., 1993). Similar results were obtained for cyclin A by immuno-electron microscopy (Sobczak-Thep et al., 1993). Colocalization was observed with DNA replication patterns typical of early, mid, and late S-phase suggesting that cyclin A–cdk2 acts throughout S-phase and not only at the beginning. This cell cycle-dependent redistribution of cyclin A and cdk2 perfectly matches the one of replication protein PCNA (Cardoso et al., 1993). Localization of cyclin–cdk complexes to specific domains might explain how they can have overlapping substrate specificity in vitro but carry out specific and nonredundant tasks in vivo. Obviously, colocalization by itself does not prove any involvement but it forms a consistent picture, together with biochemical data, showing that cyclin A and cdk2 are at the right place at the right time to “trigger” DNA replication. These results demonstrate how the functional organization of the mammalian nucleus can provide important clues about protein function in vivo and strongly suggest that cyclin A–cdk2 is the “S-phase-promoting factor” or “replication factor S” postulated by Roberts and D’Urso (1988; Reed, 1991), providing a direct link between cell cycle regulation and DNA replication.

V. Insoluble Solutions Derived from Functional Domains in the Nucleus

The most direct way to regulate an enzyme is through its abundance by either synthesis or degradation. Alternatively, the subcellular localization
of enzymes may be used to control their biological activity. A very obvious case is the cytoplasmic versus nuclear localization and well-known examples for this type of regulation are NF-κB, Rel, and dorsal, which were shown to be regulated at the level of nuclear uptake (Schmitz et al., 1991). The functional organization of the nucleus represents another potential level for regulating the activity of nuclear proteins. The activity of proteins may be positively regulated by targeting a protein to the appropriate subnuclear compartment containing the respective substrate. The local enzyme and substrate concentration is thereby dramatically increased and consequently also the rate of catalysis. The increase in velocity due to protein targeting and integration into complex protein machineries can only be estimated at this moment. On the other hand, preventing targeting of proteins to a particular subnuclear compartment may be as effective as preventing its nuclear uptake. Interestingly, several smaller replication foci appeared to have lower or no detectable levels of DNA MTase, which could have simple technical reasons or might point to a replicon-specific regulation of DNA methylation at the level of targeting (Leonhardt et al., 1992). The identification of these intranuclear targeting sequences makes it possible to screen for interacting factors and post-translational modification which may control protein targeting within the nucleus. The recently discovered SRPK1 kinase phosphorylating the RS domain of SR proteins and thus causing the break up of the speckled compartment (Gui et al., 1994a; see Section III) may fall into this category. A particular subnuclear site may also constitute a "protective environment" shielding proteins from degradation or modifying factors, which would lead to their inactivation. Finally, there are additional subnuclear compartments of still unknown function like the nuclear dots and PIKA (Ascoli and Maul, 1991; Saunders et al., 1991, respectively), which contribute to the complex organization of mammalian nuclei.

Almost by definition, the functional organization provides valuable clues about protein function and, hence, may help with model building. Since, e.g., several replication proteins were found to be specifically localized to nuclear replication foci (Table III) it seems reasonable to expect that other proteins identified at these foci might be involved in some aspects of DNA replication, just as the nuclear localization of proteins is usually taken as an indication for a role in the nucleus. One example for this approach is the identification of cyclin A and cdk2 at replication foci providing a link between cell cycle regulation and DNA replication (Cardoso et al., 1993; Section IV,D). Obviously, localization by itself is not a proof and proteins caught "in flagranti" might have actually been "innocent bystanders," but these localization studies complement biochemical and genetic data and can provide compelling evidence about protein function in vivo. The integration of various enzyme activities into complex and insoluble macromo-

molecular machineries is likely to provide the efficiency, specificity, and coordination of biological processes in vivo and is discussed throughout this chapter (see also Fig. 3). These higher order structures, however, defy the principle of traditional biochemical analysis in two ways. First, biochemistry mostly deals with the soluble fraction of the cell and hence misses all insoluble structures. Second, the traditional biochemical approach is to purify cellular components and characterize them in vitro, but this takes them out of their "natural environment" or context and may identify affinities to other cellular components that they never encounter in vivo.

For many years, research on the nuclear matrix and its associated components has been hampered by technical difficulties (Cook, 1988). It is now that a number of experimental approaches are in place to explore the functional organization of the nucleus and to determine its role in the regulation of biological processes. On the biochemical side, methods have been developed to prepare nuclear matrix samples for functional assays and matrix components are being identified by monoclonal antibodies, cloned and sequenced. Moreover, the association of nuclear components with functional domains can be studied by immunofluorescence microscopy and deletion analyses. Most recently, a method for the direct visualization of the functional organization of the nucleus also was developed. The green fluorescent protein (GFP; Chalfie et al., 1994) was fused with the targeting sequence of the DNA MTase and expressed in mammalian cells. The fusion protein, which is by itself fluorescent, is targeted to nuclear replication foci and can be followed in living cells allowing the direct visualization of nuclear structures in vivo and to follow its dynamic reorganization during the cell cycle and in response to external stimuli (Leonhardt et al., 1995). It is now feasible to study the functional organization of the nucleus in real time, altogether avoiding potential fixation and staining artifacts. Furthermore, this approach can easily be adapted for the analysis of other functional domains by replacing the targeting sequence with any of the above described nucleolar targeting signals or the RS domain of, e.g., the splicing regulator tra.

VI. Concluding Remarks

After decades of breathtaking progress at all fronts of science and technology providing a plethora of data on an ever-increasing number of molecules, we still seem to be far from really understanding living organisms. The reductive approach of science dissecting complex problems into smaller and more manageable questions was extremely successful and led to the discovery and characterization of RNA, DNA, and proteins, the smallest
building blocks of living organisms, but something still seems to be missing. At the biochemical and genetic level humans, e.g., compare to chimpanzees as, e.g., a fox to a dog, but we write poetry or fly to the moon and they do not. Likewise, putting all the known and unknown components of the nucleus together in a test tube is far from working like a nucleus in vivo and is likely to end up in a mess. The answer for both dilemmas seems to lie in higher order structures and the challenge for the next century is to piece all the individual building blocks together expecting that the cell and the organism is much more than just the sum of the properties of its components.

At the cellular level, one of the first steps in this direction is to examine the three-dimensional architecture of the nucleus and to determine the subnuclear localization of the cellular components during cell cycle and differentiation. Just as at present protein expression and post-translational modification profiles are established to understand their regulatory effects during, e.g., the cell cycle, it will be necessary to also take into account the subcellular localization and hence limited spatial availability of cellular components. In other words, the subcellular and subnuclear localization of proteins represents yet another variable in the overall regulatory networks. The functional organization of the nucleus changes the equations in enzyme kinetics; protein targeting and the assembly into huge protein machineries or factories appears to be one way to integrate complex biochemical processes and to achieve efficiency and coordination. The latter is by no means a biological luxury but rather can make all the difference in such “competitive” environments as living cells and organisms.

One of the most intriguing questions in this context is whether the individual components contain all the information necessary for the assembly of these higher order structures. Ribosomes, e.g., can be assembled in vitro, if the individual components are added in the correct order using appropriate conditions. But still a mixture of ribosomal RNA and proteins will not spontaneously assemble efficiently in vitro and even less so if all the other cellular proteins are present. This problem seems to be solved in vivo by the nucleolus, which is a subnuclear compartment dedicated to the assembly of ribosomal subunits where all the ribosomal components and participating enzymes are concentrated while other proteins are excluded. Likewise, it would make sense that interchromatin granules and coiled bodies are domains dedicated to the assembly of splicing particles (spliceosomes). Moreover, the identification of distinct targeting sequences in a variety of molecules, ranging from replication proteins to splicing regulators and nuclear proteins, suggests that the information for the higher order structures seems to be built-in at least in some cases. On the other hand, a homogenate of all cellular components does not spontaneously assemble into living cells. It is tempting to speculate that these higher order structures are duplicated and passed on to daughter cells just like chromosomes are and possibly by association with the chromosomes themselves. We would, furthermore, like to speculate that these higher order structures evolved together with the genome in an interdependent fashion, one dependent on the other, meaning the structures cannot be maintained and propagated without the genome and vice versa.

Almost two centuries ago Goethe formulated Faust’s thirst for knowledge as a desire to understand what ultimately holds the world together (“... Daß ich erkenne, was die Welt im Innersten zusammenhält...”). Besides the metaphysical component of this question, the problem could be rephrased for today’s scientists as how “simple” molecules are assembled to form higher order structures and complex systems like nuclei, cells, organs, organisms, ecosystems, and galaxies.

References


TARGETING TO NUCLEAR DOMAINS


PLATE 1 (He et al., Fig. 3). Cartoon showing possible structural roles of NuMA in the interphase nuclear matrix (A) and the mitotic spindle (B).

PLATE 2 (Moreno Díaz de la Espina, Fig. 4). Stereopair obtained after serial sectioning of the nuclear matrix shown in Fig. 5. To produce the stereopairs, the outlines of the nuclear matrix components were drawn onto transparent sheets (one for each micrograph), their (X, Y) coordinates were recorded with a digitizer and memorized in the computer. The data were computerized with a program that includes: calculation and plotting of “smooth” profiles from coordinates, rotation of the profiles at 5°, and re-plotting of new profiles from the rotated coordinates. Both originals and rotated profiles constitute a stereopair. The lamina is represented in black, the profiles of the nucleolar matrix are red, and those of the patches of the internal matrix are green. Only the profiles of the internal matrix around the nucleolar matrix have been represented in the stereopair.

PLATE 3 (Moreno Díaz de la Espina, Fig. 14). Localization of fibrillarin in the nucleolar matrix by confocal microscopy after immunolabeling with the S4 serum. Sixteen consecutive optical sections (0.4 μm) through a nuclear matrix including the nucleolar area shown after deblurring to eliminate the out-of-focus fluorescence. The labeling is not randomly distributed but accumulates on discrete domains in the central part of the nucleolar matrix (large arrow). The boundary of the nuclear matrix is outlined by its faint fluorescence only visible in some middle sections (small arrows). Bars B–E=10μm.

PLATE 4 (Panté and Aebi, Fig. 1). Schematic representation of the different components of the nuclear envelope (NE) together with the molecular trafficking occurring across the NE through the nuclear pore complexes (NPCs). The NE consists of an inner and an outer nuclear membrane enclosing the perinuclear space. The outer nuclear membrane is continuous with the endoplasmic reticulum (ER), so that the perinuclear space of the NE is continuous with the lumen of the ER. The inner nuclear membrane is lined by the nuclear lamina, a near-tetragonal meshwork made of intermediate filament-like proteins called nuclear lamins. The NPCs are interposed at irregular intervals between the inner and outer nuclear membrane. Bidirectional molecular trafficking between the nucleus and the cytoplasm occurs through the NPCs which allow passive diffusion of ions and small molecules and energy-dependent transport of nuclear proteins, RNAs, and RNP.

PLATE 5 (Panté and Aebi, Fig. 4). Surface renderings of the 52-MDa basic framework reconstructed from negatively stained NPCs released from the NE upon detergent treatment. (a) Slightly tilted view of the basic framework of the NPC which consists of eight multidomain spokes and exhibits strong 822 symmetry thus indicating that it is built of two identical halves relative to the central plane of the nuclear envelope. Note that due to its irreproducible appearance, the central plug or channel complex has been omitted in this reconstruction. When the pore membrane is positioned in this map of the basic framework of the NPC, eight ~10-nm-diameter peripheral channels are created between two adjacent spokes and the pore membrane border at a radius of ~40 nm. (b) Three different side views of one multidomain spoke cut out from the basic framework of the NPC as shown in (a). Each half-spoke is built from four distinct morphological domains termed annular (a), column (c), ring (r), and luminal (l) domain. This figure has been adapted from Hinshaw et al. (1992).

PLATE 6 (Panté and Aebi, Fig. 7). Schematic representation of a consensus model of the membrane-bound NPC. Its major structural components include the basic framework (i.e., the spoke complex as shown in Fig. 4a), the central plug or channel complex, the cytoplasmic and nuclear rings, and the cytoplasmic filaments and nuclear basket. The 52-MDa basic framework of the NPC has been adapted from the 3-D reconstruction of negatively stained detergent-released NPCs (Hinshaw et al., 1992). The cytoplasmic filaments and nuclear basket have been modeled based on EM data obtained by Ris (1991), Jarnik and Aebi (1991), and Goldberg and Allen (1992) (see also Fig. 5). In this consensus model of the NPC, we have also pictured a cytoplasmic and nuclear ring in addition to the two tenuous rings defined by the ring domain of the spoke complex (see Fig. 4a). The central plug or channel complex has been modeled as a transparent ellipsoidal particle to emphasize the fact that its definite structure remains to be determined.

PLATE 7 (Panté and Aebi, Fig. 9). Schematic diagram of the nuclear import pathway of proteins through the NPC. In the first step, the nuclear protein to be transported associates with an NLS receptor, this complex then
PLATE 8 (Basost et al., Fig. 2). Diagram of the domain architecture of peripheral membrane proteins of vertebrate NPCs deduced from their predicted amino acid sequences. Depending on the presence of N-linked N-acetylgalactosamine (O-GlcNAc) residues, two groups of peripheral membrane proteins of the NPC have been distinguished. (a) Members of the O-linked glycoprotein family which in addition to O-GlcNAc contain multiple copies of a more or less degenerate pentapeptide motif XFXPG. POM121 (shown in Fig. 3) may also be considered a member of this family. Other features of individual proteins are as follows: p62 contains a COOH-terminal coiled-coil domain, and NUP153 has four zinc finger motifs. (b) N-terminal domain containing neither O-GlcNAc nor repetitive motif sequences. The amino acid sequence of the three members of this group, NUP107, NUP155, and Nup265, is quite unrelated. In the case of Tprp265, its major feature is an 1800-residue-long coiled-coil domain beginning close to the NH2-terminus. NUP107 is distinguished by a COOH-terminal lectin zipper. For additional information concerning these proteins, see Table I and references therein.

PLATE 9 (Basost et al., Fig. 3). Diagram of the domain architecture and membrane topology of both vertebral and yeast integral membrane proteins of the NPC. The orientations shown are based on deduced amino acid sequences as well as upon direct experimental evidence (see text). Three transmembrane NPC glycoproteins, gp210, POM121, and the yeast POM152p, have been cloned so far and sequenced. The translated sequences all reveal a distinct stretch of hydrophobic residues that is predicted to traverse the NPC membrane. Gp210 contains a large NH2-terminal domain residing in the perinuclear space (PNS), the luminal of the nuclear envelope, and a small extraluminal COOH-terminal tail. In contrast, POM121 exhibits a short NH2-terminal tail residing in the PNS, and a large extraluminal COOH-terminal domain. The latter contains a number of FXPG pentapeptide motifs clustered near the COOH terminus. Like gp210, most of the mass of yeast POM152p is predicted to reside in the PNS. In this case, however, the orientation of the polypeptide is reversed such that it is the COOH terminus that lies in the PNS. The luminal domains of both gp210 and POM152p each possess consensus sites for N-glycosylation, close to the membrane-spanning segment, at least some of which are utilized (Greber et al., 1990; Wozniak et al., 1994). In addition, the amino acid sequence of POM152p contains eight repetitive segments, within the luminal domain, each 24 residues long, with the consensus sequence C-G-V-Y-L-S-P-F-Y. The functions of these repeat motifs are unknown. The outer of each of these proteins, the extraluminal domain is thought to be engaged in the inner region of the NPC.

PLATE 10 (Basost et al., Fig. 4). Diagram of the domain architecture of yeast NPC proteins deduced from their amino acid sequences. Depending on the occurrence of highly repeated motifs in their amino acid sequences, three groups of yeast NPC proteins have been distinguished. (a) The XFXPG family, Nsp11p, Nup1p, and Nup2p, members of which exhibit numerous copies of a more or less degenerate pentapeptide motif XFXPG clustered in the central part of each molecule. (b) The GLFG family, Nup49p, Nup57p (not shown), Nup100p, Nup116p, and Nup145p, members of which contain several copies of a more or less degenerate tetrapeptide GLFG within their NH2-terminal domains. Three members of this family—Nup100p, Nup116p, and Nup145p—are contained in additional related domains including an RNA-binding motif. (c) The third group by default, includes the yeast nuclear import component NIC46p, SRP1p (not shown), and NUP133p (not shown). None of these proteins exhibit any obvious interrelationships and contain no repetitive sequence elements in the vein of XFXPG or GLFG referred to above. NIC46p may be recovered in a complex with Nsp11p, Nup57p, and Nup49p (Grand et al., 1993). Interactions between these four proteins most likely involves the coiled-coil helical domains apparent in each. For more information about these proteins, see Table 1 and references therein.

PLATE 11 (Basost et al., Fig. 5). Diagram summarizing the immunolocalization of characterized NPC protein epitopes within the 3-D architecture of the consensus model of the NPC. The major structural components of the NPC include the basic framework (i.e., the spoke complex), the central plug or channel complex, the cytoplasmic and nuclear rings, and the cytoplasmic filaments and nuclear basket. The basic framework of the NPC has been adapted from a conical tilt reconstruction of negatively stained detergent-released NPCs (Hinshaw et al., 1992). The cytoplasmic filaments and nuclear basket have been modeled based on EM data obtained by Jarnik and Aebi (1991), Ris (1991), and Goldberg and Allen (1992). The central plug or channel complex has been modeled as a transparent ellipsoidal particle to indicate the fact that its definitive structure remains elusive (see chapter by Pantel and Aebi in this volume). Tyrp265, CAN/NUP214, and NUP140 exhibit epitopes residing in the cytoplasmic filaments (Willem et al., 1993; Byrd et al., 1994; Pantel et al., 1994), whereas NUP153 COOH-terminal domain epitopes have been localized at the terminal ring of the nuclear basket (Pantel et al., 1994). P62 epitopes are known at both the cytoplasmic and nuclear peripheral of the central plug or channel complex (Gian et al., 1995). The transmembrane glycoprotein gp210 exhibits several epitopes in the lumens of the NE (Greber et al., 1990) where based on its topology most of its mass resides (see Fig. 3; Greber et al., 1990). Epitopes for the transmembrane glycoproteins POM121 and POM152 are also indicated. However, these latter two should not be taken literally since their exact localization remains uncertain.

PLATE 12 (Leonard and Cardoso, Fig. 1). Targeting of β-galactosidase to replication foci (A-C) or to the speckled compartment (D-F) by fusion to the DNA methyltransferase targeting sequence (Leonard et al., 1992) or the suppressor of white-apricot RS domain (Li and Blumgrem, 1991), respectively. Mouse fibroblasts were transfected with the fusion constructs. Fusion proteins were detected with a monoclonal antibody to the galactosidase epitope (red; A, D) and compared to the DNA replication foci visualized with an antibody to DNA methyltransferase (green; B, E). Methyltransferase was previously shown to colocalize with sites of DNA replication in S phase cells (Leonard et al., 1992). Double exposures are shown to better illustrate colocalization of the DNA methyltransferase visualized at sites of DNA replication (yellow in C) and the different localization of the RS domain (red and green in F).

PLATE 13 (Leonard and Cardoso, Fig. 4). Specific localization of cyclin A (red) at subnuclear DNA replication foci visualized by BrdU incorporation and containing with BrdU-specific antibodies (green) in pulse labeled cultures of retinoblastoma murine myocytes. A localization of cyclin A (red) at subnuclear DNA replication foci visualized by BrdU incorporation and containing with BrdU-specific antibodies (green) in pulse labeled cultures of retinoblastoma murine myocytes is shown in A. The colocalization of cyclin A (red) and BrdU incorporation (green) is shown in B. The inset shows a high magnification of the foci. For further details see Cardoso et al. (1993).