1987

# Novel intermediate filaments in the nuclear envelope: molecular cloning and analysis of lamins A and C

Daniel Zev Fisher


# Novel Intermediate Filaments in the Nuclear Envelope: Molecular Cloning and Analysis of Lamins A and C

A thesis submitted to the faculty of The Rockefeller University

in partial fulfillment of the requirements

for the degree of Doctor of Philosophy

by

Daniel Zev Fisher

March, 1987

The Rockefeller University

New York, New York

## ABSTRACT

Lamins A, B and C are the major polypeptides of the nuclear lamina, the filamentous meshwork intercalated between chromatin and the inner nuclear membrane. Amino acid sequences have been deduced from cDNA clones of human lamins A and C which show identity with the exception of an additional 98 amino acid residue carboxyl-terminal tail present only in lamin A and a 6 amino acid residue stretch at the carboxyl terminus of lamin C. The lamin A precursor has a predicted MW of 74 kDa while lamin C has a predicted MW of 65 kDa. The lamins have several structural features that identify them as members of the family of intermediate filament proteins. Both lamins share an approximately 360 amino acid residue domain of α-helical conformation that, in turn, shows primary and secondary structural homology with the α-helical rod domain of the intermediate filament proteins. However, the lamin α-helical domain is 14% larger than that of the intermediate filament proteins. Analogous to the intermediate filament protein α-helical domain, the lamin α-helical domain contains three regions (1A, 1B and 2) of heptad repeats where the first and fourth residue of each heptad is usually hydrophobic or nonpolar. This results in an apolar stripe down one side of the α-helix. Lamins A and C may therefore assemble into coiled coil homodimers by interactions of their α-helical domain. Coil 1B of the lamins is approximately twice the size of Coil 1B of intermediate filament proteins, thus accounting for the larger size of the lamin α-helical domain. The highly conserved sequences found at the amino and carboxyl ends of the α-helical domains of all classes of intermediate filament proteins, as well as the highly conserved positions of the intercoil linkers, are present in the lamins as well. The lamins share an extra-helical 30 amino acid residue amino-terminal domain that bears no homology to any known intermediate filament protein sequence, yet has a slightly basic character (*Arg* residues) analogous

to nonepithelial intermediate filament proteins. The extra-helical carboxyl-terminal domain shared by both lamins contains a putative nuclear localization signal and an abundance of *Ser* residues. These *Ser* residues may be involved in the phosphorylation and dephosphorylation events that coincide with lamina disassembly and reassembly, respectively, during mitosis. The unique lamin A carboxyl-terminal tail is high in *Gly* and *Ser* residues and shows limited homology to head and tail regions of Type I and Type II keratins. Thus, nuclear lamins may initiate assembly into a filamentous lamina by interactions based on properties found among all classes of intermediate filament proteins.

## TABLE OF CONTENTS

# LIST OF FIGURES

# ABBREVIATIONS

| | |
|---|---|
| bp | base pairs |
| DEAE | diethylaminoethyl |
| DTT | dithiothreitol |
| IPTG | isopropylthio-β-galactoside |
| kb | kilobase |
| $M_r$ | apparent molecular weight |
| PAGE | polyacrylamide gel electrophoresis |
| PCL | pore complex-lamina |
| PFU | plaque forming units |
| PMSF | phenylmethylsulfonylfluoride |
| SDS | sodium dodecyl sulfate |
| TCA | trichloroacetic acid |

## ACKNOWLEDGEMENTS

I wish to thank Nilabh Chaudhary who was my partner on this lamin project, and who greatly contributed to both the sequencing and protein work that was performed. I also thank Nilabh for his patience, good humor and good will throughout all the ups and downs that we experienced.

Special thanks to the Blobel lab "nuclear" group, especially to Laura Davis and John Aris, for helpful advice and discussions. Special thanks to Rick Wozniak for the use of antibodies, and to Spyros Georgatos for helpful discussions and conveying an intense enthusiasm about the field of intermediate filaments.

I would like to thank my advisor, Dr. Gunter Blobel, for exposing me to his groundbreaking insights and visions in cell biology, and for giving me the opportunity to work in that most special environment, the Blobel lab.

Finally, I wish to thank my family and friends, and my wife, Carol, for their support.

# Chapter 1

# Introduction

## *1.1 Nuclear Envelope Components*

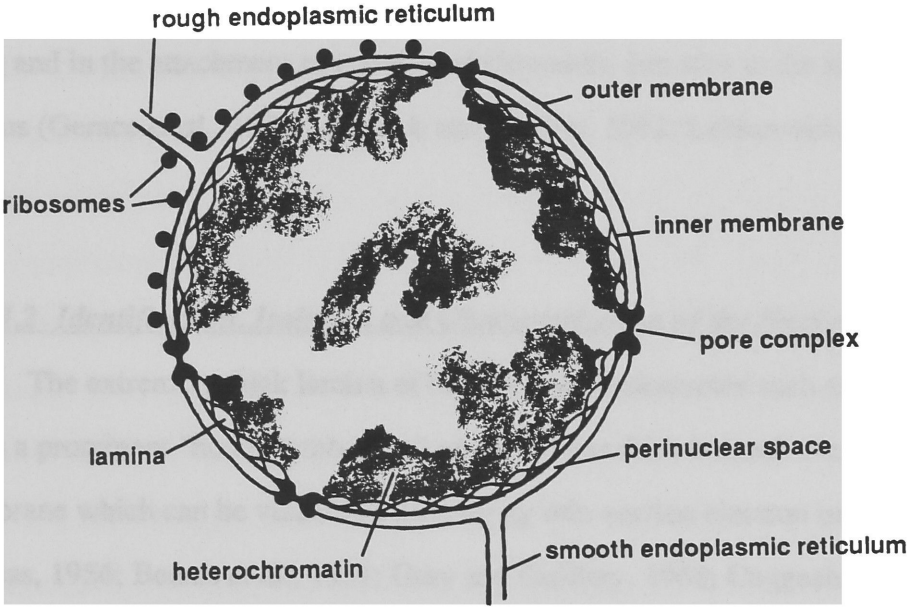
The nuclear envelope is a complex membrane system that forms the boundary between the nuclear and cytoplasmic compartments in eukaryotic cells. This membrane organelle regulates nucleocytoplasmic exchange of macromolecules and metabolites, and may be involved in the topological organization of interphase chromosome structure. The presence of a nuclear envelope which segregates nucleoplasmic DNA synthesis, transcription and RNA processing from cytoplasmic RNA translation into protein fundamentally distinguishes eukaryotic cells from prokaryotes. The primary nuclear envelope structural components include inner and outer nuclear membranes, pore complexes and the nuclear lamina (Fig. 1). The two membranes, which encircle the nucleus, are separated by a perinuclear space of usually 40-80 nm which is in direct continuity with the cisternae of the rough and smooth endoplasmic reticulum (Watson, 1955). The outer nuclear membrane, which is often studded with ribosomes, is continuous with the membranes of the endoplasmic reticulum (Watson, 1955; Afzelius, 1955) and also joins the inner nuclear membrane to from pores of approximately 80 nm diameter. The pore complexes (Watson, 1959) are large proteinaceous superstructures situated at the membrane pores which provide aqueous channels across the nuclear envelope for nucleocytoplasmic exchange of ions and molecules (reviewed by Franke, 1974; Fry, 1976; Maul, 1977; Franke *et al.*, 1981). These complexes appear to be freely permeable to ions, small molecules and small proteins while more discriminative to larger macromolecules such as large proteins and ribonucleoproteins (Paine *et al.*, 1975; Bonner, 1975a,b; DeRobertis *et al.*, 1978). Molecules in the latter class have been shown to be transported across the envelope via some selective entry

## *Figure 1*

*Major Structural Components of the Nuclear Envelope*

The nuclear envelope contains inner and outer nuclear membranes separated by a 40-80 nm perinuclear space which is in direct continuity with the cisternae of the rough and smooth endoplasmic reticulum. The two membranes join at nuclear pore complexes, supramolecular proteinaceous assemblies that provide channels across the nuclear envelope for nucleocytoplasmic transport. The outer nuclear membrane, which is continuous with the membrane of the endoplasmic reticulum, is often studded with ribosomes. The nuclear lamina is a protein meshwork situated between the inner nuclear membrane and peripheral heterochromatin. The lamina is thought to interact with chromatin, the inner nuclear membrane and the pore complexes.

# The Nuclear Envelope



mechanism(s) (Dingwall *et al.*, 1982; Feldherr and Ogburn, 1983; Feldherr *et al.*, 1984).

The nuclear envelope lamina is a protein meshwork situated between peripheral chromatin and the inner nuclear membrane of the nuclear envelope. It has been referred to as the fibrous lamina (Fawcett, 1966; Ghadially *et al.*, 1972), zonula nucleum limitans (Mazanec, 1967; Patrizi and Poger, 1967) and dense lamella (Kalifat *et al.*, 1967). The lamina is postulated to be a scaffolding involved in the organization of nuclear envelope structure (Gerace *et al.*, 1978; Gerace and Blobel, 1982) and in the attachment of compacted chromatin domains to the interphase nucleus (Gerace *et al.*, 1978; Hancock and Hughes, 1982; Lebkowski and Laemmli, 1982).

### *1.2 Identification, Isolation and Characterization of the Nuclear Lamina*

The extremely thick lamina of certain lower eukaryotes such as amoeboe forms a prominent "honeycomb layer" up to 300 nm thick beneath the inner nuclear membrane which can be visualized directly by thin-section electron microscopy (Pappas, 1956; Beams *et al.*, 1957; Gray and Guillery, 1963; Coggeshall and Fawcett, 1964; Fawcett, 1966; Daniels and Breyer, 1967). The large almost hexagonal holes seen on cross section through this layer appear to be channels that run between the nuclear interior and the pores. A 30-100 nm wide lamina of certain higher eukaryotic cells such as chondrocytes can also be visualized directly by thin-section electron microscopy (Fawcett, 1966; Kalifat *et al.*, 1967; Mazanec, 1967; Patrizi and Poger, 1967; Ghadially *et al.*, 1972; Oryschak *et al.*, 1974, 1976). However, in most cell types the lamina is much thinner and can only be visualized after nuclear subfractionation (Aaronson and Blobel, 1975; Scheer *et al.*, 1976).

Highly insoluble nuclear envelope fractions enriched in the lamina can be prepared by sequential treatment of envelopes with non-ionic detergents and high salt

concentrations (Dwyer and Blobel, 1976; Shelton *et al.*, 1980; Krohne *et al.*, 1981; Scheer *et al.*, 1976). These conditions solubilize the membranes and extract histones and residual chromatin. Among the cell types analyzed in certain vertebrates and invertebrates, the lamina contains one to three prominent 60-80 kDa polypeptides, termed "lamins" (Gerace and Blobel, 1980). The lamins are identified by their localization in the nuclear lamina by light and electron microscopic immunocytochemistry (Gerace *et al.*, 1978; Krohne *et al.*, 1978; Stick and Hausen, 1980). Rat liver nuclear envelopes contain three lamins, lamins A (70 kDa), B (67 kDa) and C (60 kDa), that comprise about 40% of the total mass of isolated rat lamina-containing fractions (Dwyer and Blobel, 1976). By electron microscopy these fractions can be seen to contain an intact lamina about 15 nm thick associated with nuclear pore complexes (the "pore complex-lamina" or "PCL" fractions) (Dwyer and Blobel, 1976). Lamins, however, do not appear to be located in nuclear pore complexes; rather, they appear to be associated as a polymeric structure (Dwyer and Blobel, 1976) within the lamina (Gerace *et al.*, 1978).

The lamins represent a family of related proteins. Peptide maps of lamins A and C are similar (Shelton *et al.*, 1980; Gerace and Blobel, 1982; Kaufmann *et al.*, 1983), although they are different from lamin B. Lamin B also has a stronger physical interaction with the inner nuclear membrane than lamins A and C as determined by its greater resistance to extraction with chemical protein perturbants (Gerace and Blobel, 1982; Gerace *et al.*, 1984). However, certain monoclonal antibodies have been shown to crossreact with all three lamins (Burke *et al.*, 1983; Krohne *et al.*, 1984), indicating that some epitopes are shared by all three lamins.

Rat (liver), Chinese hamster (CHO), human (HeLa) and bovine (MDBK) cells all apparently contain three analagous lamins as determined by probing with antibodies to rat liver lamins (Gerace and Blobel, 1982). Birds are reported to have two (Stick and Hausen, 1980) or three (Shelton and Egle, 1979; Shelton *et al.*, 1980)

lamins. Among amphibians, *Xenopus laevis* has four lamins, $L_I$-$L_{IV}$, as determined by tryptic peptide maps and two dimensional gel analysis (Benavente *et al.*, 1985; Benavente and Krohne, 1985; Krohne *et al.*, 1981; Krohne *et al.*, 1984; Stick and Krohne, 1982; Stick and Hausen, 1985), while among molluscs, the oocyte of the clam *Spisula solidissima* has only one lamin (Maul *et al.*, 1984). Two potential lamins have been detected in *Drosophila* (Fuchs *et al.*, 1983; McKeon *et al.*, 1983; Smith and Fisher, 1984; Risau *et al.*, 1981). Thus, lamins appear to be a ubiquitous component of the nuclear envelope of all vertebrates and invertebrates. However, no lamina structure has been observed in many lower eukaryotes where a pattern of complete nuclear disassembly often does not occur (Franke, 1974). For example, only partial nuclear envelope disruption occurs in *Physarum polycephalum* (Guttes *et al.*, 1968; Ryser, 1970) while no observable disruption occurs in certain dinoflagellates (Kubai and Ris, 1969). Whereas these "closed" mitoses may reflect differences in nuclear architecture between higher and lower eukaryotes, the nuclear envelope of lower eukaryotes still undergoes structural alterations even during closed mitosis, so a lamina may be present.

### *1.3 Lamin Expression*

Lamins are expressed in a cell type-specific fashion as has been demonstrated in *Xenopus laevis* by studies of lamin expression during embryonic development and cellular differentiation (Benavente *et al.*, 1985; Benavente and Krohne, 1985; Stick and Hausen, 1985). Of the four lamins, $L_I$-$L_{IV}$, both $L_I$ and $L_{II}$ are found in many somatic cells, $L_{III}$ is found in mature oocytes and certain somatic cells, and $L_{IV}$ occurs only in male germ cells. The polymerized lamina of diplotene oocytes containing $L_{III}$ breaks down during oocyte maturation and the lamins are solubilized. After fertilization, the stockpile of soluble $L_{III}$ is reutilized and progressively polymerizes to form the nuclear lamina structures of embryonic nuclei up to the midblastula

transition. Up to the midblastula stage, no newly synthesized $L_{III}$ is detectable and $L_I$ and $L_{II}$ are absent. $L_I$ and $L_{II}$ begin to be expressed at the midblastula and gastrula stages, respectively, while $L_{III}$ progressively disappears to non-detectable levels in swimming tadpoles. Interestingly, however, $L_{III}$ is expressed for a short time after the midblastula transition before falling off. Lamins $L_I$ and $L_{II}$ become the primary lamins of *Xenopus* somatic cells. Later in development, $L_{III}$ is found coexpressed with $L_I$ and $L_{II}$ in muscle cells, neurons and Sertoli cells. The only common features of the cells which express $L_{III}$ is their large size, highly differentiated state, and inability to undergo further division. In gametogenesis of *Xenopus*, lamin expression switches from $L_I/L_{II}$ to $L_{III}$ in oocytes at early diplotene (Stick and Krohne, 1982; Benavente *et al.*, 1985) and to $L_{IV}$ in spermatids after the second meiotic division (Benavente *et al.*, 1985). In early stages of meiosis, none of the lamins are detectable and, indeed, a nuclear lamina is no longer found in *Xenopus* nor in some other vertebrate species (Stick and Schwarz, 1982, 1983).

In *Drosophila*, a 74 kDa lamin is detected during early embryonic stages (*i.e.* 2-hr-old embryos) and is joined by a 76 kDa lamin in older embryos and in 16- to 21-hr-old hatching larvae (Smith and Fisher, 1984). Immunofluorescence studies during mouse embryogenesis (Schatten *et al.*, 1985) indicate an apparent shift in detectability of lamins A, B and C in pronuclei and blastomeres to lamin B alone at the morula and blastocyst stages. Whether this represents a change in lamin expression, modification of existing lamin epitopes, or other cellular changes affecting the accessibility of the antibodies used remains to be determined. It is quite apparent, nevertheless, that the lamins are a multigene family of polypeptides which exhibit developmentally regulated expression in at least some species.

### *1.4 Nuclear Lamina Dynamics During the Cell Cycle*

The lamins have been shown to be synthesized throughout the cell cycle

(Gerace *et al.*, 1984). *In vitro* translation of mRNA from mammalian (Laliberte *et al.*, 1984; Gerace *et al.*, 1984) and avian (Lehner *et al.*, 1986) cells has shown that lamins A and C are synthesized from separate mRNAs and that the translation product of the lamin A mRNA is approximately 2 kDa larger than mature lamin A. This precursor form of lamin A is converted into mature lamin A only after it is incorporated into the insoluble lamina structure, as shown by pulse-chase experiments (Gerace *et al.*, 1984). However, the precursor form of lamin A is not necessary for lamina formation since mature lamin A solubilized in mitosis is fully capable of reintegrating into the lamina polymer (see below) (Gerace and Blobel, 1980). In Buffalo Rat liver cells, the half-time of incorporation of newly synthesized lamin A is about 5 minutes while that of lamins B and C is about 60 minutes (Gerace *et al.*, 1984).

Dramatic changes occur in the location of the lamins during mitosis. Concomitant with nuclear envelope disassembly, the lamina is depolymerized during prophase (Gerace *et al.*, 1978; Krohne *et al.*, 1978) by a process that may involve transient hyperphosphorylation (see below) (Gerace and Blobel, 1980). The lamins shift from their focalized peripheral nuclear location to a diffuse distribution throughout the cell. However, lamin B remains associated with fragmented nuclear membrane vesicles (Gerace and Blobel, 1980). As the nuclear envelope reforms during telophase, the lamins progressively accumulate at the surface of the daughter chromosomes as they are reutilized for reconstruction of the daughter cell nuclear envelopes (Gerace *et al.*, 1978; Krohne *et al.*, 1978; Gerace and Blobel, 1980). In CHO cells, the lamins are disassembled into moieties that sediment at 4-5S on sucrose gradients in the presence of Triton X-100 (Gerace and Blobel, 1980). In contrast, solubilized lamin $L_{III}$ of arrested *Xenopus* oocytes sediments at 9S on sucrose gradients, although in the absence of Triton X-100 (Benavente *et al.*, 1985).

Phosphorylation and dephosphorylation of the lamins have been suggested to

regulate lamina disassembly and reassembly, respectively, during mitosis (Gerace and Blobel, 1980). Recently, lamina disassembly *in vitro* was shown to exhibit kinetics indicative of regulation by an enzymatic activity, such as a lamin kinase (Suprynowicz and Gerace, 1986; Newport and Spann, 1987). This activity was present in the soluble fraction of metaphase CHO cell homogenates (Suprynowicz and Gerace, 1986) and in mitotic extracts from Xenopus eggs (Newport and Spann, 1987). During interphase, CHO cell lamins contain 0.25-0.4 mol P/mol lamin on *Ser* and *Thr* residues on multiple sites (Ottaviano and Gerace, 1985). All detectable interphase lamin phosphate groups are found in the insoluble lamina. In striking contrast, mitotic lamins of CHO cells contain 2-3 mol P/mol lamin during disassembly *in vivo* (Gerace and Blobel, 1980; Ottaviano and Gerace, 1985) and *in vitro* (Suprynowicz and Gerace, 1986). Upon reassembly of the daughter cell nuclear envelopes, the lamins become dephosphorylated to interphase levels. In addition to regulating the reversible disassembly of the lamina during mitosis, phosphorylation may alter local lamin interactions during interphase to permit lamina growth and restructuring (Ottaviano and Gerace, 1985). The lamins may be substrates of the putative phosphorylation cascade proposed to occur during mitotic prophase which may be involved in the coordinated reorganization of many cellular structures (Wu and Gerhart, 1980; Miake-Lye *et al.*, 1983; Miake-Lye and Kirschner, 1985).

Based on the dynamic behavior of the lamina during mitosis and its ultrastructural and biochemical associations, the reversible disassembly of the lamina has been postulated to regulate nuclear envelope dynamics during mitosis (Gerace *et al.*, 1978; Gerace and Blobel, 1980). Lamina disassembly during prophase has been proposed to be a necessary yet not sufficient condition for the breakdown of the nuclear envelope into vesicles. *In vivo* and *in vitro* studies indicate that lamina depolymerization and nuclear membrane breakdown may be independent processes (Stick and Schwarz, 1983; Miake-Lye and Kirschner, 1985; Newport, 1987) with

lamina disassembly preceding nuclear membrane breakdown (Miake-Lye and Kirschner, 1985). Nuclear membrane vesiculation may subsequently occur since the inner membrane is released from its association with the lamina. Reassembly of the envelope during telophase may be directed by repolymerization of the lamina as the nuclear membrane vesicles associated with lamin B (Gerace and Blobel, 1980; Gerace *et al.*, 1984) are brought together with each other and the chromosome surface when lamin B reassociates with lamins A and C around the daughter cell chromosomes. Thus, the association of lamins A and C with chromosomes and the association of lamin B with nuclear membranes ensures the correct targeting of the membranes to the chromosome surface when the lamins reassemble at telophase. Support for this model comes from *in vivo* and *in vitro* studies. In a cell-free nuclear assembly system (Burke and Gerace, 1986), immunological depletion of either lamins A and C or lamin B resulted in strong inhibition of subsequent nuclear envelope assembly. Lamins A and C were shown to assemble around chromosomes *in vitro* without lamin B or membranes present. Conversely, depletion of lamins A and C prevented the assembly of lamin B or membranes around chromosomes. *In vivo*, microinjection of appropriate concentrations of anti-lamin antibodies into metaphase tissue culture cells was shown to prevent lamin assembly around chromosomes and proper nuclear reformation at telophase (Benavente and Krohne, 1986). Chromosomes were arrested in their condensed state at telophase. These studies clearly demonstrate the importance of lamina assembly in the reformation of the nuclear envelope in telophase. By extension, the lamins may utilize the same mechanisms on a local basis for the restructuring of the nuclear envelope during cell growth and differentiation in interphase.

Most recently, experiments monitoring nuclear assembly around protein-free DNA (bacteriophage λ DNA) in *Xenopus* oocyte extracts have suggested a two-stage model for *in vitro* assembly of the nucleus (Newport, 1987). The first stage involves

histone insertion into the DNA to form chromatids followed by condensation of the DNA into a structure similar to a metaphase chromosome. The second stage involves lamin and membrane vesicle binding to the chromatin, vesicle fusion and chromatin decondensation. Interestingly, nuclear envelope reformation occurred considerably more rapidly around metaphase chromosomes than around the phage "chromatin" structure, suggesting that the metaphase chromosomes contain receptors that initiate lamin polymerization or form a direct link to the nuclear membrane. Alternatively, the DNA itself might act as a template for lamin binding when it reaches an appropriate supercoiled configuration.

How the regular protein meshwork that constitutes the lamina is assembled from the three lamins is not known. The work reported in this thesis rested on the assumption that each lamin molecule contains multiple functional domains that might be correlated with structural domains. The lamins must contain domains that interact with each other to form the regular protein meshwork that constitutes the lamina. They must also have domains that interact with chromatin and/or chromatin-associated proteins on the nucleoplasmic side of the lamina, with the lipid bilayer and/or integral proteins of the inner nuclear membrane on the other face of the lamina, and with pore complexes, which interrupt the lamina. Since such domains could be revealed from the amino acid sequences of the lamins, I cloned and sequenced the cDNAs of lamin A and lamin C (Fisher *et al.*, 1986). Independently, the lamins were cloned elsewhere (McKeon *et al.*, 1986). The deduced amino acid sequences of lamins A and C are identical except for an extra 98 amino acid residue (9 kDa) carboxyl-terminal tail that is present only in lamin A and a 6 amino acid residue stretch present at the carboxyl terminus of lamin C. Most strikingly, both lamins A and C contain an α-helical domain of approximately 360 residues that shows primary and secondary structural homology to a corresponding α-helical rod domain that is the structural hallmark of all intermediate filament proteins.

Implications of the presence of this and other domains in lamins A and C for the assembly of the nuclear lamina will be discussed.

# Chapter 2

# Materials and Methods

## *2.1 Screening cDNA Libraries*

Guinea pig antibodies to rat liver lamins A, B, and C (gift of L. Gerace, Johns Hopkins Univ.) were used to screen a rat liver λgt11 cDNA library (gift of R. Hynes, MIT) and a human hepatoma (Li-7) λgt11 library (gift of J.R. deWet, UCSD) according to established protocols (Young and Davis, 1983; Schwarzbauer *et al.*, 1983). Eight cDNA clones were obtained from the rat library and one was obtained from the human library. The largest λ clone was obtained from the rat library and contained a 1.3 kilobase (kb) insert.

In order to identify which of the three lamins this 1.3 kb cDNA might represent, the recombinant λgt11 phage harboring the 1.3 kb cDNA was introduced into the lysogenic strain Y1089 and was induced and expressed in high abundance as a hybrid polypeptide fused to β-galactosidase (Dame *et al.*, 1984). Cells were lysed by freeze-thawing and proteins were fractionated on a 7.5%-15% gradient gel by sodium dodecyl sulfate/polyacrylamide gel electrophoresis (SDS/PAGE) (Dame *et al.*, 1984; Laemmli, 1970). The proteins were then blotted onto nitrocellulose and screened with antibodies to the rat liver PCL fraction (gift of R. Wozniak, Rockefeller Univ.) (Towbin *et al.*, 1979) . Antibodies bound to the fusion protein were eluted (Smith and Fisher, 1984) and used to screen a blot containing proteins of the rat PCL fraction (Dwyer and Blobel, 1976) (similarly blotted from a 7.5%-15% SDS/polyacrylamide gradient gel) to identify with which lamin(s) the antibodies to the fusion protein cross-reacted.

The 1.3 kb clone was then nick-translated (Rigby *et al.*, 1977) and used to screen (Grunstein and Hogness, 1975) the simian virus 40 transformed human fibroblast cDNA library of Okayama and Berg (1983) (gift of H. Okayama, NIH). A

2.6 kb partial cDNA clone of lamin A was obtained and sequenced (see *2.3*). It was clear that this was not a full-length clone when one compared its size to that predicted by RNA blot analysis (see *2.2* and *3.1*). Furthermore, the deficiency was located at the 5' end of the cDNA since the clone contained a 3' polyadenylation signal and tail (see Fig. 8, *3.2*). Therefore a 22 base pair (bp) oligonucleotide, TGGCCCCGCAGATCATGCAGCT, complementary to a 5' portion of the 2.6 kb clone (Fig. 8, *3.2*, codons 161 to 168) was synthesized (Applied Biosystems 380A DNA Synthesizer), end-labeled with [$\gamma$-$^{32}$P]ATP and phage T4 polynucleotide kinase (Maxam and Gilbert, 1980) and used to rescreen the fibroblast library for clones containing the complete 5' end. A 1.8 kb cDNA clone was obtained. This was determined to be an incomplete clone of lamin C by sequencing (see *2.3* and *3.2*). A 21 bp oligonucleotide, GCATCCCCGAGCTCGGCCTCG, complementary to a 5' portion of the 1.8kb clone (Fig. 8, *3.2*, codons 81 to 88) was synthesized and used to screen (Benton and Davis, 1977) a human hepatoma $\lambda$gt11 cDNA library (gift of M. Mueckler, MIT) for full-length cDNA clones. A 1962 bp cDNA clone of lamin C was obtained. Blots showing the isolation of this clone through primary, secondary and tertiary screens are shown in Fig. 2.

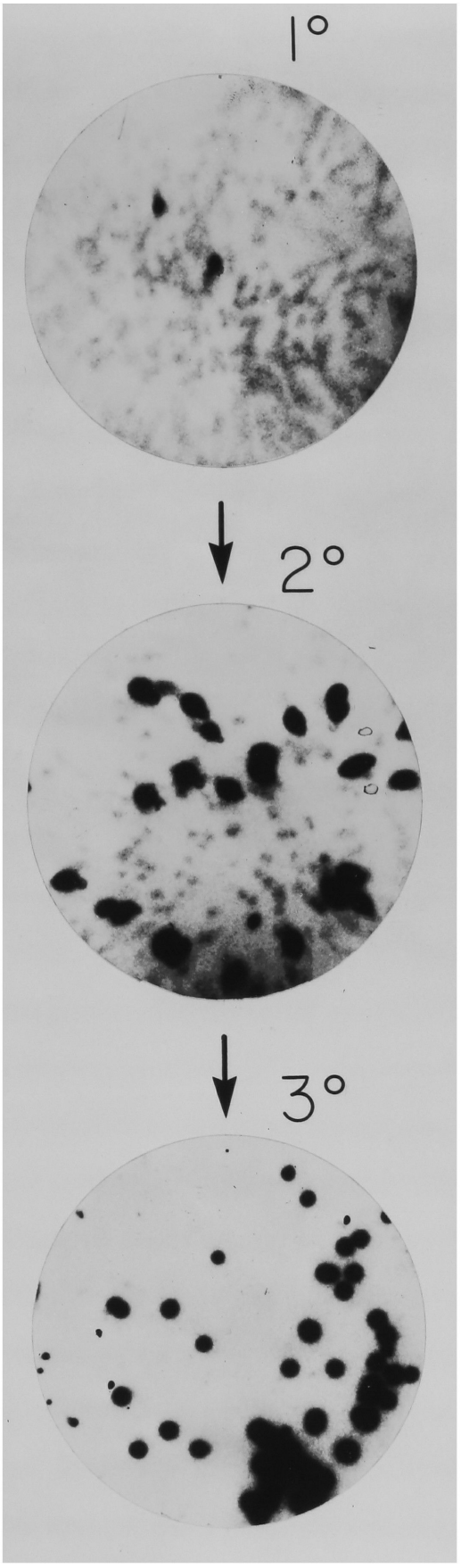
## ***2.2 RNA Blot Analysis***

HeLa cells were homogenized in guanidinium thiocyanate and total cellular RNA was pelleted through a cesium chloride cushion (Chirgwin *et al.*, 1979). Poly(A)+ RNA was isolated by oligo(dT)-cellulose chromatography (Aviv and Leder, 1972), size-fractionated by electrophoresis in a 1.2% formaldehyde/agarose gel and transferred to nitrocellulose (Thomas, 1980). Blots were hybridized (Alwine *et al.*, 1977) with nick-translated rat liver cDNA (the 1.3 kb clone, see *2.1*) or $^{32}$P-labelled oligonucleotides (the 22mer and 21mer, see *2.1*). Nick-translated cDNA clones of the heat shock protein hsp 70 (2.9 kb) and actin (2.0 kb) (gift of H. Kao, Rockefeller

## *Figure 2*

*Screening a λgt11 Recombinant Human Hepatoma cDNA Library for a Lamin A/C Recombinant*

*E. coli* Y1090 cells (Young and Davis, 1983) were infected with recombinant phage and plated for plaques on 90 mm plates at a density of 0.5-1 x $10^4$ plaque-forming-units (PFU) for the primary screen, 1 x $10^3$ PFU for the secondary screen, and 1 x $10^2$ PFU for the tertiary screen. Cells were grown for 6-10 hrs at 37°C. Dry nitrocellulose filters were placed on the bacterial lawn for transfer of phage and DNA to the filter. Filters and plates were marked for orientation by making holes with an 18-gauge needle. Filters were then removed from the plates and soaked sequentially for 5 min in denaturing solution (0.5 M NaOH/1.5 M NaCl), neutralizing solution (0.5 M Tris-HCl, pH 8.0/1.5 M NaCl) and wash solution (30 mM Na citrate/300 mM NaCl [2x SSC]). Filters were air dried and then baked in a vacuum at 80°C for 2 hrs. Baked filters were washed for ≥ 4 hrs at 65°C in 3x SSC/0.1% SDS with multiple changes. Filters were prehybridized for 1 hr at 74°C in 6x SSC/1X Denhardts (0.2% ficoll/0.2% polyvinylpyrrolidone/0.2% BSA)/0.1% SDS/2 mM $NaH_2PO_4$/18 mM $Na_2HPO_4$/100 μg/ml salmon sperm DNA, and then hybridized with radiolabelled oligonucleotide probe (21mer, see *2.1*) ($10^5$-$10^6$ counts per minute) for 3-12 hrs at 55°C. Filters were washed for 30 min at 20°C and for 15 min at 55°C in 6x SSC/0.1% SDS. Filters were then dried and autoradiographed for 24 hrs. Pen marks were made on the autoradiographs to align plaque images with plaques on the plates. Positive primary plaques were picked with the wide end of a pasteur pipette and rescreened. Single positive secondary plaques were picked with the tip of a pasteur pipette and rescreened. All plaques tested positive in the tertiary screen.


1°
2°
3°

Univ.) were hybridized as internal $M_r$ markers. Blots were washed for one hour at 22°C, air dried, and exposed to Kodak XAR-5 film for 1-14 hours.

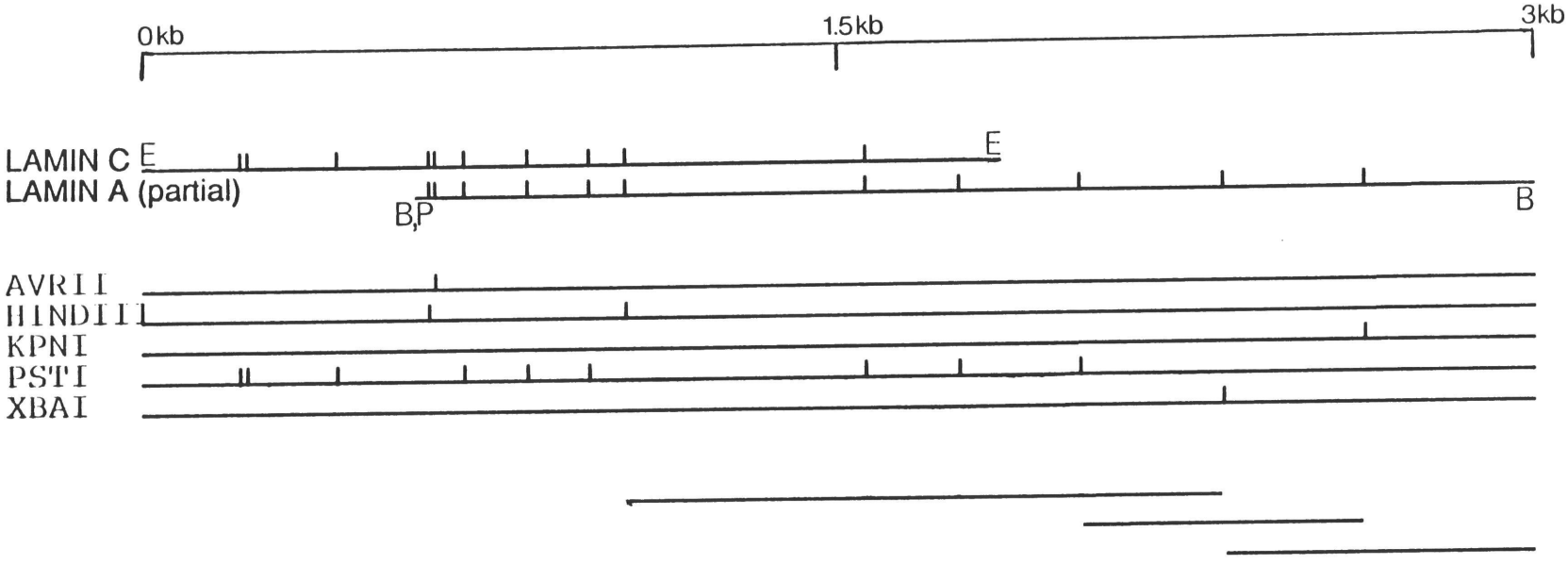
### *2.3 DNA Sequence Analysis*

Restriction maps of the lamin cDNAs were generated by digesting them with the hexanucleotide-cutting restriction endonucleases that cut within the multicloning site of the M13 vectors, M13mp18 and M13mp19 (Messing, 1983; Yanisch-Perron *et al.*, 1985) (see Fig. 3). In addition, the lamin C insert was straddled by EcoRI sites (generated from construction of the λgt11 library) while the lamin A insert was straddled by BamHI sites with an additional PstI site 5' to the insert (generated from construction of the Okayama and Berg library, see *2.1*). As can be seen in Fig. 3, multiple relatively small fragments are generated with the enzyme, Pst I. These restriction fragments, as well as fragments generated by cutting adjacent restriction sites (*i.e.* PstI+XbaI, XbaI+KpnI, *etc.*) of the plasmid and phage cDNA inserts were subcloned into the M13 vectors and sequenced in both orientations by the dideoxy chain termination method (Sanger *et al.*, 1977) using [α-$^{35}$S]dATP and buffer gradient gels (Biggin *et al.*, 1983). 25% formamide was included in some gels to resolve sequencing ambiguities generated by secondary structure alterations of DNA migration through standard 8 M urea-containing gels. Also, chemical sequencing (Maxam and Gilbert, 1980) was used to sequence regions not adequately resolved by the dideoxy method. The rapid deletion subcloning procedure of Dale *et al.* (1985) was used to sequence large fragments and to sequence across restriction sites. As depicted in Fig. 3, a HindIII/XbaI fragment, a PstI/KpnI fragment and an XbaI/BamHI fragment were sequenced by the rapid deletion procedure. Additionally, both the partial-length and full-length lamin C cDNAs were sequenced in their entirety by this method. Predictions for the secondary structure of the deduced amino acid sequences were computer-generated according to the method of

## *Figure 3*

*Restriction Maps of Isolated Lamin A and C cDNAs*

cDNA restriction maps of cloned lamins A and C were deduced by restriction analysis of the lamin cDNAs subcloned into pUC18 and pUC19 and were confirmed by sequencing. Adjacent restriction fragments along the length of the cDNAs were subcloned into M13 vectors and sequenced. Note the large number of PstI sites. Additionally, the deletion subcloning technique of Dale *et al.* (1985) was used to sequence across the restriction sites located within the three restriction fragments depicted at the bottom of the figure. This procedure was also used to sequence the 1.8 kb and full-length lamin C cDNAs. The single AvrII site was utilized to construct a full-length lamin A cDNA (see *2.5*). **B**, **P** and **E** represent BamHI, PstI and EcoRI linkers at the ends of the cDNAs.


0kb
1.5kb
3kb
LAMIN C
LAMIN A (partial)
E
E
B,P
B
AVRII
HINDIII
KPNI
PSTI
XBAI

Chou and Fasman (1978).

### *2.4 Lamin Protein Sequencing*

Rat liver nuclear PCL fraction was purified according to established procedures (Dwyer and Blobel, 1976). Lamins from approximately 6.0 x $10^9$ nuclei (2000 $A_{260}$ of nuclei) were solubilized in 5 mls of 6 M urea/10 mM Tris-HCl pH 8.0/20 mM dithiothreitol (DTT), loaded onto a 2-3 ml diethylaminoethyl (DEAE) - cellulose column equilibrated in the above buffer and eluted with 20 mls of a 0 to 200 mM NaCl gradient at 20°C. 0.5 ml fractions were collected and 20 µl aliquots were subjected to SDS/PAGE (Laemmli, 1970) (see Fig. 4).

For acid hydrolysis of *Asp-Pro* bonds only, eluted fractions enriched in lamins A and C (*i.e.* Fig. 4, fractions 14-19) were electrophoresed on 7.5%-15% SDS/polyacrylamide gradient gels and electroblotted onto trifluoroacetic acid treated glass fiber filters (Whatman GF/C) according to the low pH procedure of Aebersold *et al.* (1986). Protein bands were visualized with Coomassie Blue and lamin A and lamin C bands were cut from the filters. *Asp-Pro* cleavage was carried out by spotting 88% formic acid (30 µl/$cm^2$ of filter strip) and incubating for 20 hrs at 44°C in a sealed chamber saturated with 88% formic acid (D. Atherton, Rockefeller Univ.). Filters were then air dried for 15 minutes.

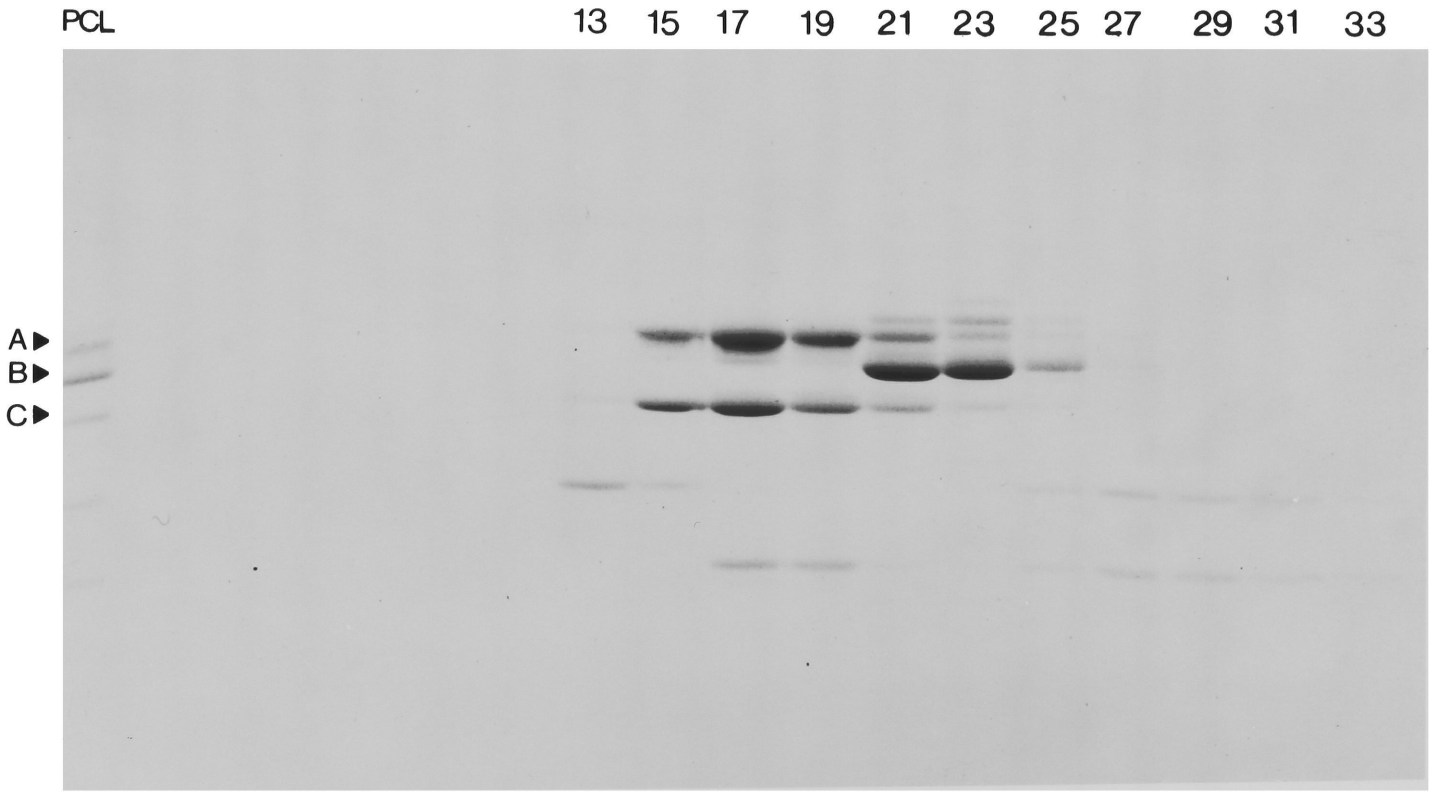
For V8 protease digestion, fractions off the DEAE-cellulose column containing lamins A and C were pooled, diluted to 2 M with $H_2O$ and precipitated on ice with 10% trichloroacetic acid (TCA); precipitates were resuspended in 0.2% SDS/125 mM Tris-HCl pH 6.8/20 mM DTT, and proteins were cleaved with V8 protease (50 ng V8/µg lamins) (Miles Laboratories) for 1 hr at 37°C, as described (Cleveland *et al.*, 1977). Samples were electrophoresed and electroblotted onto glass fiber filters as above. Protein bands were visualized with Coomassie Blue.

The lamin fragments on glass fiber filters were subjected to Edman

## *Figure 4*

*Purification of Lamins from Urea-solubilized Rat Liver PCL Fractions*

Rat liver PCL fractions were solubilized in 6 M urea and lamins were partially purified by DEAE-cellulose ion exchange chromatography (see *2.4*). Fractions were electrophoresed through a 7.5%-15% SDS/polyacrylamide gradient gel resulting in this elution profile. The fraction numbers are indicated. **A**, **B** and **C** indicate lamins A, B and C. Note the separation of lamins A and C from B. Fractions enriched in lamins A and C (*i.e.* fractions 14-19) were utilized for protein sequencing (see *2.4*).


PCL
13
15
17
19
21
23
25
27
29
31
33
A
B
C

degradation with the Applied Biosystems model 470A gas-phase sequencer. The phenylthiohydantoin amino acids were identified and quantitated using the Hewlett Packard 1084 HPLC system.

## *2.5 Lamin Expression in E. coli*

A full-length lamin A cDNA inserted in the plasmid expression vector pUC18 (Yanisch-Perron *et al.*, 1985) was generated as follows. The partial lamin A cDNA was subcloned into the pUC18 BamHI site with its 5' end closest to the vector EcoRI site, as determined by restriction mapping. The full-length lamin C cDNA was subcloned into the pUC19 EcoRI site in both orientations. The lamin A/pUC18 and lamin C/pUC19 clones were cut with AvrII+EcoRI restriction enzymes (see Fig. 3). The EcoRI/AvrII fragment corresponding to the 5' end of lamin C was then ligated to the lamin A/pUC18 clone at the AvrII site near the 5' end of the lamin A insert (Fig. 3) and at the EcoRI site further upstream in the vector. This generated a full-length lamin A clone with the 5' end (EcoRI site) positioned near the β-galactosidase initiation *Met* codon (5 codons away) but out of reading frame.

To place the lamin A cDNA in reading frame with the first 5 codons of the β-galactosidase gene product the 5' end was cut with EcoRI and the protruding ends were digested with S1 nuclease (Berk and Sharp, 1977).

The lamin A/pUC 18 construction was transfected into and expressed in *E. coli* K12 RRIΔM15 (Ruther, 1982). Cultures were grown to saturation in Luria-Bertani (LB) medium (Maniatis *et al.*, 1982) containing 100 μg/ml ampicillin in the presence or absence of 2 mM isopropylthio-β-galactoside (IPTG), which induces expression of the β-galactosidase fusion protein. Cells were pelleted at 2000*g* for 5-10 minutes at 4°C, resuspended in 50 mM Tris-HCl, pH 7.6/50 mM KAc/1 mM (DTT)/1 mM phenylmethylsulfonylfluoride (PMSF) and lysed by sonication (45 sec. at micro tip setting of 3) or pressure (4000 psi) (Carver Laboratory Press, Model B) at

4°C. The lysate was centrifuged at 350,000*g* for 15 min at 4°C. and the pellet was resuspended in cold (4°C) 50 mM Tris, pH 7.6/1% Triton X-100/1 M NaCl/1 mM DTT/1 mM PMSF in a sonicating water bath and by vortexing. The solution was centrifuged as above, and the pellet was resuspended and recentrifuged in the Triton buffer (above). Supernatant fractions were precipitated with 20% TCA, washed with acetone (4°C) and analyzed along with pellet fractions on 8% SDS/polyacrylamide gels (Laemmli, 1970). Proteins were transferred to nitrocellulose and screened (Towbin *et al.*, 1979) with rabbit antibodies to the rat liver PCL fraction.

# Chapter 3

# Results

## *3.1 cDNA Cloning*

A rat liver λgt11 library and a human hepatoma λgt11 library were screened with a mixture of antibodies to lamins A, B and C. Nine positive recombinant phages were found, the largest of which (from the rat liver library) contained a 1.3 kb insert. A simple strategy was developed to obtain preliminary identification of the 1.3 kb cDNA. The cDNA sequence was induced and expressed in high abundance in bacteria as a hybrid polypeptide fused to β-galactosidase, and the bacteria were lysed and analyzed by SDS/PAGE (Fig. 5). The proteins were blotted onto nitrocellulose and screened with rabbit antibodies to the rat nuclear PCL fraction. The antibodies bound to the fusion protein only (Fig. 5). The antibodies to the fusion protein were then affinity purified from the nitrocellulose blot and were used to screen a blot containing the proteins of the rat nuclear PCL fraction. The affinity purified antibodies bound to lamins A and C (Fig. 6). Thus, one could conclude that the expressed clone, if not a genuine lamin clone, possessed antigenic epitopes in common with lamins A and C.

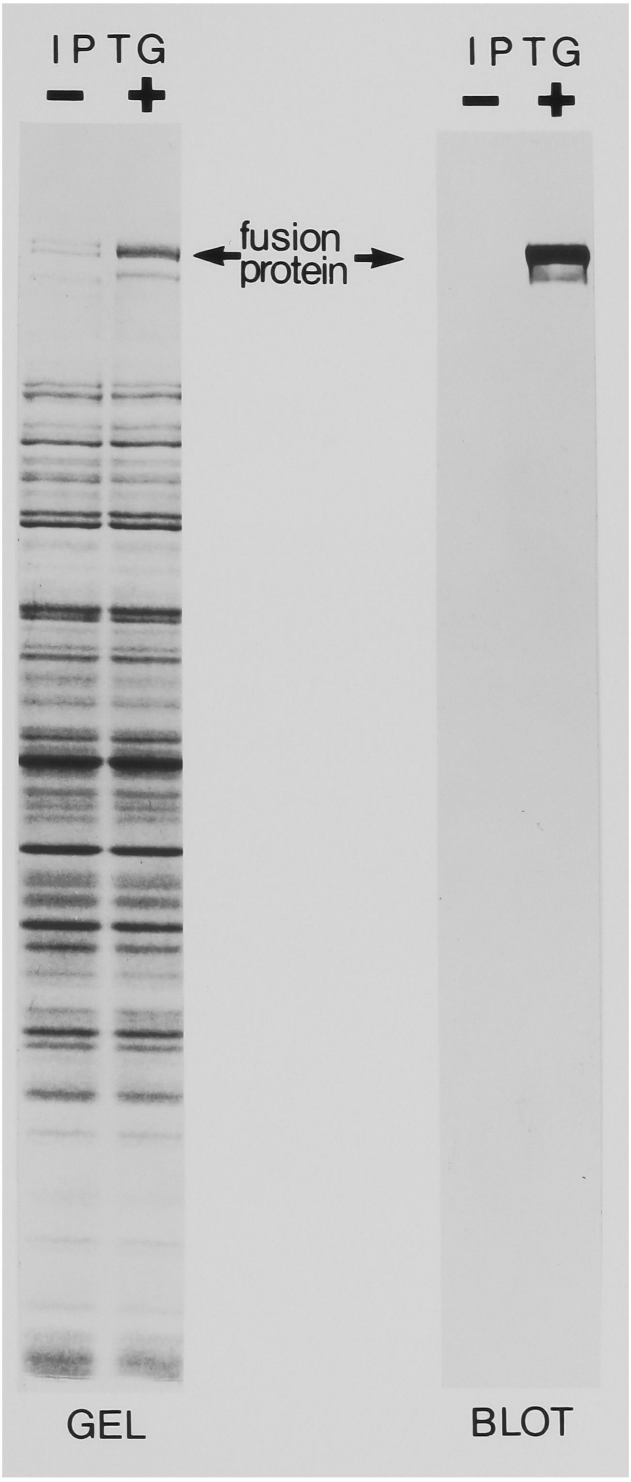
To determine the size of full-length lamin A and C clones, RNA blot analysis of poly(A)+ HeLa RNA was performed and revealed respective sizes of 2.9 kb and 2.0 kb (Fig. 7).

A human fibroblast library yielded a partial-length cDNA clone of lamin A (2.6 kb) when screened with the 1.3 kb rat liver cDNA. When rescreened with an oligonucleotide complementary to the 5' end of the partial lamin A clone, a 1.8 kb cDNA insert of lamin C was obtained. The fact that the lamin A oligonucleotide could pull out a lamin C cDNA was consistent with the observation that this oligonucleotide hybridized to the same two lamin mRNA bands as in Fig. 7 (data not

## *Figure 5*

*Antibodies to Rat Liver Nuclear PCL Fraction Bind to Fusion Protein*

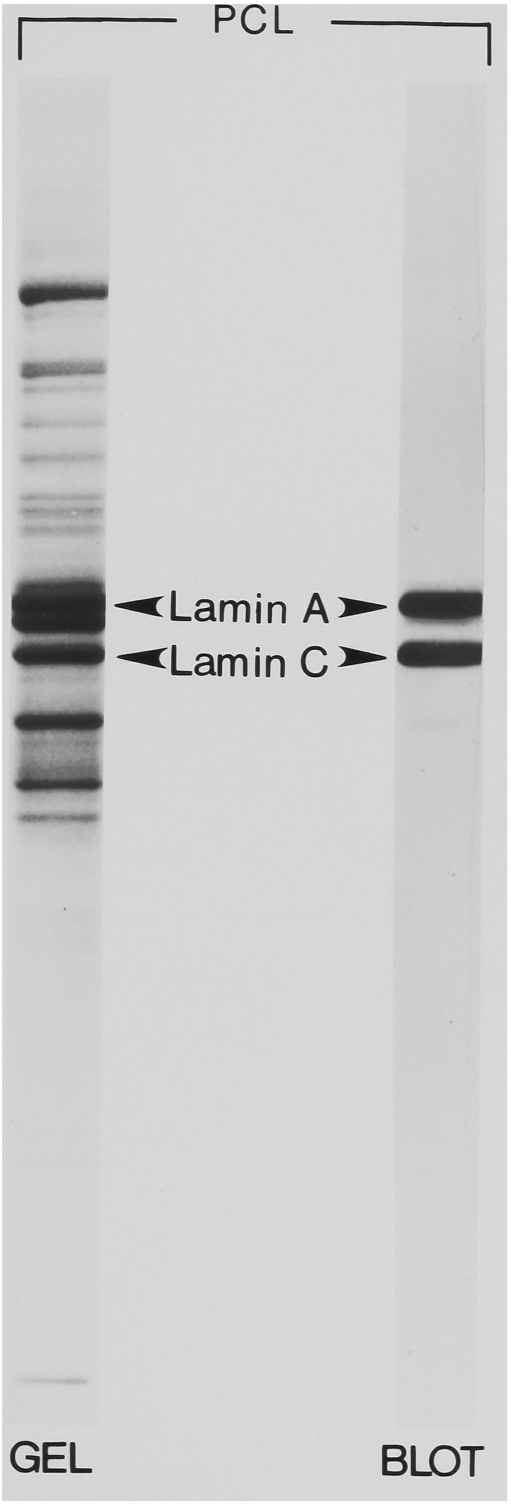
A rat liver cDNA λgt11 recombinant phage, containing a 1.3 kb cDNA insert, was lysogenized into host cells and induced to express the β-galactosidase fusion polypeptide by the addition of 5 mM IPTG. The bacterial lysate was run on a 7.5%-15% polyacrylamide gel (*left*), transferred to nitrocellulose and probed with antibodies to the rat liver PCL fraction (*right*). The antibodies bound specifically to the fusion protein. These bound antibodies were eluted from the nitrocellulose and were used to probe a blot containing proteins of the rat nuclear PCL fraction (see Fig. 6).


IPTG
− +
IPTG
− +
fusion protein
GEL
BLOT

## *Figure 6*

*Antibodies to Fusion Protein Bind to Lamins A and C*

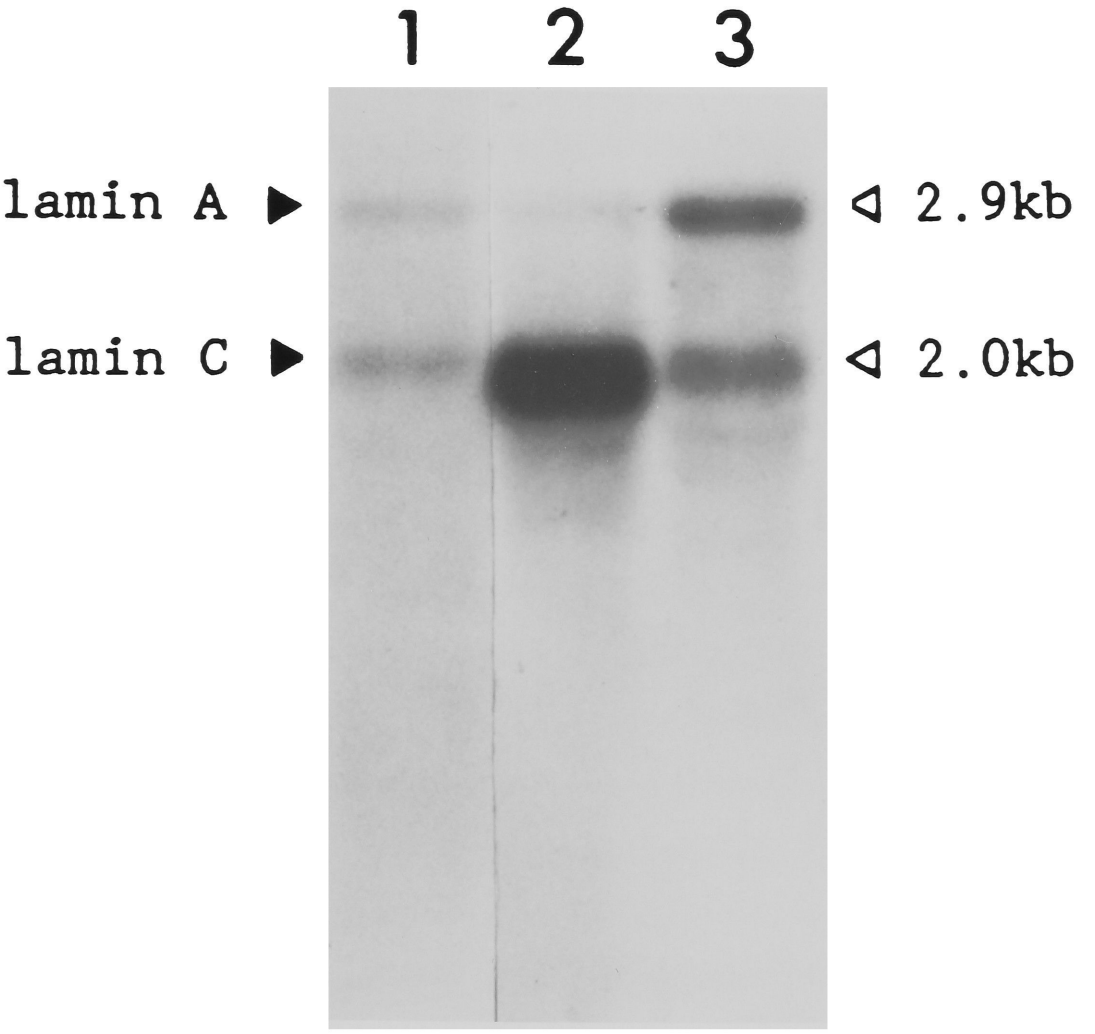
Affinity-purified antibodies to the expressed β-galactosidase fusion protein derived from a lysogen containing a rat liver cDNA λgt11 recombinant phage (see Fig. 5) were used to probe a nitrocellulose blot (*right*) containing proteins of the rat liver nuclear PCL fraction (transferred from the 7.5%-15% SDS/polyacrylamide gel, *left*). The antibodies bound to lamins A and C, indicating that the cloned cDNA, if not an actual lamin cDNA, expresses antigenic determinants in common with rat liver lamins A and C.


PCL
Lamin A
Lamin C
GEL
BLOT

## *Figure 7*

*Blot-hybridization Analysis of HeLa Cell Poly(A)+ RNA*

A single lane of RNA (25 μg) was hybridized with $^{32}$P-labeled nick-translated 1.3 kb rat liver cDNA alone (lane 1), with 1.3 kb rat liver cDNA, and then with actin cDNA (lane 2) or with 1.3 kb rat liver cDNA followed by actin cDNA and hsp 70 cDNA (lane 3). The actin (2.0 kb) and hsp 70 (2.9 kb) probes were used as internal molecular weight markers. Lamin A and C mRNA bands are seen in lane 1, the lamin A mRNA band is still visible in lane 2, but both lamin mRNAs are covered by the two molecular weight probes in lane 3. Similar results were obtained with a $^{32}$P-labeled oligonucleotide probe (22mer, see *2.1* and *3.1*).


1
2
3
lamin A
lamin C
2.9kb
2.0kb

shown). A putative full-length lamin C clone of 1962 bp was obtained by screening a human hepatoma λgt11 library (different from the one originally screened, see above and *2.1*) with an oligonucleotide complementary to the 5' region of the partial lamin C clone.

### *3.2 DNA Sequence Analysis*

Analysis of the two cloned cDNA sequences of lamin C (from both the fibroblast library and the hepatoma library) (Fig. 8) and the partial cDNA sequence of lamin A (from the fibroblast library) (Fig. 8) revealed complete identity in nucleotide sequences of lamins A and C with the exception of their 3' ends. Recent work (McKeon *et al.*, 1986) has shown, in addition, that the 5' end of full-length lamin A is identical to that of lamin C.

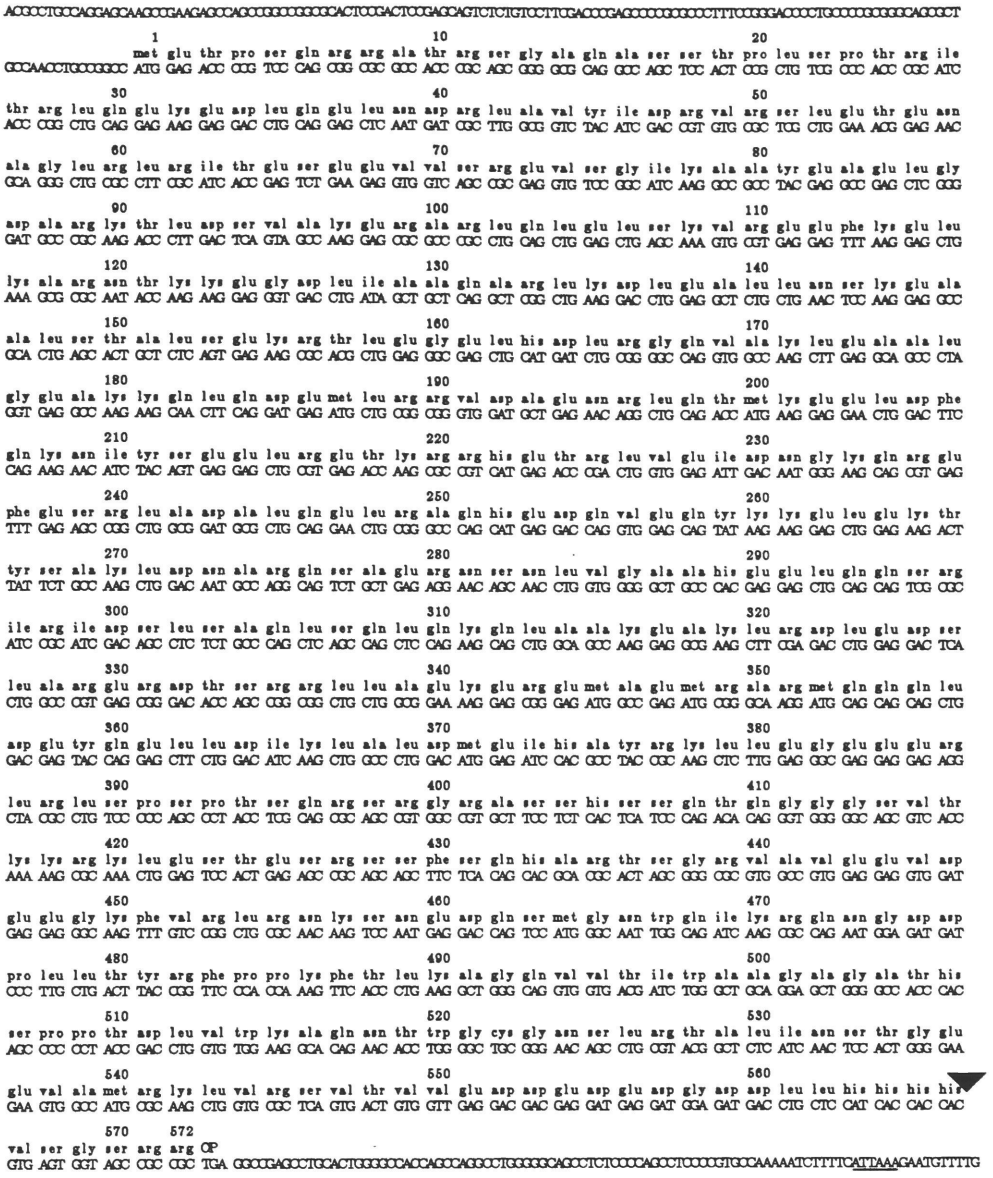
The cDNA sequence of lamins A and C (Fig. 8) shows a single open reading frame that contains an initiation codon, ATG, 135 bp from the 5' end. As there are no in-frame stop codons upstream from the ATG codon, the possibility remains that the cDNA is incomplete. However, the size of the clone (1962 bp) agrees with the predicted size based on RNA blot analysis (Fig. 7) and the initiating ATG codon is surrounded by sequences that resemble the most common eukaryotic consensus sequence, AXXATGG (Kozak, 1984). The upsteam A in the consensus sequence is replaced by a G in the lamin sequence (Fig. 8), but other eukaryotic mRNAs have a G in the same relative location (Hudson *et al.*, 1981). The 3' polyadenylation signal, AATAAA, of most eukaryotic genes is slightly modified in both the lamin A (AATCAA or AATACA) and lamin C (ATTAAA) cDNAs (Fig. 8).

The lamin A and C cDNA sequences diverge following a CAC codon corresponding to amino acid residue 566 (Fig. 8). The following six nucleotides in the lamin C sequence, GTGAGT (Fig. 8), are identical to the consensus sequence of the 5' end of introns (Cech, 1983). The two nucleotides following codon 566 in

## *Figure 8*

*cDNA Sequences and Deduced Amino Acid Sequences of Lamin C and Lamin A*

DNA sequencing was performed by the dideoxy chain-termination method on both strands of all cDNAs and by the chemical method on selected regions. Both of the isolated lamin C cDNA clones (1.8 kb and 1.9 kb) were sequenced as was the partial lamin A cDNA clone (2.6 kb). The 3' polyadenylation signals are underlined. Residue 150 is the first deduced amino acid of the incomplete lamin A cDNA sequence, and residue 566 (closed arrowhead) is the last amino acid shared by both lamins A and C before their sequences diverge. The six nucleotides following codon 566 in lamin C, GTGAGT, represent the consensus sequence of the 5' end of introns (Cech, 1983). Residues 551-560 at the carboxyl end of the lamin C protein (but within the carboxyl domain of the lamin A protein) are highly acidic.



## Lamin A Carboxyl Terminus

lamin A, GG (Fig. 8), represent a consensus sequence of the 5' end of spliced exons (Cech, 1983). These sequences suggest that the two lamin mRNAs may result from alternate forms of transcriptional processing (see *4.5*).

The predicted $M_r$ of the encoded lamin C protein is 65,145, slightly larger than its apparent $M_r$ of 60,000 as estimated by SDS/PAGE (Gerace and Blobel, 1980) and in agreement with McKeon *et al.* (1986).

The complete lamin A sequence (Fig. 8) has a predicted $M_r$ of 74,152. This size is in close agreement with the apparent $M_r$ of 72,000 of the lamin A precursor as estimated by SDS/PAGE (Gerace and Blobel, 1980) and is about 5300 daltons smaller than the size predicted by McKeon *et al.* (1986).

### *3.3 Protein Purification and Sequencing*

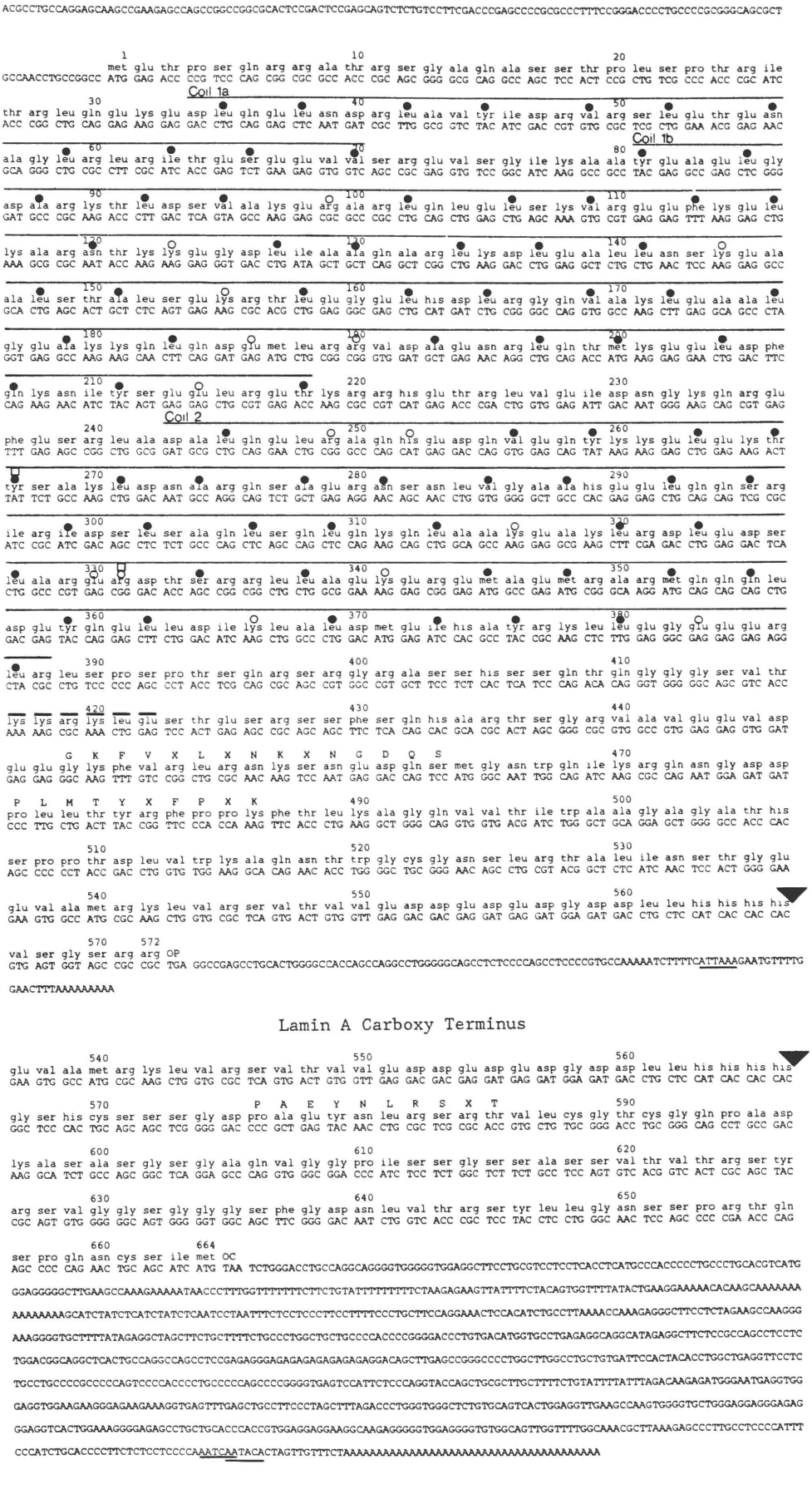
In order to confirm the identities of the cloned cDNAs, amino terminal sequencing of purified rat liver lamin A and C polypeptides was attempted. Both proteins were found to be blocked. Therefore protease fragments of the lamins were generated and sequenced. Since lamin A and lamin C have common fragments when subjected to V8 protease digestion (Gerace and Blobel, 1982), a 14 kDa fragment shared by both rat liver lamins was generated by V8 digestion and sequenced. Fifteen amino acids sequenced from this rat liver fragment, *Gly-Lys-Phe-Val-X-Leu-X-Asn-Lys-X-Asn-Gly-Asp-Gln-Ser*, corresponded to residues 449-463 in human lamins A and C (Fig. 9). This sequence immediately follows a *Glu* residue, consistent with V8 protease cleavage after *Glu*.

In addition to V8 digestion, both lamins were subjected to mild acid hydrolysis to cleave between *Asp-Pro* residues. Only one sequence was obtained from lamin C as predicted by the nucleotide sequence (Fig. 9) since there is only one *Asp-Pro* site in the predicted lamin C sequence and the amino terminus was blocked. This sequence, *Pro-Leu-Met-Thr-Tyr-X-Phe-Pro-X-Lys*, corresponded to residues

## *Figure 9*

*Analysis of Deduced Amino Acid Sequences of Lamin C and Lamin A*

The deduced lamin A and lamin C amino acid sequences are identical up to residue 566 (closed arrowhead) after which the lamin A sequence continues with 98 amino residues and the lamin C sequence continues with 6 amino acid residues. Coils 1a, 1b and 2 denote α-helical domains with heptad repeats capable of forming coiled coils. At the first and fourth positions of the heptads, closed circles mark hydrophobic and nonpolar residues, and open circles mark polar residues. Open boxes (residues 267 and 331) indicate phase changes in the repeating heptad sequences. Amino acid residues sequenced from V8-digested and formic acid-hydrolyzed rat lamin A and C fragments are indicated by single letter amino acid symbols above the corresponding amino acids predicted by the human cDNA sequences. "X" denotes amino acids not detected. A stretch of six amino acid residues (residues 417-422, broken line), similar to the nuclear localization signal of the simian virus 40 large T antigen (Kalderon *et al.*, 1984), is also indicated. The 3' polyadenylation signals are underlined.




## Lamin A Carboxy Terminus

477-486 (Fig. 9). The third amino acid in this sequence, *Met*, differed from the corresponding amino acid, *Leu*, and represents a single nucleotide switch of *A* to *C* in the first codon position. This same sequence was obtained from hydrolyzed lamin A; in addition, a second sequence was obtained from lamin A corresponding to the only other *Asp-Pro* site in the predicted lamin A sequence. This sequence, *Pro-Ala-Glu-Tyr-Asn-Leu-Arg-Ser-X-Thr*, corresponded exactly to residues 576-585 in the predicted lamin A sequence (Fig. 9). The last three residues of this sequence (583-585) (Fig. 9) confirmed the predicted lamin A sequence where it diverged from that of McKeon *et al.* (1986), who reported an additional nucleotide (cytosine) following codon 582. Thus, the lamin A sequence extends only 82 residues from residue 583 as opposed to 120 (McKeon *et al.*, 1986) residues.

These protein sequence data clearly established the identity of the two lamin clones and indicated that the lamin protein sequence is at least partly conserved from rat to human.

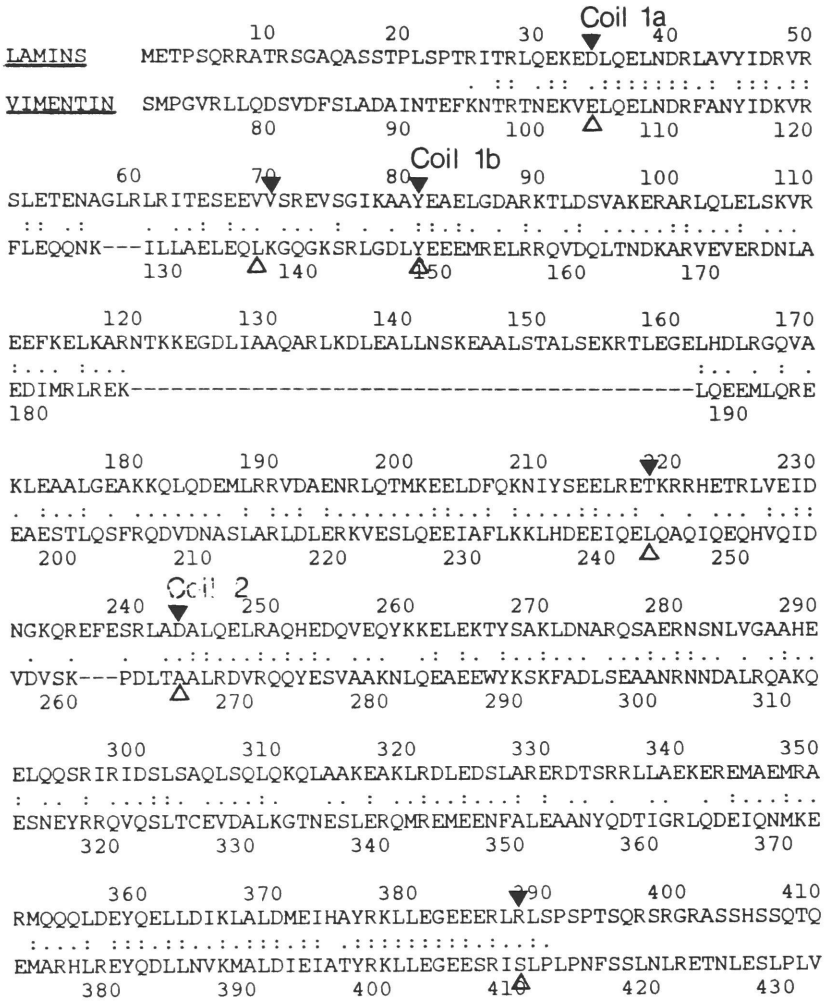
### *3.4 Lamin Homology to Intermediate Filament Proteins*

When the lamin A and C sequences were compared with other known protein sequences, a striking sequence homology was found with all types of intermediate filament proteins. In Fig. 10, the sequences of lamins A and C are compared to that of hamster vimentin (Quax *et al.*, 1983) (serving as a representative of the family of intermediate filament proteins). To achieve maximum alignment of lamins A/C and vimentin, three gaps (3, 42, and 3, respectively) were introduced into the vimentin sequence. An overlap of 366 residues between lamins A and C and 318 residues of vimentin showed 27.6% identity. The amino terminal region of this overlap, comprising 20 residues, showed greater than 70% homology (Fig. 10). The carboxy-terminal end of this overlap, comprising 35 residues, also showed greater than 70% homology (Fig. 10). When the lamin sequences were compared to those of other

## *Figure 10*

*Primary and Secondary Structural Homology between Lamins A/C and Vimentin*

The deduced amino acid sequence common to lamins A and C was aligned to maximize homology with that of hamster vimentin (Quax *et al.*, 1983). The lamins shared 27.6% identity with vimentin over a 366 amino acid residue overlap. A gap of 42 residues was introduced into the vimentin sequence to optimize this homology. This region of sequence homology coincides with predicted α-helical domains capable of forming coiled coils present in lamins A/C (between closed arrowheads) and vimentin (between open arrowheads). The first 20 and last 35 amino acids within this stretch each show greater than 70% sequence homology.


Coil 1a
LAMINS
VIMENTIN
Coil 1b
Coil 2

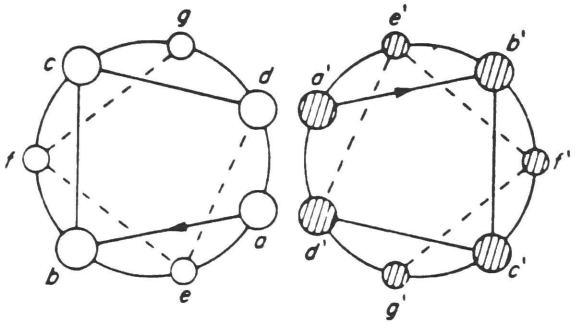
intermediate filament proteins, the sizes of the homologous stretches and percent homology were consistent with that of vimentin (data not shown).

Chou and Fasman analysis (Chou and Fasman, 1978) to predict the secondary structure of lamins A and C revealed a 360-residue stretch of α-helical conformation extending from residue 31 to residue 390 (Fig. 9). Most strikingly, this region of predicted a–helical conformation coincides closely with the region of sequence homology between lamins A and C and vimentin (Fig. 10) and that of other intermediate filament proteins (Geisler *et al.*, 1984) (data not shown). An approximately 310-residue central domain of about 30% sequence homology and α-helical conformation is shared by all intermediate filament proteins (reviewed by Fuchs and Hanukoglu, 1983). The lamins contain a repeating heptad sequence within this a–helical domain of the form, $(a\text{-}b\text{-}c\text{-}d\text{-}e\text{-}f\text{-}g)_n$, where residues *a* and *d* are usually hydrophobic residues. In Fig. 9, this repeating heptad organization is labeled at positions *a* and *d* by closed or open circles, representing apolar or hydrophilic residues, respectively. Three regions within the α-helical domain, containing these heptad repeats, are separated by short stretches of amino acid residues that break up this repeating pattern. The α-helical domain of all intermediate filament proteins contains the same repeating heptad organization as the lamins. The three regions of heptad repeats are labelled Coil 1a, Coil 1b, and Coil 2 in Figs. 9 and 10 since such domains are thought to be capable of forming coiled coils with analagous heptad-repeating α-helical domains of other molecules (Crick, 1953). Residues *a* and *d* of each repeating heptad are believed to interact with residues *a'* and *d'* of the heptads of an adjacent molecule to form a hydrophobic seam between the two molecules (Fig. 11). This places the hydrophobic residues within the core of the coiled coil and places the hydrophilic residues at the surface, thus stabilizing the structure within an aqueous milieu. The regions between the heptad repeats, although α-helical in the lamin A/C sequence, are not believed to be capable of forming coiled coils. Within

## *Figure 11*

*End-on View of a Coiled Coil*

Two α-helices interlock to form a coiled coil. Heptad positions *a* and *d* interlock with *a'* and *d'* stabilizing the structure. These amino acid residues are usually hydrophobic or nonpolar and form a greasy seam within the core of the structure. Hydrophilic residues occupy the surface. This view looks from the amino to the carboxyl end. The chain are in register and parallel.




McLachlan and Stewart, 1975

the intermediate filament family, these so-called "linker" regions are not α-helical, and thus are also not capable of forming coiled coils. Coil 2 contains two phase changes in its repeating heptad sequence at residues 267 and 331 (Fig. 9).

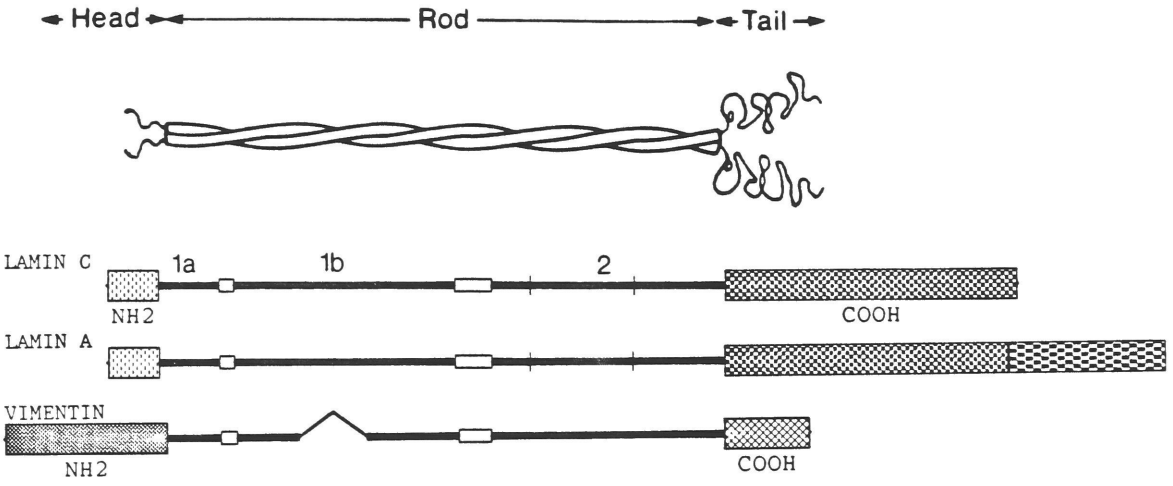
Fig. 10 reveals the precise alignment of these coiled coil forming regions between lamins A/C and vimentin (delimited by closed and open triangles for the lamins and vimentin, respectively) with the exception of Coil 1b, which is about 50% longer in the lamins than in vimentin (and all other intermediate filament proteins, data not shown). The entire α-helical domain of the lamins is therefore approximately 14% larger than the α-helical domain of intermediate filament proteins. Flanking the a-helical domain are an amino terminal "head" domain and a carboxy terminal "tail" domain which are variable in size and composition in the intermediate filament family. The lamin A/C head domain is approximately 4 kDa, while the tail domains of lamins A and C are approximately 30 kDa and 21 kDa, respectively (Fig. 9). By Chou and Fasman analysis these end domains were predicted to contain secondary structures of all types. The central domain is termed the "rod" domain based on studies of intermediate filaments where proteolytic cleavage of these proteins produced a morphological rod-like structure corresponding to the highly α-helical domain (Geisler *et al.*, 1982). Fig. 12 depicts this structural organization schematically.

The lamin A tail has a glycine- and serine-rich stretch of residues (571-652) and also shares approximately 20%-35% sequence homology with amino- and carboxyl-terminal portions of a human type II epidermal keratin (67 kDa) (Johnson *et al.*, 1985), with a mouse type II keratin (60 kDa) (Steinert *et al.*, 1984), and with mouse (59 kDa) and human (52 kDa) type I keratins (Hanukoglu and Fuchs, 1982; Steinert *et al.*, 1983) (Fig. 13).

## *Figure 12*

*Common Structural Model of Lamins A/C and Vimentin*

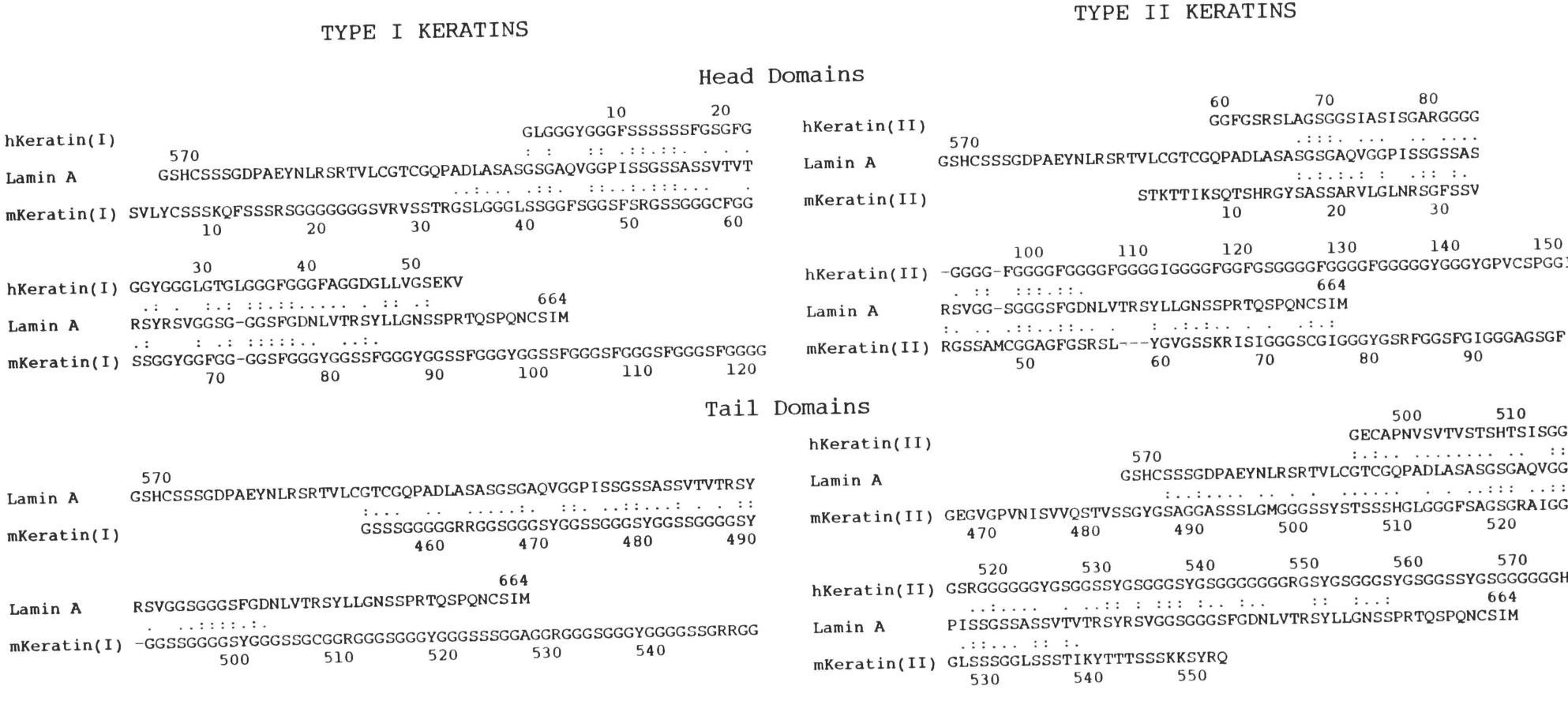
(*Top*) The α-helical rod domain is flanked by hypervariable non-helical head and tail domains located at the amino- and carboxy-terminal ends, respectively. The helices of the coiled coil dimer are parallel and in register. (*Bottom*) Amino, carboxyl, and coiled coil domains of lamins A and C were aligned with the predicted secondary structure of hamster vimentin (Geisler *et al.*, 1984). All molecules are drawn to scale. Lamins A and C have identical sequences spanning the central α-helical domain (black lines) and the carboxyl domain of lamin C. The lamin C carboxyl domain extends 6 amino acids beyond the point of divergence from lamin A whereas the lamin A carboxyl domain extends 98 amino acids further (see Fig. 8). Putative globular domains are located at both ends of all three molecules. The central α-helical domains are 14% longer in lamins A/C than in vimentin. The coiled coil-forming subdomains of the rod (1a, 1b and 2) are separated by intercoil linkers. Coils 1a and 2 and the two intercoil linkers are conserved in size between the lamins and vimentin whereas Coil 1b is about 50% longer in the lamins than in vimentin. Coils 1 and 2 are nearly equal in length, about 21 nm in intermediate filament proteins, and about 25 nm and 21 nm, respectively, in the lamins. Vertical lines mark stutters in the heptad repeats.


Head
Rod
Tail
LAMIN C
1a
1b
2
NH2
COOH
LAMIN A
VIMENTIN
NH2
COOH

## *Figure 13*

*Lamin A Sequence Homology to Type I and Type II Keratin Head and Tail Domains*

A 92 amino acid residue stretch (residues 571-663) in the carboxyl-tail region of lamin A, not present in lamin C (see Fig. 8), contains regions showing between 20% and 35% identity with head and tail regions of human type I (52 kDa) and type II (67 kDa) epidermal keratins (Hanukoglu and Fuchs, 1982; Johnson *et al.*, 1985), and with mouse type I (59 kDa) and type II (60 kDa) keratins (Steinert *et al.*, 1983; Steinert *et al.*, 1984). Four of the five lamin A *Cys* residues are located in this tail domain.


TYPE I KERATINS
Head Domains
hKeratin(I)
Lamin A
mKeratin(I)
hKeratin(I)
Lamin A
mKeratin(I)
Tail Domains
Lamin A
mKeratin(I)
Lamin A
mKeratin(I)
TYPE II KERATINS
hKeratin(II)
Lamin A
mKeratin(II)
hKeratin(II)
Lamin A
mKeratin(II)
hKeratin(II)
Lamin A
mKeratin(II)

### *3.5 Nuclear Localization Site*

Lamins A and C contain a sequence of 6 amino acids (*Lys-Lys-Arg-Lys-Leu-Glu*) that is similar to a stretch of amino acids in the simian virus 40 large T (tumor) antigen (*Lys-Lys-Arg-Lys-Val-Glu*) that contributes to nuclear localization of the viral protein (Kalderon *et al.*, 1984). Located approximately 30 amino acids outside the carboxyl end of the α-helical domain (Fig. 9), the lamin sequence differs from the viral sequence only by the substitution of *Leu* for *Val* in the fifth position. Although it has been shown that the consensus sequences *Lys-Lys-Arg-Lys-Val-Glu-Asp-Pro* and *Pro-Lys-Lys-Lys-Arg-Lys-Val* can each promote nuclear accumulation of various proteins (Kalderon *et al.*, 1984), their common sequence, *Lys-Lys-Arg-Lys-Val*, has not been tested. With the conservative change of *Val* to *Leu*, the lamins may represent examples of cellular nuclear proteins targeted to the nucleus by an internal karyophilic signal sequence.

### *3.6 Expression of Lamin A in E. coli*

A full-length lamin A cDNA was constructed in order to have both complete lamin A and lamin C cDNA clones. The full-length lamin A cDNA was engineered into the bacterial expression vector, pUC18, and expressed in *E. coli* (see *2.5*) in order to verify that the cDNA was translation-competent and in order to determine the appropriate conditions for engineering the cDNA into other expression systems (*i.e.* eukaryotic expression vectors). As shown in Fig. 14 (gel), the protein is expressed in abundance when induced with 2 mM IPTG and is clearly visible as a heavy band with an apparent $M_r$ of approximately 80 kDa. Antibodies to the rat nuclear PCL fraction react exclusively with this protein in the bacterial lysate and do not react with any bacterial proteins (Fig. 14, blot). However, upon overexposure, a band corresponding to the expressed 80 kDa protein is present in the lane of the cell lysate harboring the cDNA-containing plasmid, not induced with IPTG (Fig. 14, lane

## *Figure 14*

*Lamin A Expression in E. coli*

A full-length lamin A cDNA was constructed in the bacterial expression vector, pUC18. The recombinant plasmid was introduced into *E. coli* strain K12 RR1ΔM15 and was induced for protein expression. 5 ml cultures were grown for 10-12 hrs at 37°C in the presence (lanes 3 and 4) or absence (lanes 2 and 5) of 2 mM IPTG. Cells were pelleted, resuspended in 250 µl 6% SDS/0.5 M Tris base, freeze-thawed and mixed with 0.2 M DTT/40% glycerol/0.16% bromphenol blue/0.01% $NaN_3$. Samples were boiled for 3 min and electrophoresed on an 8% SDS/polyacrylamide gel (*left*). 20 µl were loaded per lane. Molecular weight markers (kDa) are indicated.

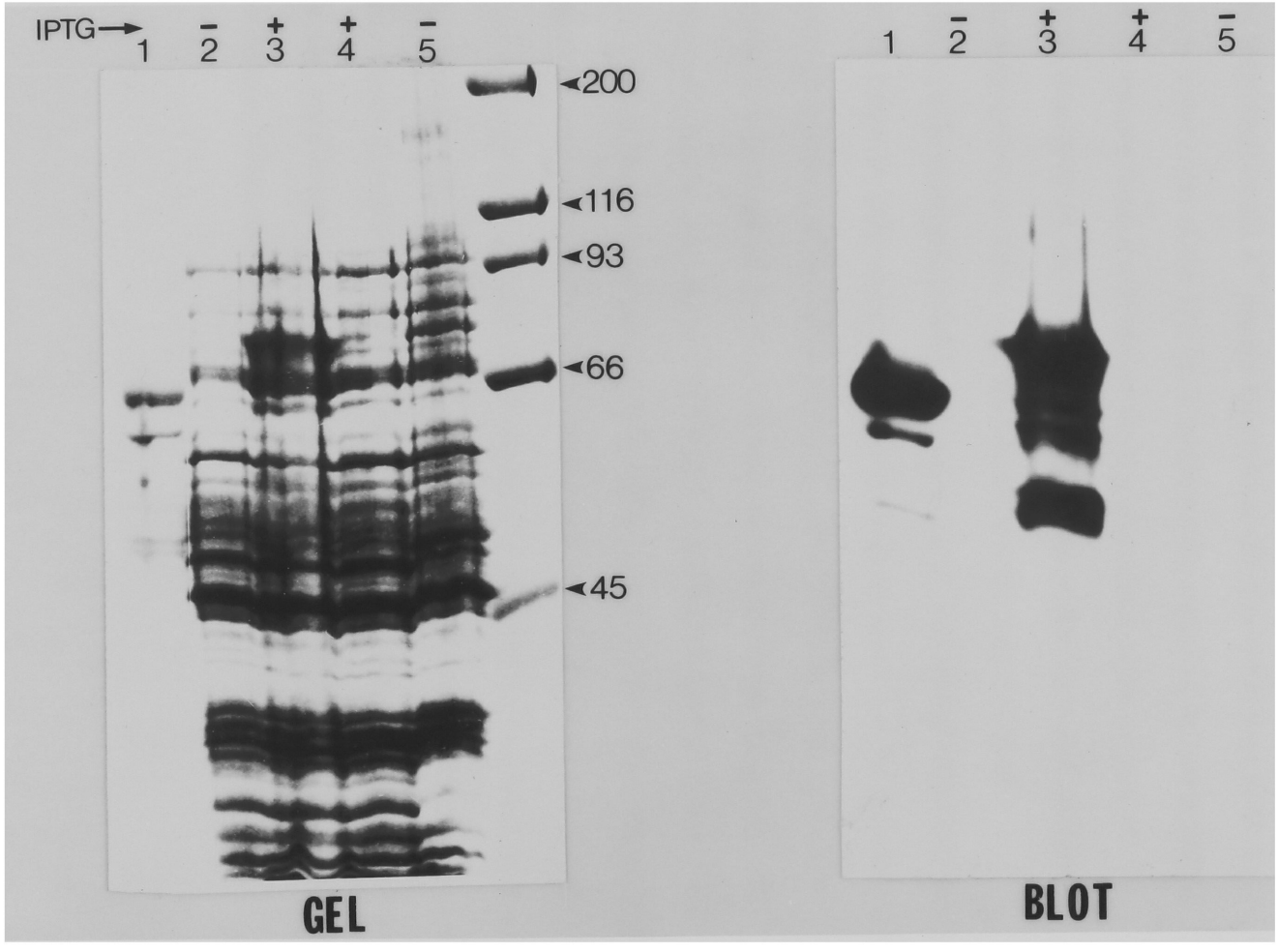
*Lane 1*.....rat liver nuclear lamins

*Lanes 2 and 3*.....*E. coli* transformed with pUC18/lamin A

*Lane 4*.....*E. coli* transformed with pUC18/lamin A (out of frame)

*Lane 5*.....*E. coli* strain K12 RR1ΔM15

Proteins from the gel were blotted onto nitrocellulose and probed with antibodies to the rat nuclear PCL fraction (*right*). The antibodies bound to the expressed lamin A protein, which is about 6 kDa larger than predicted by the cDNA sequence. This is due to translation of the 5' untranslated region of the cDNA which is in frame with the initiation ATG, plus 5 additional codons from the 5' end of the β-galactosidase gene. Note the extensive proteolysis of the expressed protein.


IPTG
1
2
3
4
5
200
116
93
66
45
GEL
1
2
3
4
5
BLOT

2, data not shown). Thus, there are low levels of β-galactosidase-related gene expression even in the absence of IPTG induction. Considerable proteolysis of the expressed lamin protein is also apparent by immunoblotting (Fig. 14).

The expressed lamin A* protein migrates at 80 kDa rather than at the expected 74 kDa due to expression of the 135 bp 5' untranslated region of the lamin cDNA (Fig. 8) plus 5 codons from the 5' end of the β-galactosidase gene of pUC18. As mentioned previously (*3.2*), the 5' untranslated region of the lamin A/C cDNAs contains no stop codons in frame with the initiating ATG codon (Fig. 8).

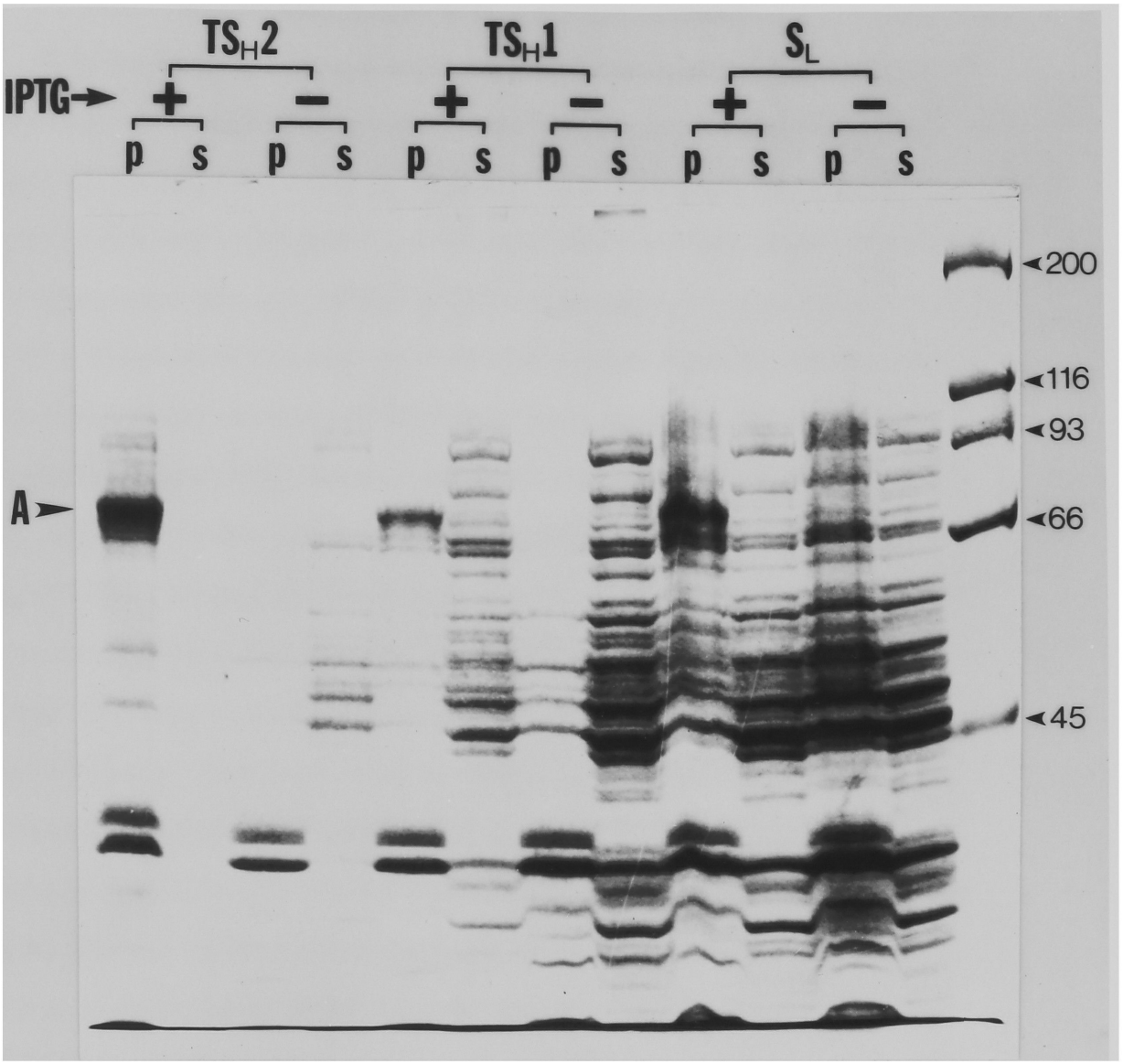
In order to develop a scheme for obtaining large amounts of expressed lamin A* protein for future *in vitro* analysis of lamin protein domains (*i.e.* by deletion- and site-directed-mutagenesis of the lamin cDNAs), and to examine some properties of this protein in *E. coli*, a partial purification scheme was developed (Fig. 15). The lamin protein was found to be highly insoluble in a buffer containing Triton X-100 and 1 M NaCl, enabling it to be potentially purified based on this insolubility. In this regard, the expressed protein exhibits physical properties similar to those of native lamin proteins, despite the addition of 50 amino acid residues at the amino terminus of the expressed protein. However, although this insolubility may reflect formation of a genuine lamin filament structure within *E. coli*, it may also reflect aggregation of overexpressed protein material, a common occurrence in bacterial expression systems (see, for example, Eilers and Schatz, 1986).

## *Figure 15*

*Partial Purification of Lamin A Expressed in E. coli*

A full-length lamin A cDNA engineered into the bacterial expression vector, pUC18, was expressed in *E. coli* strain K12 RR1ΔM15 in the presence (**+**) or absence (**-**) of 2 mM IPTG. 5 ml cultures grown for 10-12 hrs were spun down at 2000*g* for 10 min at 4°C. Pellets were resuspended in 1 ml of 50 mM KAc/50 mM Tris-HCl, pH 7.6/1 mM DTT/1 mM PMSF (**$S_L$**), sonicated and centrifuged at 350,000*g* for 15 min. Pellets were resuspended in 1 ml of 1% Triton X-100/1 M NaCl/1 mM DTT/1 mM PMSF (**$TS_H1$**) and centrifuged as above. Pellets were resuspended in the Triton buffer (**$TS_H2$**) and centrifuged again. Supernatant proteins were TCA-precipitated and all pellets were solubilized in 100 µl 6% SDS/0.5 M Tris base plus 75 µl 0.2M DTT/40% glycerol/0.16% bromphenol blue/0.01% $NaN_3$. 10 µl samples were loaded per lane and run on an 8% SDS/polyacrylamide gel. Molecular weight markers (kDa) are indicated. Note the appearance of the expressed lamin A protein (**A**) in the pellet fractions.

**s** = *supernatant* **p** = *pellet*


TS$_H$2
TS$_H$1
S$_L$
IPTG
+
−
p
s
A
200
116
93
66
45

# Chapter 4
# Discussion

## *4.1 Lamins are Intermediate Filament Proteins*

The intermediate filament proteins are a large multigene family of proteins consisting of 19 nonepidermal epithelial keratins (40-68 kDa), 8 "hard" keratins of hair and nail forming cells, vimentin (54 kDa, in cells of mesenchymal origin), desmin (53 kDa, in muscle), glial fibrillary acidic protein (51 kDa, in astrocytes) and 3 neurofilament proteins (63, 160 and 200 kDa, in neurons) (reviewed by Steinert *et al.*, 1985). Based on various analyses of their primary and secondary structure, all intermediate filament proteins have been shown to be comprised of a common molecular architecture (Fig. 12). Each has a central α-helical rod domain of approximately 310 amino acid residues with heptad repeats capable of forming a two-stranded coiled coil (see *4.3*). This domain is subdivided by two non-α-helical linker regions into three subdomains termed Coils 1A, 1B and 2 (Fig. 12) The overall length of the rod domain is 45-48 nm with Coils 1 and 2 of approximately equal length (21-23 nm). The central domain is flanked by end domains of variable sizes and chemical character which are presumed to account for the variability in size, physical properties and physiological functions of the different intermediate filament subunits. The central domain contains highly conserved regions of sequence homology at the beginning of Coil 1A and at the end of Coil 2; the positions of the linker regions as well as a phase change in the heptad repeat pattern of Coil 2 are also highly conserved.

Lamins A and C can be classified as members of the intermediate filament protein family based on several characteristic features of their amino acid sequences. The lamins share an α-helical region of approximately 360 amino acid residues that, in turn, shows striking homology with the canonical 310-residue α-helical rod

domain characteristic of all intermediate filament proteins (Fig. 10). Like the intermediate filament proteins, these lamin domains contain heptad repeats capable of forming coiled coils separated by linker regions that, although α-helical, are not capable of forming coiled coils. These linkers, as well as the phase changes in the heptad repeat pattern of Coil 2, are located in the conserved positions characteristic of all intermediate filament protein rod domains. The lamins share approximately 30% sequence homology with the other classes of intermediate filament proteins across their α-helical domain; approximately 30% homology across this domain is found between all classes of cytoplasmic intermediate filament proteins, with desmin, vimentin and glial fibrillary acidic protein sharing 60%-70% identity across this region (reviewed by Fuchs and Hanukoglu, 1983). The amino acid sequences located at the beginning of Coil 1A and at the end of Coil 2 are also highly conserved (> 70% identity) between the lamins and all intermediate filament proteins.

Structural and biochemical analyses of lamin structures formed *in vivo* and *in vitro* are in agreement with predictions of the lamin sequence data. Isolated lamins A and C as well as B from rat liver have been shown by electron microscopy to be rod-shaped dimers comprised of two globular heads attached to an approximately 50 nm rod-like tail (Aebi *et al.*, 1986). Based on the sequence data, the globular head domains probably correspond to the approximately 20-30 kDa carboxy-terminal portions of the lamin polypeptides, which are are larger than the approximately 4 kDa amino-terminal portions. The 50 nm rod domain probably corresponds to the lamin α-helical domain. The observed length of this rod domain (50 nm) is larger than that of cytoplasmic intermediate filament proteins (45-48 nm) probably due to the larger Coil 1B subdomain of the lamins, which contains 42 more amino acid residues than their cytoplasmic counterparts (Figs. 10 and 12). These isolated lamin dimers have Stokes radii of 10-11 nm (Aebi *et al.*, 1986). Sedimentation studies of lamin tetramers indicate that they form 4.5S-4.9S species (Aebi *et al.*, 1986), values that

suggest that the 4-5S solubilized lamins of mitotic cells may be dimers or tetramers (see *1.4*). However, since the mitotic lamins were sedimented in the presence of the mitotic extract plus Triton X-100 (Gerace and Blobel, 1980), these sedimentation values cannot be directly compared to those of the isolated rat liver lamins which were sedimented under very different conditions. *In vitro* reconstitution experiments (Aebi *et al.*, 1986) have shown that under certain conditions, such as pH 6.5 and higher than physiological salt concentrations, the lamin dimers can assemble into long filaments approximately 10 nm in diameter which are indistinguishable from typical intermediate filaments. Decreasing the ionic strength yields thicker paracrystals with an approximately 24-25 nm axial periodicity, slightly larger than that seen for intermediate filaments in general (21-23 nm) (see *4.3*). Again, this larger periodicity present in lamin filaments can be accounted for by the larger Coil 1B subdomain. Detergent-extracted nuclear envelopes from *Xenopus laevis* were shown to contain a lamina composed of a near-orthagonal meshwork of approximately 10nm filaments with a 52 nm crossover spacing and the same axial periodicity (25 nm) as that formed by *in vitro* reconstituted lamin filaments (Aebi *et al.*, 1986). Thus, the *in vivo* and *in vitro* data seem consistent with each other and with the predictions from the sequence data. However, there are some noteworthy differences between the *in vivo* and *in vitro* data as well. First, the *Xenopus* oocyte lamina consists of only one lamin species, $L_{III}$, while the rat liver lamina consists of three lamins. Thus only one lamin is required for *Xenopus* oocyte lamina assembly while at least two (lamins A and C) are apparently utilized for rat lamin filament assembly, based on the *in vitro* data (Aebi *et al.*, 1986). The near-orthagonal lamin lattice of oocyte nuclear envelopes also may or may not be a feature common to all lamina structures since it has yet to be observed in other nuclear envelopes, such as in rat liver. Furthermore, $L_{III}$ disassembles into a soluble form that sediments at 9S, not 4-5S as with rat lamins (see *1.4* and *4.4*), although these studies were carried out

under different conditions.

A final characteristic that lamins share with intermediate filament proteins is that they all belong to developmentally regulated gene families (intermediate filament families reviewed by Franke *et al.*, 1982; lamin families, see *1.3*). With the present availability of lamin cDNA probes it should now be feasible to study the developmental regulation of lamin expression more precisely than previous immunological studies by monitoring RNA levels throughout development.

In summary, the sequence data together with the structural data clearly indicate that the lamins are a newly discovered class of intermediate filament proteins.

### *4.2 Lamins Differ from Intermediate Filament Proteins*

In spite of the structural relatedness between the lamins and the intermediate filament proteins, there are some noteworthy differences. As mentioned above (*see 4.1*), the lamin α-helical domain is approximately 14% larger than the rod domain of intermediate filament proteins due to a longer Coil 1B domain (Figs. 10 and 12). This is reflected by the crossover spacing of 50-52 nm in the lamina of *Xenopus* oocytes and the 50 nm rod tail of *in vitro* isolated lamins as compared with the 45-48 nm rod domain of typical intermediate filament proteins. Also, the larger axial periodicity (25 nm) of native and *in vitro* reconstituted lamin filaments versus cytoplasmic intermediate filaments (21-23 nm) (see *4.3*) reflects this enlarged lamin domain. Interestingly, the extra 42 amino acid residues located in Coil 1B of the lamins (Fig. 10) are located precisely at the position of an intron conserved in all non-neuronal classes of intermediate filaments (Weber, 1986). This suggests that the lamins and intermediate filament proteins share a common evolutionary precursor. Lamins probably were expressed earlier than cytoplasmic intermediate filament proteins in the evolutionary tree as eukaryotes acquired their distinguishing nuclear

envelope. Intermediate filament expression may have been a later diversification of eukaryotic cell evolution since they are not expressed in all cells of vertebrates (*i.e.* some tissue culture cell lines) nor have they been found in a number of invertebrates (Bartnik *et al.*, 1986), although this may reflect lack of suitable antibody probes. Another difference between the lamins and intermediate filament proteins is that although both belong to developmentally regulated gene families (see *1.3* and *4.1*) different forms of lamins appear to be expressed in nondividing versus dividing cells (i.e. $L_{III}$ in oocytes, muscle and neurons, see *1.3*) whereas different intermediate filament proteins are expressed in cells depending on their embryonic origin (*i.e.* vimentin in mesenchymal-derived cells, desmin in muscle, *etc.*).

Unlike all other intermediate filaments, the lamin filaments are located and assembled in the cell nucleus, not in the cytoplasm. Lamins A and C contain a putative nuclear localization signal in their carboxy-terminal domain similar to that of the simian virus 40 large T antigen (Fig. 9) which might explain their accumulation in the nucleus. Deletion of and site-directed mutagenesis of this putative signal in a eukaryotic lamin expression system may help resolve this issue (see *4.6*). Since lamins must be synthesized in the cytoplasm it remains to be determined how they avoid being assembled on to preexisting cytoplasmic intermediate filaments or into lamin cytoplasmic filaments. Microinjection experiments with keratin mRNA, monitoring the translation products by immunofluorescence and immunoelectron microscopy, has shown that the expressed keratin is assembled into separate intermediate filament arrays from pre-existing cell type-specific non-keratin networks, yet they will coassemble with pre-existing keratin networks of different cell types and species (Kreis *et al.*, 1983; Franke *et al.*, 1984). The speed of lamin transport into the nucleus coupled with the rate of lamin synthesis may be such that cytoplasmic concentrations of lamins may be too low to permit polymerization into filaments (see *4.3*). Postsynthetic modifications, such as phosphorylation (see *1.4*)

may also prevent cytoplasmic lamin filament formation.

Unlike the cytoplasmic intermediate filament proteins, the lamins undergo complete and reversible depolymerization during mitosis (see *1.4*). This process may involve transient hyperphosphorylation of the lamins (see *1.4*). The large number of *Ser* residues outside the α-helical domain and in the lamin A tail (Fig. 9) may be substrates for such kinase activity.

Finally, the lamin α-helical domains are predicted to exhibit stronger ionic interactions to stabilize the coiled coil than the cytoplasmic intermediate filament rod domains (Parry *et al.*, 1986). Also, lamins can be depolymerized into stable dimers much more readily than typical intermediate filaments (Aebi *et al.*, 1986). Indeed, the smallest, stable subfilamentous oligomer of cytoplasmic intermediate filament proteins that can exist in solution - the so-called "protofilament," "structural unit" or "repeating unit" - is a tetramer, a pair of coiled coils (Ahmadi and Speakman, 1978; Woods and Gruen, 1981; Geisler and Weber, 1982; Gruen and Woods, 1983; Woods, 1983; Crewther *et al.*, 1983; Quinlan *et al.*, 1984; Woods and Inglis, 1984; Ip *et al.*, 1985b). Recently, Quinlan *et al.* (1986) have been able to prepare dimers of various intermediate filament molecules by depolymerization of intermediate filaments and by reconstitution from fully denatured molecules, using buffers containing 3M guanidine-HCl. Thus, rather harsh denaturants are necessary to disassemble intermediate filament protein tetramers into dimers, conditions not necessary to destabilize tetrameric lamins. These differences in physical properties between lamins and intermediate filament proteins will be discussed below (*4.3*).

### *4.3 Lamin Assembly into a Lamina*

The striking homology of lamins A and C with intermediate filament proteins suggests that at least some aspects of their assembly into a lamina may resemble that of intermediate filament proteins into intermediate filaments. However, the

mechanism of intermediate filament assembly *in vitro* and *in vivo* is still poorly understood. The first step in filament assembly is the formation of the coiled coil molecule. Since an α-helix is generally intrinsically unstable in water the formation of a coiled coil structure between two adjacent α-helices serves to stabilize these proteins. Crick (1953) pointed out that the physical basis of the stability of these coiled coils is a combination of "knobs into holes" packing of apolar side chains forming the core of the protein structure and exposure of charged groups at the surface of the double-stranded rope. The protein subunits of intermediate filament coiled coils have been predicted (on the basis of calculated interchain ionic interactions) (Parry *et al.*, 1977) and shown (Woods and Inglis, 1984; Parry *et al.*, 1985; Quinlan *et al.*, 1986) to be arranged in parallel and in exact axial register. Similarly, the electron microscopy showing two globular heads attached to an approximately 50 nm rod-like tail clearly reveals the parallel unstaggered orientation of purified lamin dimers (Aebi *et al.*, 1986; see *4.1*). This is also compatible with the sequence data and studies showing that lamins can be oxidatively crosslinked into oligomers (Lam and Kasper, 1979; Shelton *et al.*, 1982), since the sequence data show that all *Cys* residues of lamins A (5) and C (1) are located in their carboxyl-terminal domains (Fig. 9). Interestingly, the number of predicted interchain ionic interactions is much higher for a parallel unstaggered arrangement of lamin chains than it is for intermediate filament chains (Parry *et al.*, 1986; see *4.2*) suggesting that lamin dimers are more stable molecular structures than intermediate filament protein dimers. Both the longer Coil 1B and Coil 2 of the lamins are more highly charged than Coils 1B and 2 of intermediate filament proteins. This may at least partially explain why at certain concentration lamins exist as dimers while intermediate filament proteins exist as tetramers (see *4.2*): at these concentrations the ionic interactions between the two lamin chains and the surrounding medium stabilize the molecules to a greater degree than the interactions between adjacent subunits of a

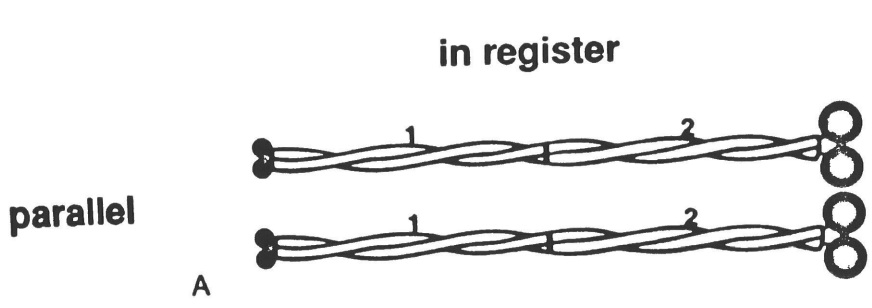
lamin tetramer, while intermediate filament tetramers are much more stable than intermediate filament dimers and hence do not depolymerize into less stable dimeric structures. Head and tail domains probably play a role, ranging from small to large, in stabilizing or destabilizing these structures. Purified rod domains of certain intermediate filament proteins (Geisler *et al.*, 1982) are capable of forming only small abnormal filament aggregates. Naturally occurring intermediate filament proteins with very small end domains (Wu *et al.*, 1982), however, are capable of forming normal intermediate filaments. Also, the thermal stability of isolated rod domains is very similar to that of the parent molecules, indicating that the end domains do not greatly affect this aspect of the stability of the α-helical rod domain (Woods and Inglis, 1984). Interestingly, the dissociation of myosin filaments *in vitro* by high salt buffer gives rise mainly to coiled coil dimers, not tetramers, although tetramers are believed to be intermediates in myosin assembly (Davis *et al.*, 1982). Thus, lamins may be considered more structurally analagous to myosin than intermediate filament proteins in this regard. (Indeed, the appearance of the lamin dimer in electron micrographs as a 50 nm rod with two globular heads [Aebi *et al.*, 1986] can be described as a "minimyosin" structure.)

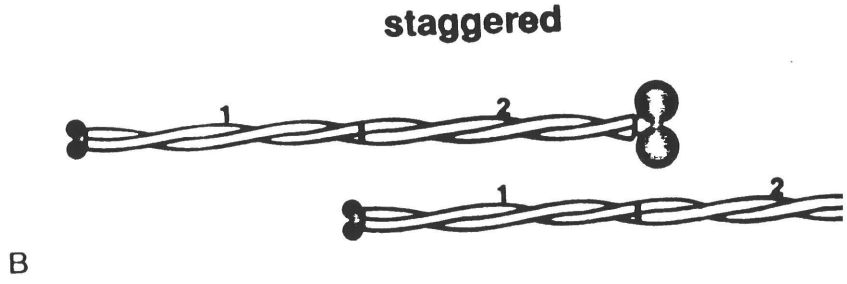
The formation and alignment of four-chain tetramers from two-chain coiled coil molecules are also postulated to be based on the number of possible ionic interactions between coiled coil neighboring segments (Crewther *et al.*, 1983). The maximum number of favorable ionic interactions occur when segments 1B and/or 2 are adjacent. Thus, there are five possible alignments resulting from parallel or antiparallel half-staggered or in-register neighboring dimers (Fig. 16). While X-ray data and proteolytic experiments favor a 60-70 nm particle corresponding to Fig. 16D (Woods and Gruen, 1981; Woods and Inglis, 1984; Parry *et al.*, 1985; Fraser *et al.*, 1986), electron microscopy seems to reveal the shorter 45-48 nm particle corresponding to Fig. 16C (Quinlan *et al.*, 1984; Ip *et al.*, 1985a; Geisler *et al.*, 1982,

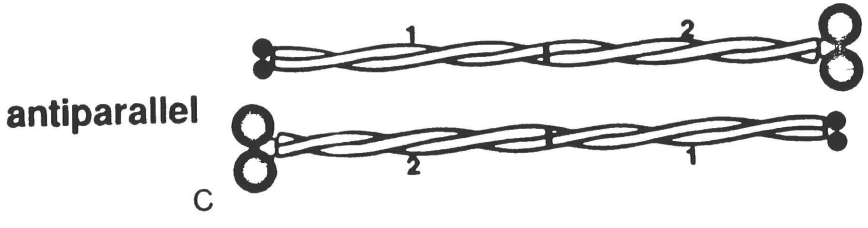
## *Figure 16*

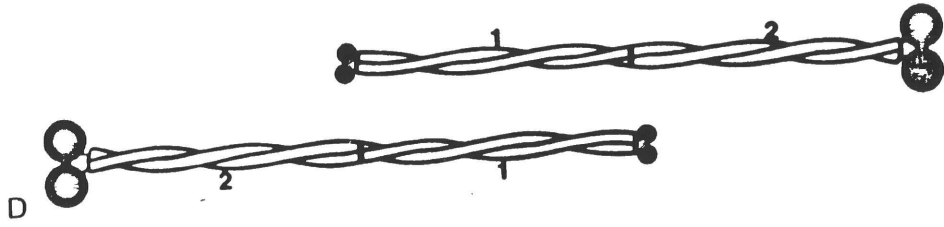
*Possible Orientations of Intermediate Filament Tetramers*

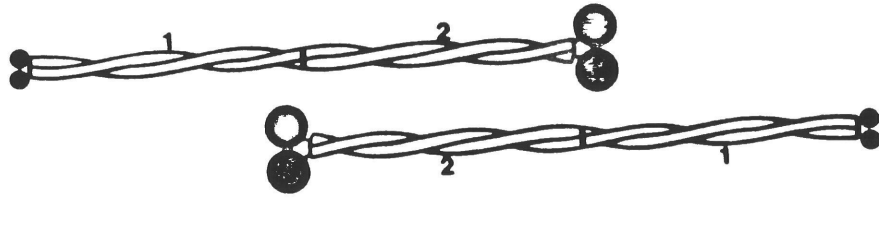
The possible alignments of two-chain coiled coil molecules to form the putative four-chain building block are based on the maximum numbers of favorable ionic interactions that can occur between neighboring coiled coils (Crewther *et al.*, 1983). These occur when Coils 1B and/or 2 of the neighboring dimers are adjacent, either in parallel or antiparallel, as shown. Models A and C would be about 45-48 nm long, and models B, D and E would be between 60-70 nm long. Evidence for both models C and D exist (see *4.3*).


in register
parallel
1
2
1
2
A


staggered
1
2
1
2
B


antiparallel
1
2
2
1
C


1
2
2
1
D


1
2
2
1
E

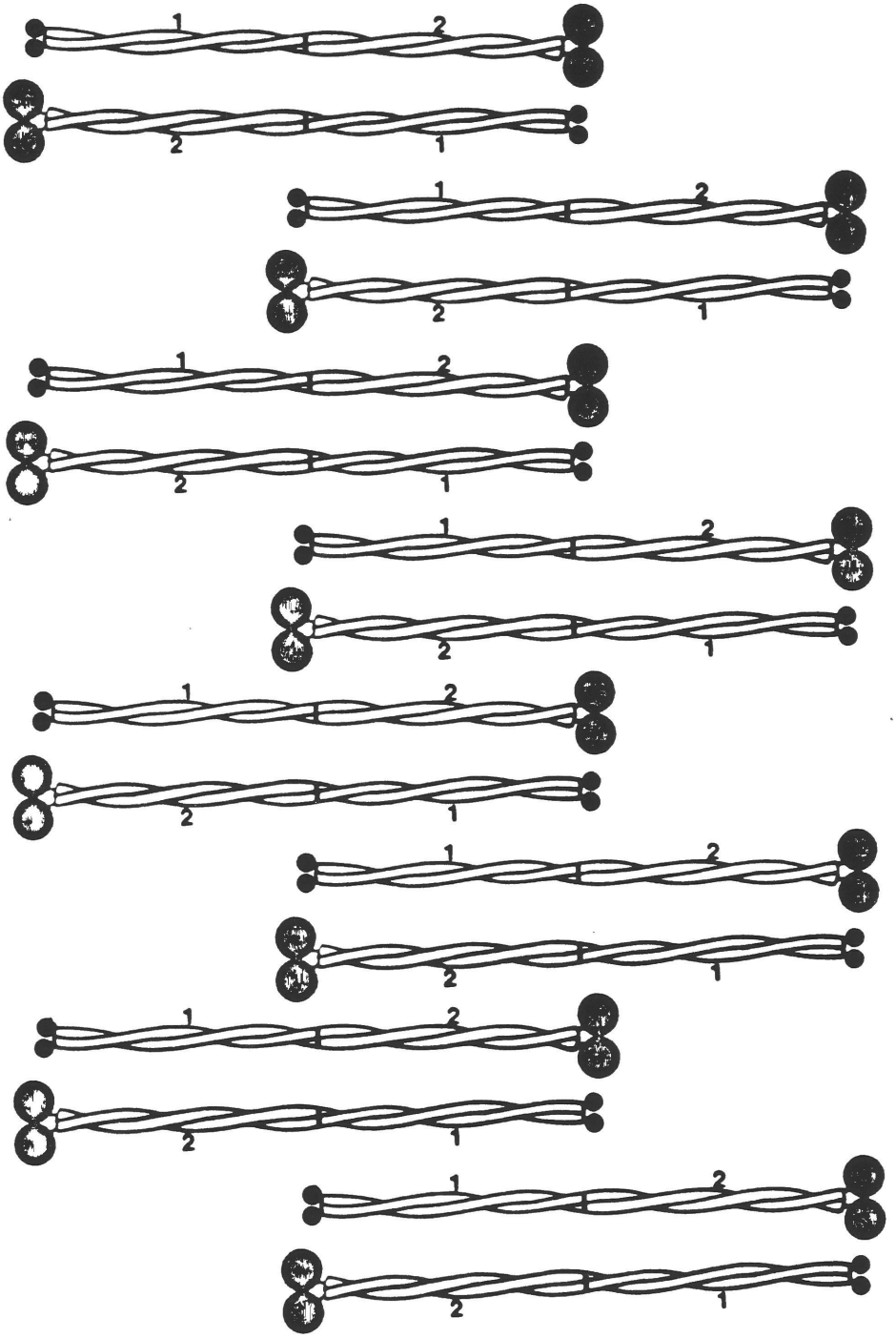
1985; Kaufmann *et al.*, 1985). These apparently conflicting models are actually compatible when one considers that both tetrameric forms may result from depolymerization of intermediate filaments composed of 45-48 nm tetramers arranged in a half-staggered orientation (Fig. 17) or from depolymerization of 62 nm tetramers arranged in register (Fraser *et al.*, 1986). Thus, the half-stagger (which corresponds to the observed axial repeats, see below) may occur at the level of the octamer or the tetramer. Depending on the conditions used for depolymerization, both tetrameric forms may be obtained. In support of this, both forms (45-50 nm and 60-70 nm) recently could be identified by electron microscopy within the same sample (Kaufmann *et al.*, 1985; Potschka, 1986). Lamin tetramers have also been reconstituted *in vitro* in a concentration-dependent fashion (Aebi *et al.*, 1986) although the orientation of the neighboring coiled coils has not been investigated.

Keratin filaments are obligate heteropolymers (Lee and Baden, 1976; Steinert *et al.*, 1976; Franke *et al.*, 1983) consisting of Type I (acidic) and Type II (basic) keratin molecules in a 1:1 stoichiometry (Milstone, 1981; Gruen and Woods, 1983; Quinlan *et al.*, 1984; Woods and Inglis, 1984; Hatzfeld and Franke, 1985). While all other intermediate filament proteins form homopolymers *in vivo* and *in vitro* (Geisler and Weber, 1980, 1981; Huiatt *et al.*, 1980; Renner *et al.*, 1981; Liem *et al.*, 1982; Steinert *et al.*, 1982), they are capable of forming heteropolymers *in vitro* and even *in vivo* (Steinert *et al.*, 1981, 1982; Geisler and Weber, 1981; Liem *et al.*, 1982; Zackroff *et al.*, 1982; Quinlan and Franke, 1982, 1983; Sharp *et al.*, 1983; Wang *et al.*, 1984) . There is evidence that a keratin tetramer consists of a Type I homodimer plus a Type II homodimer (Gruen and Woods, 1983; Hatzfeld and Franke, 1985; Quinlan *et al.*, 1986), but there is also evidence for keratin heterodimers (Woods and Inglis, 1984; Parry *et al.*, 1985). Similarly, it remains to be determined if lamins A and C form heteropolymers or homopolymers *in vivo* and, if heteropolymers, whether they form at the level of dimers, tetramers or some higher level of organization. *In*

## *Figure 17*

*Intermediate Filament Formation from Subunits*

Lateral associations between 8 unstaggered antiparallel tetramers approximately 45 nm in length complete the ring of subunits present per 10 nm intermediate filament diameter. This is an average number based on mass-per-unit-length measurements of filaments determined by scanning transmission electron microscopy (STEM) (Steven *et al.*, 1982; see *4.3*). These tetramers may be arranged in a half-staggered orientation accounting for the 21-25 nm axial repeats observed in metal-shadowed intermediate filament protein and lamin filaments (Henderson *et al.*, 1982; Milam and Erickson, 1982; Aebi *et al.*, 1986). Additionally, this arrangement might result, under appropriate conditions, in the isolation of tetramers of two different orientations: the unstaggered antiparallel tetramer as in Fig. 16C and the staggered antiparallel tetramer of Fig. 16D. The subunits presumably interact end-to-end as well as laterally to elongate the filament. In contrast to this model, which generates essentially an apolar structure, tetramers might be formed in an unstaggered parallel arrangement, and the filament might be elongated by half-staggered and end-to-end associations in parallel. This would generate a polarized filament (*i.e.* amino-termini on one end, carboxy-termini on the other end).


1
2
2
1
1
2
2
1
1
2
2
1
1
2
2
1
1
2
2
1
1
2
2
1
1
2
2
1
1
2
2
1

*vitro* studies with lamins A and C provide a unique opportunity to investigate subdomain organization of these proteins. Lamin C essentially is a truncated form of lamin A, with the addition of only six amino acid residues at its carboxyl terminus. By testing if lamin C can or cannot form homopolymers or heteropolymers with lamin A, one can determine the extent to which the 98 amino acids of the lamin A tail are necessary for filament formation. Also, one can then determine the extent to which the 98 amino acids contribute to the stabilization of various filament intermediates, such as dimers and tetramers. A homopolymer of lamin $L_{III}$ probably comprises the entire lamina of *Xenopus* oocytes (see *1.3* and *4.1*).

Based on the kinetics of *in vitro* filament polymerization, it has been suggested (Steinert *et al.*, 1976; Zackroff and Goldman, 1979) that filament assembly is a two stage "nucleation-condensation" process, similar to actin polymerization, involving the rate limiting formation of short intermediate filament nucleation sites followed by rapid filament elongation from these short intermediate filaments. As formalized by Oosawa and Kasai (1962), this two stage reaction would result from exceeding a critical subunit protein concentration, thereby stabilizing the nucleation sites so they can support elongation. Above this critical concentration, steady state filament polymerization would be a linear function of the total protein concentration. Studies of keratin assembly (Eichner *et al.*, 1985) suggest that the rate limiting nucleation step may involve the formation of stable 4.5 nm-wide "protofibrils" consisting of two tetramers, or possibly the subsequent formation of a short filament composed of two to four of these protofibril octamers (see also Ip *et al.*, 1985a,b concerning vimentin and desmin intermediates). Subsequent filament elongation is believed to be stabilized by different interactions than those that maintain the dimers and tetramers. As mentioned above, intermediate filament protein head and tail domains may play a role in filament assembly and stability since isolated rod domains cannot form normal filaments (Geisler *et al.*, 1982) and the desmin molecule

lacking its amino terminus cannot form filaments (Kaufmann *et al.*, 1985). This role is probably variable, however, due to the wide variability in end-domains and due to the existence of intermediate filaments formed from native proteins virtually devoid of head and/or tail domains (Wu *et al.*, 1982). "Headless" desmin can still co-assemble into filaments with full-length desmin (Kaufmann *et al.*, 1985). In possible agreement with the two stage assembly model suggested for intermediate filament formation, *in vitro* lamin filament formation has also been shown to arise in a concentration-dependent manner (Aebi *et al.*, 1986).

Intermediate filaments clearly have been shown to contain a half-unit-length stagger corresponding to either Coil 1 or Coil 2 of the coiled coils and thus to the 21-23 nm beading pattern of metal-shadowed filaments (Fig. 17) (Henderson *et al.*, 1982; Milam and Erickson, 1982; Aebi *et al.*, 1983). Similarly, *in vitro* reconstituted lamin filaments generate an approximately 24-25 nm beading pattern upon dialysis against low ionic strength buffer, changing from filaments to thicker beaded cables and ultimately to paracrystalline arrays (Aebi *et al.*, 1986). As mentioned in *4.1* and *4.2*, the lamin beading pattern is slightly larger than that of intermediate filaments due to their Coil 1B domain (Fig. 12). Utilizing scanning transmission electron microscopy (STEM) to obtain mass-per-unit-length measurements of unstained intermediate filament specimens (Steven *et al.*, 1982, 1983a,b; Eichner *et al.*, 1985) and X-ray diffraction data (Fraser *et al.*, 1986), it has been determined that each intermediate filament contains on the average seven or eight tetramers (28-32 polypeptides) per diameter. However, there is considerable polymorphism in intermediate filament structure both *in vivo* and *in vitro* resulting in thinner and thicker filaments as well as filled-in and hollow filaments (Eichner *et al.*, 1985; Steven *et al.*, 1985). Most strikingly, this polymorphism seems to occur in increments of eight polypeptides per cross section, or two tetramers, suggesting that an octamer might be the building block of these filaments (Eichner *et al.*, 1985). That

lamins assemble into filaments similarly to the cytoplasmic intermediate filament proteins is thus evident from the hierarchy of lamin assembly intermediates - dimeric rods, tetrameric rods, 10 nm filaments, filaments with a 25 nm axial repeat - which mimics the *in vitro* assembly pattern of the cytoplasmic intermediate filament proteins (Quinlan *et al.*, 1984, 1986; Aebi *et al.*, 1983; Ip *et al.*, 1985a) and possibly their *in vivo* pattern (Soellner *et al.*, 1985).

While the assembly of lamins into filaments can be modeled after intermediate filament protein polymerization, the assembly of lamin filaments into a near-tetragonal lattice poses new questions of structural organization. Intuitively it seems that such a highly ordered lattice would require some type of an organizing center to initiate and/or direct polymerization at the end of mitosis or following fertilization. It does not seem likely that lamin filaments could organize themselves into such an ordered meshwork. A possible source for such external "growth/organizing centers" may be DNA or DNA-associated proteins with which the lamins are already believed to interact (see *1.1* and *1.4*). Lamin A and C dimers present throughout the mitotic cell may bind to one or more of these DNA sites in anaphase or telophase, lamin A and C filaments may then assemble from these dimers, and lamin B dimers associated with vesicles may then join the lamin A and C polymer to form a heteropolymer of all three lamins, thus positioning the vesicles alongside each other and the DNA for fusion back into a double membrane. Since assembly of lamin filaments may be spontaneous at certain physiological conditions, dephosphorylation of mitotic lamins may be required for lamin binding to DNA or associated proteins to obtain an ordered near-tetragonal lattice. Since lamina disassembly appears to require a kinase activity (see *1.4* and *4.4*), lamina reassembly may require a phosphatase activity. Dephosphorylated lamins may be required for assembly into filaments and/or for binding to DNA or DNA-associated proteins. Thus, lamina assembly *in vivo* may mimic *in vitro* lamin filament assembly by being

driven by mass-action, yet *in vivo* lamin assembly may be regulated by the putative lamin kinase and a putative phosphatase. It is tempting to speculate that, in line with the two stage nucleation-condensation reaction proposed for filament assembly (see above), the putative DNA organizing center suggested above may act as a nucleation site for lamin filament assembly. Lamin-DNA binding may then proceed by "zipper"-like lateral associations as in other DNA-ligand interactions (*i.e.* DNA-actinomycin, DNA-protamines, *etc.*) (McGhee and von Hippel, 1974). A lamin lattice would result from such lateral associations with DNA plus end-to-end associations with lamin filament subunits. Termination of lattice growth would be imposed by gaps between heterochromatin "islands" beneath the nuclear envelope. Release from the putative nucleation and lateral DNA binding sites, along with the lamin filament kinase activity, would facilitate lamina disassembly in mitosis. These events may also occur on a focal basis during interphase to effect the restructuring and growth of the lamina. Such lamin-DNA restructuring may be a means by which the three-dimensional structure of the genome reorganizes throughout cellular differentiation (Blobel, 1985). However, it still remains to be seen whether a near-tetragonal lattice is a ubiquitous feature of all nuclear lamina structures. Also, it remains to be seen whether lamin $L_{III}$ of the *Xenopus* oocyte is capable of all the interactions ascribed to mammalian lamins A, B and C. Lamin B presumably anchors the lamina to the inner nuclear membrane, and lamins A and C may interact with higher order chromatin domains such as the growth centers mentioned above to form the primary lamina scaffolding.

### *4.4 Lamina Disassembly into Lamins*

Unlike the cytoplasmic intermediate filaments, the nuclear lamina is disassembled in dramatic fashion during mitotic prophase (see *1.4*). Since enzymatically driven hyperphosphorylation has been implicated in this process (see

*1.4*), the factors that affect depolymerization *in vitro* - substrate concentration, pH and ionic strength - may modulate the physiological process. However, as suggested above (*4.3*), in addition to catalyzing lamin filament disassembly, lamin phosphorylation may be involved in destabilizing lamin interactions with DNA or DNA-associated proteins. Also, if DNA or a DNA receptor acts as a nucleation site for lamin filament assembly (see *4.3*), the dissociation of the lamina from the chromatin may shift the equilibrium between assembly and disassembly toward the latter. Disassembly may then proceed spontaneously in a cooperative fashion; as the local lamin concentration decreases, the equilibrium will be shifted toward disassembly thus propagating the disassembly process itself. While this may occur on a large scale at mitosis, local phosphorylation and dephosphorylation events during interphase may trigger focal restructuring of the lamina by the same mechanism.

Cytoplasmic intermediate filaments are also phosphorylated during mitosis, and a transient disintegration into granular aggregates has been observed (Horwitz *et al.*, 1981; Franke *et al.*, 1982; Lane *et al.*, 1982). However, this is not a general phenomenon and the aggregation has been shown not to be caused chiefly by phosphorylation (Celis *et al.*, 1983). Rather, as suggested above for lamin-DNA interactions, phosphorylation of intermediate filament proteins has been suggested to play a role in modulating the various proposed associations of these filaments with organelles and other cellular structures (Celis, *et al.*, 1983).

It remains to be investigated why lamin $L_{III}$ is disassembled into a 9S moiety in the *Xenopus* oocyte suggestive of a partial oligomeric state, while rat lamins disassemble into 4-5S subunits, perhaps dimers or tetramers.

## *4.5 One Gene. Two Messages*

The lamin A and C cDNA sequences are identical until the lamin A sequence

diverges following a CAC codon corresponding to amino acid residue 566 (Fig. 8). The following six nucleotides in the lamin C sequence, GTGAGT, are identical to the consensus sequence of the 5' end of introns (see *3.2*), and the following two nucleotides in the lamin A sequence, GG, represent a consensus sequence of the 5' end of spliced exons (see *3.2*). In combination with evidence from Southern analysis that there may be only one gene coding for the two lamins (McKeon *et al.*, 1986), these consensus sequences suggest that the two lamins may arise from alternate forms of processing. There are two possibilities of what might occur. In the first possibility, the RNA polymerase may synthesize a single lamin C primary transcript which gets polyadenylated and spliced yielding the lamin C mRNA. For lamin A transcription, the polymerase may somehow continue through the lamin C polyadenylation and termination sites to synthesize a large primary transcript which always gets spliced at the consensus site following codon 566 to produce lamin A mRNA. In the second possibility, a single long primary transcript is always synthesized and then a choice is made for polyadenylation to produce either lamin C RNA or lamin A RNA. There is precedent for this method of RNA processing in the production of membrane and secreted forms of IgM RNA (Alt *et al.*, 1980; Rogers *et al.*, 1980). For the lamins, this choice could also be between polyadenylation producing lamin C RNA, and splicing at the codon 566 consensus site which would then result in lamin A RNA production. With the lamin cDNA probes it should be possible to resolve this issue by obtaining a lamin genomic clone and carrying out transcription studies.

Why two nearly identical lamin proteins - lamins A and C - are utilized in mammalian cells also remains to be determined. The presence of a functional lamina composed of only one or two lamins in non-mammalian cells (see *1.2*) further complicates the issue. However, lamins A and C have clearly been shown to differ in their rate of incorporation into the insoluble lamina structure - half times of 5 and 60

minutes, respectively, - in pulse-chase experiments (Gerace *et al.*, 1984; see *1.4*). This may reflect a DNA or DNA-protein binding capacity for lamin A not present in lamin C (or lamin B) which also promotes its assembly into the lamina meshwork. Conversely, lamin C may have a DNA binding capacity which slows down its assembly into the lamina. Alternatively, lamin A may preferentially bind to lamin B within the lamina. The presence of a highly acidic tail in lamin C that is buried within the lamin A tail domain, (Fig. 8) may provide a structural clue to these functional differences.

### ***4.6 Future Prospects***

With the cDNA clones of lamins A and C it should be possible to obtain insight into a variety of issues concerning the structure and function of the nuclear lamina. The cDNAs can be manipulated into a eukaryotic expression system such as a CHO cell line and lamin expression and behavior throughout the cell cycle can be monitored with antibodies specific to these human cDNA protein products. Deletion mutagenesis and site-directed mutagenesis could be carried out to determine both *in vitro* and *in vivo* what domains and specific sequences are responsible for filament formation and elongation, lamina assembly or disassembly and nuclear translocation (*i.e.* a karyophilic sequence, see Fig. 9), and which domains and sequences are phosphorylation sites and DNA-associating sites. Based on the lamin sequences, peptides can be synthesized to raise antibodies to particular domains. These antibodies could be used to investigate the question of filament orientation (*i.e.* parallel or antiparallel, staggered or unstaggered) as well as to determine domains involved in lamina assembly and disassembly *in vivo* and *in vitro* complementing the mutagenesis approach mentioned above. The lamin cDNA probes enable one to determine the genomic organization of the lamin gene family and to investigate its relationship with the intermediate filament gene family. Finally, the cDNA probes

can be used to monitor lamin RNA expression and processing throughout cellular differentiation and embryonic development.

## References

Aaronson, R.P. and Blobel, G. (1975). *Proc. Natl. Acad. Sci.* USA **72**, 1007-1011.

Aebersold, R., Teplow, D., Hood, L. and Kent, S. (1986). *J. Biol. Chem.* **261**, 4229-4238.

Aebi, U., Cohn, J., Buhle, L. and Gerace, L. (1986). *Nature* **323**, 560-564.

Aebi, U., Fowler, W.E., Rew, P. and Sun, T.-T. (1983). *J. Cell Biol.* **97**, 1131-1143.

Afzelius, B.A. (1955). *Exp. Cell Res.* **8**, 147-158.

Ahmadi, B. and Speakman, P.T. (1978). *FEBS Lett.* **94**, 365-367.

Alt, F.W., Bothwell, A.L.M., Knapp, M., Siden, E., Mather, E., Koshland, M. and Baltimore, D. (1980). *Cell* **20**, 293-301.

Alwine, J.C., Kemp, D.J. and Stark, G.R. (1977). *Proc. Natl. Acad. Sci. USA* **74**, 5350-5354.

Aviv, H. and Leder, P. (1972). *Proc. Natl. Acad. Sci. USA* **69**, 1408-1412.

Bartnik, E., Osborn, M. and Weber, K. (1986). *J. Cell Biol.* **102**, 2033-2041.

Beams, H.W., Tahmisian, T.N., Devine, R. and Anderson, E. (1957). *Exp. Cell Res.* **13**, 200-204.

Benavente, R. and Krohne, G. (1985). *Proc. Natl. Acad. Sci. USA* **82**, 6176-6180.

Benavente, R. and Krohne, G. (1986). *J. Cell Biol.* **103**, 1847-1854.

Benavente, R., Krohne, G. and Franke, W.W. (1985). *Cell* **41**, 177-190.

Benton, W.D. and Davis, R. (1977). *Science* **196**, 180-182.

Berk, A.J. and Sharp, P.A. (1977). *Cell* **12**, 721-732.

Biggin, M., Gibson, T. and Hong, G. (1983). *Proc. Natl. Acad. Sci. USA* **80**, 3963-3965.

Blobel, G. (1985). *Proc. Natl. Acad. Sci. USA* **82**, 8527-8529.

Bonner, W.M. (1975a). *J. Cell Biol.* **64**, 421-430.

Bonner, W.M. (1975b). *J. Cell Biol.* **64**, 431-437.

Burke, B. and Gerace, L. (1986). *Cell* **44**, 639-652.

Burke, B., Tooze, J. and Warren, G. (1983). *EMBO J.* **2**, 361-367.

Cech, T.R. (1983). *Cell* **34**, 713-716.

Celis, J.E., Mose Larsen, P., Fey, S.J. and Celis, A. (1983). *J. Cell Biol.* **97**, 1429-1434.

Chirgwin, J., Przybyla, A., MacDonald, R. and Rutter, W. (1979). *Biochemistry* **18**, 5194-5199.

Chou, P.Y. and Fasman, G.D. (1978). *Adv. Enzymol.* **47**, 45-148.

Cleveland, D., Fischer, S., Kirschner, M. and Laemmli, U. (1977). *J. Biol. Chem.* **252**, 1102-1106.

Coggeshall, R.E. and Fawcett, D.W. (1964). *J. Neurophysiol.* **27**, 229-289.

Crewther, W.G., Dowling, L.M., Steinert, P.M. and Parry, D.A.D. (1983). *Int. J. Biol. Macromol.* **5**, 267-274.

Crick, F.H. (1953). *Acta Cryst.* **6**, 689-697.

Dale, R., McClure, B. and Houchins, J. (1985). *Plasmid* **13**, 31-40.

Dame, J., Williams, J., McCutchan, T., Weber, J., Wirtz, R., Hockmeyer, W., Maloy, W.L., Haynes, J.D., Schneider, I., Roberts, D., Sanders, G., Reddy, E.P., Diggs, C. and Miller, L. (1984). *Science* **225**, 593-599.

Daniels, E.W. and Breyer, E.P. (1967). *J. Protozool.* **14**, 167-175.

Davis, J.S., Buck, J. and Greene, E.P. (1982). *FEBS Lett.* **140**, 293-297.

DeRobertis, E.M., Longthorne, T. and Gurdon, J.B. (1978). *Nature (London)* **272**, 254-256.

Dingwall, C., Sharnick, S.V. and Laskey, R.A. (1982). *Cell* **30**, 449-458.

Dwyer, N. and Blobel, G. (1976). *J. Cell Biol.* **70**, 581-591.

Eichner, R., Rew, P., Engel, A. and Aebi, U. (1985). *Ann. N.Y. Acad. Sci.* **455**, 381-402.

Eilers, M. and Schatz, G. (1986). *Nature (London)* **322**, 228-232.

Fawcett. D.W. (1966). *Am. J. Anat.* **119**, 129-146.

Feldherr, C.M. and Ogburn, J.A. (1983). *J. Cell Biol.* **96**, 1486-1490.

Feldherr, C.M., Kallenbach, E. and Schultz, N. (1984). *J. Cell Biol.* **99**, 2216-2222.

Fisher, D.Z., Chaudhary, N. and Blobel, G. (1986). *Proc. Natl. Acad. Sci. USA* **83**, 6450-6454.

Franke, W., Schmid, E., Schiller, D., Winter, S., Jarasch, E., Moll, R., Denk, H., Jackson, B. and Illmensee, K. (1982). *Cold Spring Harbor Symp. Quant. Biol.* **46**, 431-453.

Franke, W.W. (1974). *Int. Rev. Cytol. Suppl.* **4**, 71-236.

Franke, W.W., Scheer, U., Krohne, G. and Jarasch, E. (1981). *J. Cell Biol.* **91**, 39s-50s.

Franke, W.W., Schiller, D.L., Hatzfeld, M. and Winter, S. (1983). *Proc. Natl. Acad. Sci. USA* **80**, 7113-7117.

Franke, W.W., Schmid, E. and Grund, C. (1982). *Cell* **30**, 103-113.

Franke, W.W., Schmid, E., Mittnacht, S., Grund, C. and Jorcano, J.L. (1984). *Cell* **36**, 813-825.

Fraser, R.D.B., MacRae, T.P., Parry, D.A.D. and Suzuki, E. (1986). *Proc. Natl. Acad. Sci. USA* **83**, 1179-1183.

Fry, D.J. (1976). *In* "Mammalian Cell Membranes" Vol. 2., Jamieson, G.A. and Robinson, D.M. (*eds.*), Butterworth, Woburn, 197-265.

Fuchs, E. and Hanukoglu, I. (1983). *Cell* **34**, 332-334.

Fuchs, J.P., Giloh, H., Kuo, C.H., Saumweber, H. and Sedat, J. (1983). *J. Cell Sci.* **64**, 331-349.

Geisler, N. and Weber, K. (1980). *Eur. J. Biochem.* **111**, 425-433.

Geisler, N. and Weber, K. (1981). *J. Mol. Biol.* **151**, 565-571.

Geisler, N. and Weber, K. (1982). *EMBO J.* **1**, 1649-1656.

Geisler, N., Fischer, S., Vandekerckhove, J., Plessmann, U. and Weber, K. (1984). *EMBO J.* **3**, 2701-2706.

Geisler, N., Kaufmann, E. and Weber, K. (1982). *Cell* **30**, 277-286.

Geisler, N., Kaufmann, E. and Weber, K. (1985). *J. Mol. Biol.* **182**, 173-177.

Gerace, L. and Blobel, G. (1980). *Cell* **19**, 277-287.

Gerace, L. and Blobel, G. (1982). *Cold Spring Harbor Symp. Quant. Biol.* **46**, 967-978.

Gerace, L., Blum, A. and Blobel, G. (1978). *J. Cell Biol.* **79**, 546-566.

Gerace, L., Comeau, C. and Benson, M. (1984). *J. Cell Sci. Suppl.* **1**, 137-160.

Ghadially, F.N., Bhatnager, R. and Fuller, J. (1972). *Arch. Pathol.* **94**, 303-307.

Gray, E.G. and Guillery, R.W. (1963). *Z. Zellforsch. Mikrosk. Anat.* **59**, 738-745.

Gruen, L.C. and Woods, E.F. (1983). *Biochem. J.* **209**, 587-595.

Grunstein, M. and Hogness, D. (1975). *Proc. Natl. Acad. Sci. USA* **72**, 3961-3965.

Guttes, S., Guttes, E. and Ellis, R.A. (1968). *J. Ultrastruct. Res.* **22**, 508-529.

Hancock, R. and Hughes, M. (1982). *Biol. Cell* **44**, 201-212.

Hanukoglu, I. and Fuchs, E. (1982). *Cell* **31**, 243-252.

Hatzfeld, M. and Franke, W.W. (1985). *J. Cell Biol.* **101**, 1826-1841.

Henderson, D., Geisler, N. and Weber, K. (1982). *J. Mol. Biol.* **155**, 173-176.

Horwitz, B., Kupfer, B., Eshar, Z. and Geiger, B. (1981). *Exp. Cell Res.* **134**, 281-290.

Hudson, P., Haley, J., Cronk, M., Shine, J. and Niall, H. (1981). *Nature (London)* **291**, 127-131.

Huiatt, T.W., Robson, R.M., Arakawa, N. and Stromer, M.H. (1980). *J. Biol. Chem.* **255**, 6981-6989.

Ip, W., Hartzer, M.K., Pang, Y.S. and Robson, R.M. (1985a). *J. Mol. Biol.* **183**, 365-375.

Ip, W., Heuser, J.E., Pang, Y.S., Hartzer, M.K. and Robson, R.M. (1985b). *Ann N.Y. Acad. Sci.* **455**, 185-199.

Johnson, L., Idler, W.W., Zhou, X., Roop, D.R. and Steinert, P.M. (1985). *Proc. Natl. Acad. Sci. USA* **82**, 1896-1900.

Kalderon, D., Roberts, B., Richardson, W. and Smith A. (1984). *Cell* **39**, 499-509.

Kalifat, S.R., Bouteille, M. and Delarue, J. (1967). *J. Microsc. (Paris)* **6**, 1019-1026.

Kaufman, S., Gibson, W. and Shaper, J. (1983). *J. Biol. Chem.* **258**, 2710-2719.

Kaufmann, E., Weber, K. and Geisler, N. (1985). *J. Mol. Biol.* **185**, 733-742.

Kozak, M. (1984). *Nature (London)* **308**, 241-246.

Kreis, T.E., Geiger, B., Schmid, E., Jorcano, J.L. and Franke, W.W. (1983). *Cell* **32**, 1125-1137.

Krohne, G., Dabauvalle, M. and Franke, W.W. (1981). *J. Mol. Biol.* **151**, 121-141.

Krohne, G., Debus, E., Osborn, M., Weber, K. and Franke, W. (1984). *Exp. Cell Res.* **150**, 47-59.

Krohne, G., Franke, W., Ely, S., D'Arcy, A. and Jost, E. (1978). *Cytobiologie* **18**, 22-38.

Kubai, D.F. and Ris, H. (1969). *J. Cell Biol.* **40**, 508-528.

Laemmli, U.K. (1970). *Nature (London)* **227**, 680-685.

Laliberte, J., Dagenais, A., Filion, M., Bibor-Hardy, V., Simard, R. and Royal, A. (1984). *J. Cell Biol.* **96**, 980-985.

Lam, K.S. and Kasper, C.B. (1979). *J. Biol. Chem.* **254**, 11713-11720.

Lane, E.B., Goodman, S.I. and Trejdosiewicz, I.K. (1982). *EMBO J.* **1**, 1365-1372.

Lebkowski, J. and Laemmli, U. (1982). *J. Mol. Biol.* **156**, 325-344.

Lee, L.D. and Baden, H.P. (1976). *Nature (London)* **264**, 377-378.

Lehner, C., Furstenberger, G., Eppenberger, H. and Nigg, E. (1986). *Proc. Natl. Acad. Sci. USA* **83**, 2096-2099.

Liem, R.K.H., Keith, C.H., Leterrier, J.F., Trenkner, E. and Shelanski, M.L. (1982). *Cold Spring Harbor Symp. Quant. Biol.* **46**, 341-350.

Maike-Lye, R., Newport, J. and Kirschner, M.W. (1983). *J. Cell Biol.* **97**, 81-91.

Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982). "Molecular Cloning: A Laboratory Manual" Cold Spring Harbor Lab., N.Y., 68.

Maul, G.G. (1977). *Int. Rev. Cytol. Suppl.* **6**, 76-186.

Maul, G.G., Baglia, F.A., Newmeyer, D.D. and Ohlsson-Wilhelm, B.H. (1984). *J. Cell Sci.* **67**, 69-85.

Maxam, A. and Gilbert, W. (1980). *Meth. Enzymol.* **65**, 499-560.

Mazanec, K. (1967). *J. Microsc. (Paris)* **6**, 1027-1032.

McGhee, J.D. and von Hippel, P.H. (1974). *J. Mol. Biol.* **86**, 469-489.

McKeon, F., Kirschner, M. and Caput, D. (1986). *Nature (London)* **319**, 463-468.

McKeon, F.D., Tuffanelli, D.L., Fukuyama, K. and Kirschner, M.W. (1983). *Proc. Natl. Acad. Sci. USA* **80**, 4374-4378.

Messing, J. (1983). *Meth. Enzymol.* **101**, 20-78.

Miake-Lye, R. and Kirschner, M.W. (1985). *Cell* **41**, 165-175.

Milam, L. and Erickson, H.P. (1982). *J. Cell Biol.* **94**, 592-596.

Milstone, L. (1981). *J. Cell Biol.* **88**, 317-322.

Newport, J. (1987). *Cell* **48**, 205-217.

Newport, J. and Spann, T. (1987). *Cell* **48**, 219-230.

Okayama, H. and Berg, P. (1983). *Mol. Cell. Biol.* **3**, 280-289.

Oosawa, F. and Kasai, M. (1962). *J. Mol. Biol.* **4**, 10-21.

Oryschak, A.F., Mitchell, D.M. and Ghadially, F.N. (1974). *J. Anat.* **118**, 511-515.

Oryschak, A.F., Mitchell, D.M. and Ghadially, F.N. (1976). *Arch. Path. Lab. Med.* **100**, 218-220.

Ottaviano, Y. and Gerace, L. (1985). *J. Biol. Chem.* **260**, 624-632.

Paine, P.L., Moore, L.C. and Horowitz, S.B. (1975). *Nature (London)* **254**, 109-114.

Pappas, G.D. (1956). *J. Biophys. Biochem. Cytol. Suppl.* **2**, 431-434.

Parry, D.A.D., Conway, J.F. and Steinert, P.M. (1986). *Biochem. J.* **238**, 305-308.

Parry, D.A.D., Crewther, W.G., Fraser, R.D.B. and MacRae, T.P. (1977). *J. Mol. Biol.* **113**, 449-454.

Parry, D.A.D., Steven, A.C. and Steinert, P.M. (1985). *Biochem. Biophys. Res. Commun.* **127**, 1012-1018.

Patrizi, G. and Poger, M. (1967). *J. Ultrastruc. Res.* **17**, 127-136.

Potschka, M. (1986). *Biophys. J.* **49**, 129-130.

Quax, W., Egberts, W., Hendriks, W., Quax-Jeuken, Y. and Bloemendal, H. (1983). *Cell* **35**, 215-223.

Quinlan, R.A. and Franke, W.W. (1982). *Proc. Natl. Acad. Sci. USA* **79**, 3452-3456.

Quinlan, R.A. and Franke, W.W. (1983). *Eur. J. Biochem.* **132**, 477-484.

Quinlan, R.A., Cohlberg, J.A., Schiller, D.L., Hatzfeld, M. and Franke, W.W. (1984). *J. Mol. Biol.* **178**, 365-388.

Quinlan, R.A., Hatzfeld, M., Franke, W.W., Lustig, A., Schulthess, T. and Engel, J. (1986). *J. Mol. Biol.* **192**, 337-349.

Renner, W., Franke, W.W., Schmid, E., Geisler, N., Weber, K. and Mandelkow, E. (1981). *J. Mol. Biol.* **149**, 285-306.

Rigby, P.W.J., Dieckmann, M., Rhodes, C. and Berg, P. (1977). *J. Mol. Biol.* **113**, 237-251.

Risau, W., Saumweber, H. and Symmons, P. (1981). *Exp. Cell Res.* **133**, 47-54.

Rogers, J., Early, P., Carter, C., Calame, K., Bond, M., Hood, L. and Wall, R. (1980). *Cell* **20**, 303-312.

Ruther, U. (1982). *Nucl. Acids. Res.* **10**, 5765-5772.

Ryser, U. (1970). *Z. Zellforsch. Mikrosk. Anat.* **110**, 108-130.

Sanger, F., Nicklen, S. and Coulson, A.R. (1977). *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467.

Schatten, G., Maul, G.G., Schatten, H., Chaly, N., Simerly, C., Balczon, R. and Brown, D.L. (1985). *Proc. Natl. Acad. Sci. USA* **82**, 4727-4731.

Scheer, U., Kartenbeck, J., Trendelenburg, M., Stadler, J. and Franke, W.W. (1976). *J. Cell Biol.* **69**, 1-18.

Schwarzbauer, J.E., Tamkun, J., Lemischka, I. and Hynes, R. (1983). *Cell* **35**, 421-431.

Sharp, G., Osborn, M. and Weber, K. (1982). *Exp. Cell Res.* **141**, 385-395.

Shelton, K.R. and Egle, P.M. (1979). *Biochem. Biophys. Res. Commun.* **90**, 425-430.

Shelton, K.R., Guthrie, V.H. and Cochran, D.L. (1982). *J. Biol. Chem.* **257**, 4328-4332.

Shelton, K.R., Higgins, L., Cochran, D.L., Ruffolo, D. and Egle, P. (1980). *J. Biol. Chem.* **255**, 10978-10983.

Smith, D.E. and Fisher, P.A. (1984). *J. Cell Biol.* **99**, 20-28.

Soellner, P., Quinlan, R.A. and Franke, W.W. (1985). *Proc. Natl. Acad. Sci. USA* **82**, 7929-7933.

Steinert, P.M., Idler, W.W. and Zimmerman, S.B. (1976). *J. Mol. Biol.* **108**, 547-567.

Steinert, P.M., Idler, W.W., Aynardi-Whitman, M., Zackroff, R. and Goldman, R.D. (1982). *Cold Spring Harbor Symp. Quant. Biol.* **46**, 465-474.

Steinert, P.M., Idler, W.W., Cabral, F., Gottesman, M.M. and Goldman, R.D. (1981). *Proc. Natl. Acad. Sci. USA* **78**, 3692-3696.

Steinert, P.M., Parry, D.A.D., Racoosin, E., Idler, W.W., Steven, A.C., Trus, B.L. and Roop, D.R. (1984). *Proc. Natl. Acad. Sci. USA* **81**, 5709-5713.

Steinert, P.M., Rice, R., Roop, D.R., Trus, B.L. and Steven, A.C. (1983). *Nature (London)* **302**, 794-800.

Steinert, P.M., Steven, A.C. and Roop, D.R. (1985). *Cell* **42**, 411-419.

Steven, A.C., Hainfeld, J.F., Trus, B.L., Wall, J.S. and Steinert, P.M. (1983a). *J. Biol. Chem.* **258**, 8323-8329.

Steven, A.C., Hainfeld, J.F., Trus, B.L., Wall, J.S. and Steinert, P.M. (1983b). *J. Cell Biol.* **97**, 1939-1944.

Steven, A.C., Trus, B.L., Hainfeld, J.F., Wall, J.S. and Steinert, P.M. (1985). *Ann. N.Y. Acad. Sci.* **455**, 371-380.

Steven, A.C., Wall, J.S., Hainfeld, J.F. and Steinert, P.M. (1982). *Proc. Natl. Acad. Sci. USA* **79**, 3101-3105.

Stick, R. and Hausen, P. (1980). *Chromosoma* **80**, 219-236.

Stick, R. and Hausen, P. (1985). *Cell* **41**, 191-200.

Stick, R. and Krohne, G. (1982). *Exp. Cell Res.* **138**, 319-330.

Stick, R. and Schwarz, H. (1982). *Cell Diff.* **11**, 235-243.

Stick, R. and Schwarz, H. (1983). *Cell* **33**, 949-958.

Suprynowicz, F.A. and Gerace, L. (1986). *J. Cell Biol.* **103**, 2073-2081.

Thomas, P.S. (1980). *Proc. Natl. Acad. Sci. USA* **77**, 5201-5205.

Towbin, H., Staehelin, T. and Gordon, J. (1979). *Proc. Natl. Acad. Sci. USA* **76**, 4350-4354.

Wang, E., Cairncross, J.G. and Liem, R.K.H. (1984). *Proc. Natl. Acad. Sci. USA* **81**, 2102-2106.

Watson, M.L. (1955). *J. Biophys. Biochem. Cyt.* **1**, 257-270.

Watson, M.L. (1959). *J. Biophys. Biochem. Cyt.* **6**, 147-156.

Weber, K. (1986). *Nature* **320**, 402.

Woods, E.F. (1983). *Biochem. Int.* **7**, 769-774.

Woods, E.F. and Gruen, L.C. (1981). *Aust. J. Biol. Sci.* **34**, 515-526.

Woods, E.F. and Inglis, A.S. (1984). *Int. J. Biol. Macromol.* **6**, 277-283.

Wu, M. and Gerhart, J. (1980). *Dev. Biol.* **79**, 465-477.

Wu, Y.J., Parker, L.M., Binder, N.E., Beckett, M.A., Sinard, J.H., Griffiths, C.T. and Rheinwald, J.G. (1982). *Cell* **31**, 693-703.

Yanisch-Perron, C., Vieira, J. and Messing, J. (1985). *Gene* **33**, 103-119.

Young, R.A. and Davis, R.W. (1983). *Science* **222**, 778-782.

Zackroff, R.V. and Goldman, R.D. (1979). *Proc. Natl. Acad. Sci. USA* **76**, 6226-6230.

Zackroff, R.V., Idler, W.W., Steinert, P.M. and Goldman, R.D. (1982). *Proc. Natl. Acad. Sci. USA* **79**, 754-757.

End