The Lamin B Receptor of the Nuclear Envelope Inner Membrane: A Polytopic Protein with Eight Potential Transmembrane Domains

Howard J. Worman, Carole D. Evans, and Günter Blobel

Laboratory of Cell Biology, Howard Hughes Medical Institute, The Rockefeller University, New York 10021

Abstract. The lamin B receptor is a previously identified integral membrane protein in the nuclear envelope of turkey erythrocytes that associates with the nuclear intermediate filament protein lamin B (Worman, H. J., J. Yuan, G. Blobel, and S. D. Georgatos. 1988. Proc. Natl. Acad. Sci. USA. 85:8531-8534). In the present report, we use cell fractionation and antibodies against the lamin B receptor to localize it to an 8-M urea-extracted membrane fraction of chicken liver nuclei, supporting an inner nuclear membrane localization. We deduced the amino acid sequence of the chicken lamin B receptor from overlapping clones obtained by screening cDNA libraries with a probe generated by the polymerase chain reaction with primers based on the partial protein sequence of the isolated protein. The mature lamin B receptor has a calculated molecular mass of

THREE morphologically distinct membrane domains can be discerned in the nuclear envelope of eukaryotic cells. The first of these domains is the outer nuclear membrane, which contains ribosomes on its cytoplasmic side and resembles the rough endoplasmic reticulum (RER)¹ with which it is continuous at certain points (15, 18, 35). The second is the inner nuclear membrane, which is associated on its nucleoplasmic side with the nuclear lamina (12, 15, 18), a meshwork of intermediate filament proteins termed lamins (1, 14, 25, 32). The third distinct membrane is the nuclear pore membrane, which forms numerous annular transcisternal connections between the outer and inner nuclear membranes and is associated with the nuclear pore complexes (15, 18).

Because viral integral membrane proteins synthesized in the RER have access to the inner nuclear membrane presumably by lateral diffusion via the nuclear pore membrane (6,

73,375 D and eight segments of hydrophobic amino acids that could function as transmembrane domains as determined by hydropathy analysis. Preceding the first putative transmembrane segment is a highly charged 204-residue-long amino terminal region that contains two consensus sites for phosphorylation by protein kinase A. Since the lamin B receptor has been shown to be phosphorylated by protein kinase A in vitro and in vivo and this phosphorylation affects lamin B binding (Applebaum, J., G. Blobel, and S. D. Georgatos. 1990. J. Biol. Chem. 265:4181-4185), it is likely that this amino terminal region faces the nucleoplasm. The amino terminal region also contains three DNA-binding motifs that are found in gene regulatory proteins and histones, suggesting that the lamin B receptor may additionally play a role in gene regulation and/or chromatin organization.

41, 42), the distribution of integral membrane proteins between the RER and the three membrane domains of the nuclear envelope could be random. However, the unique association of specific structures (ribosomes, pore complexes, and lamina) with specific membrane domains of the nuclear envelope suggests a nonrandom distribution of at least those integral membrane proteins that may be involved in the anchorage of these structures. In support of this notion, several integral membrane proteins associated with these structures have been identified and localized to specific nuclear envelope membrane domains. An integral membrane glycoprotein (gp210) has been localized to the nuclear pore membrane where it presumably functions in the anchorage of the pore complexes (19, 44). Another integral membrane protein with an apparent molecular mass of 58 kD (p58), likely located in the inner nuclear membrane, has been identified as a lamin B receptor (43), which presumably functions in the anchorage of the lamina to this membrane. Three other proteins (p75, p68, and p55) that are recognized by a single monoclonal antibody have also been characterized as integral proteins of the inner nuclear membrane and proposed to be associated either directly or indirectly with the lamina (38). These three proteins and the lamin B receptor (whether or not they are related remains to be investigated) are all markers for the inner nuclear membrane, much as gp210 is

Dr. H. Worman's address after 1 October 1990 will be the Department of Medicine, Mount Sinai School of Medicine, New York 10029. Dr. C. Evans' permanent address is the Department of Nuclear Medicine, Veterans Administration Medical Center, New York University Dental School, New York 10010.

^{1.} Abbreviations used in this paper: PCR, polymerase chain reaction; RER, rough ER.

a marker for the nuclear pore membrane (19), and RER proteins are markers for the outer nuclear membrane (2, 35). Whether the outer nuclear membrane domain also possesses marker proteins that distinguish it from the RER proper is presently not known.

Except for gp210 of the nuclear pore membrane (44), the primary structures for integral inner nuclear membrane marker proteins and other pore complex membrane marker proteins have not been determined. Here we report the complete primary structure of the lamin B receptor, an inner nuclear membrane protein, deduced from the sequences of cDNA clones. We further discuss structural features of this protein with regards to its topology and its multiple potential functions.

Materials and Methods

Homogenization and Cellular Fractionation of Chicken Liver

Liver was removed from 3-d-old chicks and homogenized at 4° C in 250 mM sucrose, 50 mM triethanolamine (pH 7.5), 25 mM KCl, 5 mM MgCl₂, 1 mM DTT, and 0.2 mM PMSF. The homogenate was passed through two layers of cheesecloth and then centrifuged at 1,200 g for 10 min, yielding a pellet enriched in nuclei and a postnuclear supernatant. The latter was centrifuged at 10,000 g for 10 min yielding a pellet enriched in mitochondria and a postmitochondrial supernatant. The postmitochondrial supernatant was centrifuged at 10,000 g 60 min giving a pellet enriched in microsomes and a postmicrosomal supernatant.

Electrophoresis and Immunoblotting

Electrophoresis on SDS-polyacrylamide slab gels was performed under reducing conditions as described by Laemmli (27). Immunoblotting was performed as previously described (43). For immunoblots with guinea pig antiserum that was raised against gel-purified turkey erythrocyte lamin B receptor (43), the serum was preabsorbed against human skin keratins (Sigma Chemical Co., St. Louis, MO) immobilized on nitrocellulose strips as described (20) to remove contaminating antikeratin antibodies. Guinea pig antiserum raised against gel-purified turkey lamin B (43) and rabbit antiserum raised against purified pig UDP-glucuronosyltransferase (antiserum supplied by Dr. A. Dannenberg, Cornell University Medical College, New York) were used without treatment or purification.

Isolation of Lamin B Receptor from Turkey Erythrocyte Nuclear Envelopes

Isolation of nuclei from turkey erythrocytes, preparation of nuclear envelopes, and extraction of nuclear envelopes with 1 M NaCl and 8 M urea to yield "urea-extracted nuclear envelopes" were performed as previously described (17, 43). Proteins of urea-extracted nuclear envelopes were separated by electrophoresis under reducing conditions on SDS-polyacrylamide slab gels and the Coomassie blue-stained band corresponding to the lamin B receptor (43) was isolated by electroelution as described by Hunkapiller et al. (23).

Proteolytic Cleavage of Lamin B Receptor Protein

V8 proteolytic cleavage of the lamin B receptor and separation of proteolytic fragments was performed as described by Cleveland et al. (10). Briefly, lamin B receptor protein isolated by electroelution as described above was resuspended at a concentration of 0.5 mg/ml in 0.5% SDS, 125 mM Tris-HCl (pH 6.8), and incubated with 75 μ g/ml of V8 protease (Boehringer Mannheim Diagnostics, Inc., Houston, TX) at 37°C for 45 min. The V8 proteolytic cleavage peptides were separated by electrophoresis under reducing conditions on SDS-polyacrylamide gels (15% acrylamide).

Amino Acid Sequencing

Amino acid sequencing of the lamin B receptor was performed on both electroeluted lamin B receptor and on the protein migrating at 58 kD in ureaextracted turkey erythrocyte nuclear envelope fractions transferred to polyvinylidene difluoride membrane (Immobilon Transfer 0.45- μ m pore size; Millipore Continental Water Systems, Bedford, MA). Electrophoretic transfer to polyvinylidene difluoride membrane, protein visualization and preparation for sequencing were performed according to Matsudaira (31). V8 cleavage peptides separated by electrophoresis on SDS-polyacrylamide slab gels were also electrophoretically transferred to polyvinylidene difluoride membrane for amino terminal sequencing. These samples were sequenced on a model 470A gas-phase sequenator (Applied Biosystems Inc., Foster City, CA) as described (22).

RNA Isolation and cDNA Preparation

RNA was isolated from livers of 2-d-old chicks by the method of Chirgwin et al. (9). Poly (A^+)-enriched RNA was isolated by one cycle of oligothymidylic acid-cellulose (Boehringer Mannheim Diagnostics, Inc.) chromatography according to the method of Aviv and Leder (4). Poly (A^+)-enriched RNA was reverse transcribed to cDNA using the Invitrogen (San Diego, CA) Reverse Transcription kit following the manufacturer's instructions.

Polymerase Chain Reaction (PCR)

The polymerase chain reaction was performed as previously described (36) using the Gene Amp Taq-polymerase kit (Perkin Elmer Corp./Cetus Corp., Emeryville, CA). The reaction was run for 25 cycles in a DNA thermal cycler (Perkin Elmer Corp./Cetus Corp.) with the following incubation conditions: denaturation at 94°C for 1 min, annealing at 45°C for 2 min, and extension at 72°C for 3 min. Primers (synthesized on a model 380B oligonucleotide synthesizer [Applied Biosystems Inc.] and purified by electrophoresis on 20% polyacrylamide gels), template (chicken liver cDNA), deoxynucleotides, and Taq polymerase were used at the concentrations recommended in the Gene Amp Kit instructions. The amplified reaction mixture was subjected to electrophoresis on agarose gels, and the DNA migrating at the expected size was eluted, phosphorylated, and subcloned into the Sma I site of PUC-mp19 by standard methods (30) for double-stranded sequencing.

cDNA Library Screening

A synthetic 25-mer oligonucleotide corresponding to a portion of the amplified PCR product was end labeled with [³²P]gamma-ATP (New England Nuclear, Boston, MA) using T4 polynucleotide kinase (New England Biolabs, Beverly, MA) as described (30). The ³²P-labeled probe was purified by electrophoresis on a 20% polyacrylamide gel and used to screen a previously characterized (28) chick embryo cDNA library (supplied by Dr. D. Johnson, University of California, San Francisco) constructed in lambda Zap (Stratagene, La Jolla, CA). Library screening, plaque purification, and DNA isolation by CsCl-gradient centrifugation were by standard methods (30). The cDNA insert of the largest positive clone from this library (termed DJ-5) was ³²P-labeled using the Boehringer Mannheim Random Primed DNA Labeling kit and used to isolate cross-hybridizing overlapping clones from a previously described (33) chick embryonal erythrocyte lambda gt11 cDNA library (supplied by Dr. R. Moon, University of Washington, Seattle, WA).

RNA Analysis (Northern Blotting)

Northern blotting of chick liver $poly(A^+)$ -enriched RNA, isolated as described above, was performed by the method of Thomas (40).

DNA Sequencing

DNA sequencing was performed by the dideoxy chain termination method (37) using the Sequenase kit (United States Biochemical Corp., Cleveland, OH) on either CsCl-gradient prified double-stranded DNA or single-stranded DNA subcloned into M13. Synthetic oligonucleotides were used as primers in the sequencing reactions.

Sequence Analysis

DNA sequence analysis was performed using the DNA Star software package (DNA Star, Inc., Madison, WI) on a Personal Computer AT (IBM). Hydropathy analysis was performed by the method of Kyte and Doolittle (26) using the program in the DNA Star package.

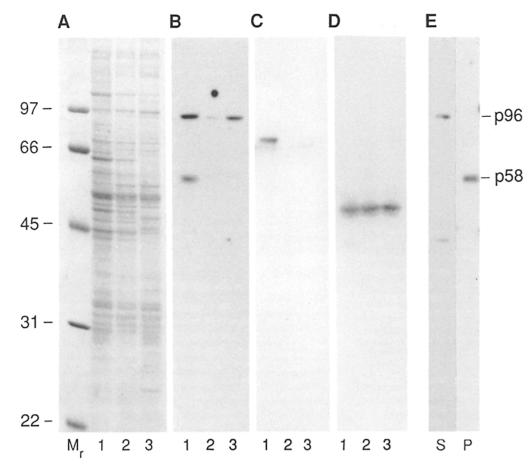


Figure 1. Subcellular fractionation of the lamin B receptor in chick liver cells. (A) Coomassie blue-stained SDS-polyacrylamide slab gel (10% acrylamide) of crude nuclear fraction (lane 1), crude mitochondrial fraction (lane 2), and microsomal fraction (lane 3) obtained from homogenates of chicken liver. Migration of molecular mass standards are shown at left. (B-D) Autoradiograms of immunoblots of fractions identical to those shown in A probed with guinea pig serum containing antibodies against the lamin B receptor (B); guinea pig serum containing antibodies against turkey lamin B (C); and rabbit serum containing antibodies against UDP-glucuronosyltransferase (D). (E) Autoradiogram of immunoblot of proteins of chick liver nuclei that were in the supernatant fraction (S) or remained in the pellet (P) after extraction with 8 M urea. Nuclei were suspended in 100× volume excess of 8 M urea, 10 mM Tris-HCl (pH 8.0), 2 mM EDTA, 1 mM DTT, and 0.2 mM PMSF by bath sonication, incubated at room temperature for 40 min and centrifuged at 365,000 g 20 min to yield supernatant (S) and pellet (P) fractions. Bands corresponding to the lamin B receptor (p58) and cross-reactive 96 kD proteins (p96) are indicated at the right. Immunoblotting was performed using 1:500 dilutions of sera. Sera were used without purification; however, the guinea pig anti-lamin B receptor serum used in B and E was first preabsorbed against human skin keratins to remove contaminating antibodies.

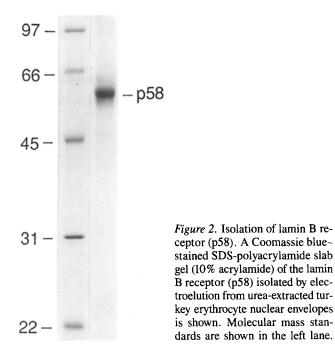
Materials

Unless otherwise specified, routine chemical reagents were obtained from Sigma Chemical Co. or Fisher Scientific Co. (Pittsburgh, PA). Enzymes other than those in the kits described above were obtained from either Boehringer Mannheim Diagnostics, Inc. or New England Biolabs. Unless otherwise indicated, bacteria, cloning vectors, and plasmids were obtained from either Stratagene, Bethesda Research Laboratories (Gaithersburg, MD), Pharmacia Fine Chemicals (Piscataway, NJ), or Clonetech Laboratories, Inc. (La Jolla, CA).

Results

Localization of the Lamin B Receptor to the Inner Nuclear Membrane

The lamin B receptor (p58) has been previously identified in turkey erythrocytes using urea-extracted nuclear envelopes and ¹²⁵I-labeled lamin B in both solution binding and solid-phase ligand blotting assays (43). Antibodies raised against p58 yielded a smooth nuclear rim staining on immunofluorescence microscopy of avian erythrocytes identical to that obtained with antibodies against lamin B, consistent with a localization of p58 in the inner nuclear membrane (43). Because avian erythrocytes contain few other intracellular membranes, it could not be ruled out that, in cells with a more complex network of intracellular membranes, p58 might also be located elsewhere than in the inner nuclear membrane. To address this question, we fractionated a chicken liver homogenate by differential centrifugation (Fig. 1) into nuclear (lanes 1), mitochondrial (lanes 2), and microsomal (lanes 3) fractions; resolved their proteins by SDS-PAGE (Fig. 1 A); and probed corresponding protein blots with antibodies against p58 (Fig. 1 B), lamin B (Fig. 1 C), and UDP-glucuronosyltransferase (Fig. 1 D), a marker for the ER and outer nuclear membrane (2, 16). The lamin B receptor (p58) fractionated exclusively into the nuclear fraction (Fig. 1 B, lane 1) and not into the mitochondrial or



microsomal fractions (Fig. 1 *B*, lanes 2 and 3). Identical fractionation was seen for lamin B (Fig. 1 *C*, lanes 1-3). These data strongly suggest that p58 is located in the inner rather than the outer nuclear membrane. If p58 were located in the outer nuclear membrane, it would be expected to be present also in the RER and therefore would be found not only in the nuclear fraction but also in the mitochondrial and microsomal fractions. All of these fractions contain the ER/ outer nuclear membrane marker enzyme UDP-glucuronosyl-transferase (Fig. 1 *D*, lanes 1-3) and have also been shown previously to contain several other proteins of the ER and outer nuclear membrane (2).

It should be noted that the anti-p58 antibodies are apparently not monospecific. Although they reacted only with p58 when used to probe subcellular fractions of turkey erythrocytes (43), they also cross-reacted with a protein with an apparent molecular mass of 96 kD (p96) when chicken liver fractions were probed (Fig. 1 B). However, p96 is not an integral membrane protein since it can be extracted by 8 M urea (Fig. 1 E, lane S), whereas p58 remains associated with Figure 4. Northern blot of $poly(A^+)$ -enriched RNA from chicken liver probed with the ³²Plabeled cDNA insert of clone DJ-5. 2 μ g of poly(A⁺) RNA were loaded on the gel. Hybridization was performed at 42°C for 24 h in 6 × SSC, 2 × Denhardt's solution, 50% formamide, 0.1% SDS, 20 mM NaPhosphate (pH 7.4), and 0.1 mg/ml of denatured salmon sperm DNA. Nitrocellulose sheets were then washed three times (10 min each time) at room temperature in 2 × SSC and then at 50°C with 1 × SSC for 30 min each wash. Migration of 28S and 18S rRNA in parallel lane of total RNA is indicated.

membranes after 8 M urea extraction (Fig. 1 E, lane P). The nature of this cross-reactive 96 kD protein remains to be investigated.

cDNA Cloning of the Lamin B Receptor

28S -

18S-

Nuclear envelopes from turkey erythrocytes were extracted with 1 M NaCl and 8 M urea and subjected to SDS-PAGE. The polypeptide band corresponding to p58 was electroeluted from the gels. The electroeluted p58 migrated as a single Coomassie blue-stained band on SDS-polyacrylamide gels (Fig. 2), and was recognized by anti-lamin B receptor antibodies (data not shown).

The amino acid sequences of the amino terminus (Fig. 3 A) and also three V8 proteolytic fragments (see Fig. 5) derived from electroeluted p58 were determined by gas-phase microsequencing. A pair of partially degenerate oligonucleotides for use as PCR primers was synthesized based on the amino terminal sequence of p58 (Fig. 3 B, Pl and P2). Chicken liver cDNA was subjected to 25 cycles of PCR amplification using primers P1 and P2, and a product of the expected size was obtained. This product encoded the turkey p58 amino acid sequence situated between the two primers determined by protein sequencing, except for the Arg at position 15, which was deduced to be a Trp based on the cDNA coding sequence (Fig. 3 C). This may represent an interspecies difference between turkeys (protein sequence) and chickens (cDNA sequence). A 25-mer oligonucleotide (P3) corresponding to a portion of the sense strand of the PCR product was synthesized (Fig. 3 D) and used in subsequent chicken library screening and Northern blotting.

1 A. ProAsnArgLysTy	10 yrAlaAspGlyGluValValMetGlyArg <u>Arg</u> Pro		20 LeuTyrTyrG	27 luValGlnValThr
B.5'CCCAACCGAAAAT/			5	TCCAGGTTCA 5'
TAGG C T	C (P1)	(P2)	G G	ТСС А
C.	5'CGATGGCGAGGTGGTGATGGGTCGTTGGCC. 3'GCTACCGCTCCACCACTACCCAGCAACCGG			
D.	5'CGATGGCGAGGTGGTGATGGGTCGT 3'	(P3)		

Figure 3. Strategy for the synthesis of a specific PCR amplification probe for the lamin B receptor (p58). (A) Amino terminal amino acid sequence of the lamin B receptor determined by gas-phase microsequencing either of electroeluted protein (see Fig. 2) or of p58 transferred to polyvinylidene difluoride membrane from

SDS-PAGE preparations of urea-extracted turkey erythrocyte nuclear envelopes. The underlined Arg (position 15) is the only amino acid residue that is different than those deduced from the cDNA sequence (see C, below) that encodes for a Trp at this position. (B) The amino acid sequence in A was used to synthesize partially degenerate sense (P1) and antisense (P2) primers for PCR. (C) Nucleotide sequence of the PCR amplification product using primers P1 and P2 and cDNA reverse transcribed from chicken liver $poly(A^+)$ -enriched RNA as template. (D) Sense strand oligonucleotide (P3) used to screen cDNA library. These sequence data are available from EMBL/GenBank/DDBJ under accession number Y0082.

-88	GCAC	CGCC	CTC	GGGG	CGGG	TCCC	GCCG	CACC	GGCA	GCCC	CCTC	CTGC	CTGI	CCCG	GTGG	GGCC	GGCG	CAGO	TATI	GCTI	TTGO	SAAA
1 1	ATG MET	CCA Pro	AAC Asn	CGG Arg	AAG Lys	TAT Tyr	GCC Ala	GAT Asp	GGC Gly	GAG Glu	GTG Val	GTG Val	ATG Met	GGT Gly	CGT Arg	TGG Trp	CCA Pro	G GA Gly	AGC Ser	GTC Val	CTG Leu	TAC Tyr
67 23	TAT Tyr	GAA Glu	GTG Val	C AA Gln	GTG Val	ACT Thr	AGT Ser	TAT Tyr	GAT Asp	GAT Asp	GCT Ala	TCT Ser	C AT His	CTT Leu	TAT Tyr	ACT Thr	GTG Val	AAG Lys	TAC Tyr	AAA Lys	GAT Asp	GGT Gly
133 45	ACC Thr	GAA Glu	CTT Leu	GCT Ala	TTG Leu	AA G Lys	G AA Glu	AGT Ser	GAT Asp	ATA Ile	AGG Arg	TTA Leu	C AG Gln	TCA Ser	TCA Ser	TTC Phe	AAG Lys	C AG Gln	AGG Arg	AAA Lys	AGC Ser	CAG Gln
199 67	TCT Ser	TCC Ser	TCA Ser	AGT Ser	TCT Ser	CCT Pro	TCC Ser	AGA Arg	AGA Arg	AGT Ser	AGA Arg	AGC Ser	AGA Arg	TCT Ser	CGA Arg	TCC Ser	AGA Arg	TCT Ser	CCT Pro	GGT Gly	CGG Arg	CCA Pro
89	GCA Ala	Lys	Gly	Arg	Arg	Arg	Ser	Ser	Ser	His	Ser	Arg	Glu	His	Lys	Glu	Asp	Lys	Lys	Lys	Ile	Ile
111	CAG Gln	Glu	Thr	Ser	Leu	Ala	Pro	Pro	Lys	Ρτο	Ser	Glu	Asn	Asn	Thr	Arg	Arg	Tyr	Asn	Gly	Glu	Pro
133	GAC Asp	Ser	Thr	Glu	Arg	Asn	Asp	Thr	Ser	Ser	Lys	Leu	Leu	Glu	Gln	Gln	Lys	Leu	Lys	Pro	Asp	Val
155	GAA Glu	Met	Glu	Arg	Val	Leu	Asp	Gln	Tyr	Ser	Leu	Arg	Ser	Arg	Arg	Glu	Glu	Lys	Lys	Lys	Glu	Glu
177	ATC Ile	Tyr	Ala	Glu	Lys	Lys	Ile	Phe	Glu	Ala	Ile	Lys	Thr	Pro	Glu	Lys	Pro	Ser	Ser	Lys	Thr	Lys
199	GAG Glu	Leu	Glu	Phe	Gly	Gly	Arg	Phe	Gly	Thr	Phe	Met	Leu	Met	Phe	Phe	Leu	Pro	Ala	Thr	Val	Leu
221	TAC Tyr	Leu	Val	Leu	Met	Cys	Lys	Gln	Asp	Asp	Pro	Ser	Leu	Met	Asn	Phe	Pro	Pro	Leu	Pro	Ala	Leu
243	GAA Glu	Ser	Leu	Trp	Glu	Thr	Lys	Val	Phe	Gly	Val	Phe	Leu	Leu	Trp	Phe	Phe	Phe	Gln	Ala	Leu	Phe
265	TAC Tyr CGG	Leu	Leu	Pro	Ile	Gly	Lys	Val	Val	Glu	Gly	Leu	Pro	Leu	Ser	Asn	Pro	Arg	Lys	Leu	Gln	Tyr
287	Arg	Ile								Leu	Thr											
0.0 E	mme	C 3 3		C 3 77	TT A TT	TTC.	TT A TT	CAT	CAC	ጥጥር	GTG	CAG	ጥጥጥ	GCA	GTG	TCA	GCT	GCA	GCT	TTT	TCT	ATG
309	Phe	Glu	Leu	His		Leu	Tyr	Asp	His	Phe	Val	Gln	Phe	Ala	Val	Ser	Ala	Ala	Ala	Phe	Ser	Met
309 991 331	Phe GCA Ala	Glu TTG Leu	Leu AGC Ser	His ATC Ile	Tyr TAT Tyr	Leu TTG Leu	Tyr TAC Tyr	Asp ATT Ile	His CGA Arg	Phe TCC Ser	Val TTG Leu	Gln AAG Lys	Phe GCA Ala	Ala CCT Pro	Val GAG Glu	GAA Glu	Ala GAC Asp	Ala CTA Leu	Ala GCG Ala	Phe CCA Pro	Ser GGT Gly	GGA Gly
309 991 331 1057 353 1123	Phe GCA Ala AAT Asn GAC	Glu TTG Leu TCT Ser CTC	Leu AGC Ser GGA Gly AAA	His ATC Ile TAT Tyr TAC	Tyr TAT Tyr CTT Leu TTC	Leu TTG Leu GTT Val TGT	Tyr TAC Tyr TAT Tyr GAG	Asp ATT Ile GAC Asp TTA	His CGA Arg TTC Phe CGT	Phe TCC Ser TTT Phe CCA	Val TTG Leu ACT Thr GGA	Gln AAG Lys GGA Gly TTA	Phe GCA Ala CAC His ATT	Ala CCT Pro GAA Glu GGC	Val GAG Glu TTG Leu TGG	GAA Glu AAC Asn GTT	Ala GAC Asp CCT Pro GTC	Ala CTA Leu CGT Arg ATA	Ala GCG Ala ATT Ile AAC	Phe CCA Pro GGC Gly TTG	GGT Gly AGT Ser GCA	Met GGA Gly TTT Phe ATG
309 991 331 1057 353 1123 375 1189	Phe GCA Ala AAT Asn GAC Asp CTC	Glu TTG Leu TCT Ser CTC Leu TTG	Leu AGC Ser GGA Gly AAA Lys GCT	His ATC Ile TAT Tyr TAC Tyr GAG	Tyr TAT Tyr CTT Leu TTC Phe ATG	Leu TTG Leu GTT Val TGT Cys AAG	Tyr TAC Tyr TAT Tyr GAG Glu ATA	Asp ATT Ile GAC Asp TTA Leu CAC	His CGA Arg TTC Phe CGT Arg AAT	Phe TCC Ser TTT Phe CCA Pro	Val TTG Leu ACT Thr GGA Gly AGT	Gln AAG Lys GGA Gly TTA Leu ATG	Phe GCA Ala CAC His ATT Ile CCA	Ala CCT Pro GAA Glu GGC Gly TCG	Val GAG Glu TTG Leu TGG Trp CTG	Ser GAA Glu AAC Asn GTT Val TCA	Ala GAC Asp CCT Pro GTC Val ATG	Ala CTA Leu CGT Arg ATA Ile ATA	Ala GCG Ala ATT Ile AAC Asn CTC	Phe CCA Pro GGC Gly TTG Leu GTG	Ser GGT Gly AGT Ser GCA Ala AAC	Met GGA Gly TTT Phe ATG Met AGC
309 991 331 1057 353 1123 375 1189 397 1255	Phe GCA Ala AAT Asn GAC Asp CTC Leu TTT	Glu TTG Leu TCT Ser CTC Leu TTG Leu	Leu AGC Ser GGA Gly AAA Lys GCT Ala	His ATC Ile TAT Tyr TAC Tyr GAG Glu CTA	Tyr TAT Tyr CTT Leu TTC Phe ATG Met	Leu TTG Leu GTT Val TGT Cys AAG Lys GTG	Tyr TAC Tyr TAT Tyr GAG Glu ATA Ile GTG	Asp ATT Ile GAC Asp TTA Leu CAC His GAT	His CGA Arg TTC Phe CGT Arg AAT Asn GCT	Phe TCC Ser TTT Phe CCA Pro CAA Gln CTT	Val TTG Leu ACT Thr GGA Gly AGT Ser TGG	Gln AAG Lys GGA Gly TTA Leu ATG Met	Phe GCA Ala CAC His ATT Ile CCA Pro GAG	Ala CCT Pro GAA Glu GGC Gly TCG Ser GAA	Val GAG Glu TTG Leu TGG Trp CTG Leu GCT	GAA Glu AAC Asn GTT Val TCA Ser GTT	Ala GAC Asp CCT Pro GTC Val ATG Met	Ala CTA Leu CGT Arg ATA Ile ATA Ile	Ala GCG Ala ATT Ile AAC Asn CTC Leu ACA	Phe CCA Pro GGC Gly TTG Leu GTG Val ATG	Ser GGT Gly AGT Ser GCA Ala AAC Asn GAC	Met GGA Gly TTT Phe ATG Met AGC Ser ATT
309 991 331 1057 353 1123 375 1189 397 1255 419	Phe GCA Ala AAT Asn GAC Asp CTC Leu TTT Phe	Glu TTG Leu TCT Ser CTC Leu TTG Leu CAG Gln	Leu AGC Ser GGA Gly AAA Lys GCT Ala CTT Leu	His ATC Ile TAT Tyr TAC Tyr GAG Glu CTA Leu	Tyr TAT Tyr CTT Leu TTC Phe ATG Met TAT Tyr	Leu TTG Leu GTT Val TGT Cys AAG Lys GTG Val	Tyr TAC Tyr TAT Tyr GAG Glu ATA Ile GTG Val	Asp ATT Ile GAC Asp TTA Leu CAC His GAT Asp	His CGA Arg TTC Phe CGT Arg AAT Asn GCT Ala	Phe TCC Ser TTT Phe CCA Pro CAA Gln CTT Leu	Val TTG Leu ACT Thr GGA Gly AGT Ser TGG Trp	Gln AAG Lys GGA Gly TTA Leu ATG Met AAT Asn	Phe GCA Ala CAC His ATT Ile CCA Pro GAG Glu	Ala CCT Pro GAA Glu GGC Gly TCG Ser GAA Glu	Val GAG Glu TTG Leu TGG Trp CTG Leu GCT Ala	GAA Glu AAC Asn GTT Val TCA Ser GTT Val	Ala GAC Asp CCT Pro GTC Val ATG Met TTG Leu	Ala CTA Leu CGT Arg ATA Ile ATA Ile ACT Thr	Ala GCG Ala ATT Ile AAC Asn CTC Leu ACA Thr	Phe CCA Pro GGC Gly TTG Leu GTG Val ATG Met	GGT Gly AGT Ser GCA Ala AAC Asn GAC Asp	Met GGA Gly TTT Phe ATG Met AGC Ser ATT Ile
309 991 331 1057 353 1123 375 1189 397 1255 419 1321 441 1387	Phe GCA Ala AAT Asn GAC Asp CTC Leu TTT Phe ACC Thr CTA	Glu TTG Leu TCT Ser CTC Leu TTG Leu CAG Gln His CAT	Leu AGC Ser GGA Gly AAA Lys GCT Ala CTT Leu GAT ASP GCC	His ATC Ile TAT Tyr TAC Tyr GAG Glu CTA Leu GGA Gly TTC	Tyr TAT Tyr CTT Leu TTC Phe ATG Met TAT Tyr TTT Phe TAT	Leu TTG Leu GTT Val TGT Cys AAG Lys GTG Val GGA Gly TTA	Tyr TAC Tyr TAT Tyr GAG Glu ATA Ile GTG Val TTC Phe GTT	Asp ATT Ile GAC Asp TTA Leu CAC His GAT Asp ATG Met GGA	His CGA Arg TTC Phe CGT Arg AAT Arg CGT Arg CGT Arg CTA Leu CAC	Phe TCC Ser TTT Phe CCA Gln CTT Leu GCC Ala CCT	Val TTG Leu ACT Thr GGA Gly AGT Ser TGG Trp TTT Phe ATT	Gln AAG Lys GGA Cly TTA Leu ATG Met AAT Asn GGA Gly GCA	Phe GCA Ala CAC His ATT Ile CCA Pro GAG Glu GAT ASP	Ala CCT Pro GAA Glu GGC Gly TCG Ser GAA Glu TTG Leu TCT	Val GAG Glu TTG Leu CTG Leu GCT Ala GTG Val	Ser GAA Glu AAC Asn GTT Val TCA Ser GTT Val TGG Trp CCG	Ala GAC Asp CCT Pro GTC Val ATG Met TTG Leu GTT Val GTT	Ala CTA Leu CGT Arg Arg Ile ATA Ile ATA Ile CCA Pro GCA	Ala GCG Ala ATT Ile AAC Asn CTC Leu ACA Thr TTT Phe GCT	Phe CCA Pro GGC Gly TTG Leu Val ATG Met GTC Val GCA	Ser GGT Gly AGT Ser GCA Ala AAC Asn GAC Asp TAC Tyr ATT	Met GGA Gly TTT Phe ATG Met AGC Ser ATT Ile AGC Ser ACT
309 991 331 1057 353 1123 375 1189 397 1255 419 1321 441 1387 463 1453	Phe GCA Ala AAT Asn GAC Leu TTT Phe ACC Thr CTA Leu ATT	Glu TTG Leu TCT Ser CTC Leu TTG Leu CAG Gln CAT His CAG Gln CTG	Leu AGC Ser GGA Gly AAAA Lys GCT Ala CTT Leu GAT Asp GCC Ala AAC	His ATC Ile TAT Tyr CAG Glu CTA Leu GGA Gly TTC Phe	Tyr TAT Tyr CTT Leu TTC Phe ATG Met TAT Tyr TTT Phe TAT Tyr	Leu TTG Leu GTT Val TGT Cys AAG Lys GTG Val GGA GIy TTA Leu GGT	Tyr TAC Tyr TAT Tyr GAG Glu ATA Ile GTG Val TTC Phe GTT Val TAT	Asp ATT Ile GAC Asp TTA Leu CAC His GAT Asp ATG Met GGA GIY TAC	His CGA Arg TTC Phe CGT Arg AAT Asn GCT Ala CTA Leu CAC His ATA	Phe TCC Ser TTT Phe CCA Pro CCA CTT Leu GCC Ala CCT TTC	Val TTG Leu ACT Thr GGA Gly AGT Ser TGG Trp TTT Phe ATT Ile CGT	Gln AAG Lys GGA Gly TTA Leu ATG Met AAT ASn GGA Gly GCA Ala AGT	Phe GCA Ala CAC His ATT Ile CCA Pro GAG Glu GAT Asp ATT Ile GCA	Ala CCT Pro GAA Glu GGC Gly TCG Ser GAA Glu TTG Leu TCT Ser AAT	Val GAG Glu TTG Leu TGG Trp CTG Leu GCT Ala GTG Val TGG Trp TCT	Ser GAA Glu AAC Asn GTT Val TCA Ser TCA Ser CCG Trp CCG Pro CAG	Ala GAC Asp CCT Pro GTC Val ATG Met TTG Leu GTT Val GTT Val AAA	Ala CTA Leu CGT Arg ATA Ile ATA Ile ATA Thr CCA Pro GCA Ala AAC	Ala GCG Ala ATT Ile AAC Asn CTC Leu ACA Thr TTT Phe GCT Ala AAT	Phe CCA Pro GGC Gly TTG Leu GTG Val ATG Met GTC Val GCA Ala TTC	GGT Gly AGT Ser GCA Ala AAC Asn GAC Asp TAC Tyr Ile CGA	Met GGA Gly TTT Phe ATG Met AGC Ser ATT Ile AGC Ser ACT Thr AGG
309 991 331 1057 353 1123 375 1189 397 1255 419 1321 441 1387 463 1453 485 1519	Phe GCA Ala AAT Asn GAC Asp CTC Leu TTT Phe ACC Thr CTA Leu ATT Ile AAT	Glu TTG Leu TCT Ser CTC Leu TTG Leu CAG Gln CAT His CAG Gln CTG Leu CCA	Leu AGC Ser GGA Gly AAA Lys GCT Ala CTT Leu GAT Asp GCC Ala AAC Asn GCA	His ATC Ile TAT Tyr CAG Glu CTA Leu GGA Gly TTC Phe TGT Cys GAT	Tyr TAT Tyr CTT Leu TTC Phe ATG Met TAT Tyr TYr TAT Tyr TAT TILE CCC	Leu TTG Leu GTT Val TGT Cys Lys GTG Val GGA Gly TTA Leu GGT Gly AAA	Tyr TAC Tyr TAT Tyr GAG Glu ATA Ile GTU Val TTC Phe GTT Val TAT Tyr TTG	Asp ATT Ile GAC Asp TTA Leu CAC His GAT Asp ATG GAT Asp TTA CIC TCC	His CGA Arg TTC Phe CGT Arg Arg Arg Arg Arg Arg Arg Arg CTA Leu CAC His ATA Ile TAT	Phe TCC Ser TTT Phe CCA Gln CTT Leu GCC Ala CCT Pro TTC Phe CTG	Val TTG Leu ACT Thr GGA Gly AGT Ser TGG Trp TTT Phe ATT Ile CGT AAA	Gln AAG Lys GGA Gly TTA Leu ATG Met Asn Gly GCA Ala AGT Ser GTT	Phe GCA Ala CAC His ATT Ile CCA Pro GAG Glu GAT Asp ATT Ile GCA Ala	Ala CCT Pro GAA Glu GGC Gly TCG Ser TCG Ser TTG Leu TCT Ser AAT Asn CCC	Val GAG Glu TTG Leu TTG Trp CTG Leu CTG Leu GCT Ala GTG Val TGG Trp TCT Ser ACT	Ser GAA Glu AAC Asn GTT Val TCA Ser TTP CCG GTT Val CCG GIn GCG	Ala GAC Asp CCT Pro Val ATG Met TTG Leu GTT Val GTT Val AAA Lys ACT	Ala CTA Leu CGT Arg ATA Ile ATA Ile ATA Thr CCA Pro GCA Ala AAC Asn GGA	Ala GCG Ala ATT Ile AAC ASn CTC Leu ACA Thr TTT Phe GCT Ala AAT AAA	Phe CCA Pro GGC Gly TTG Leu GTG Val ATG Met GTC Val TTC Phe GGG	Ser GGT Gly AGT Ser GCA Ala AAC Asn TAC Tyr TAC Tyr CGA ATG CTT	Met GGA Gly TTT Phe ATG Met AGC Ser ATT Ile AGC Ser ACT Thr AGG Arg CTT
309 991 331 1057 353 1123 375 1189 397 1255 419 1321 441 1387 463 1453 4R5 1519 507	Phe GCA Ala AAT Asn GAC Asp CTC Leu TTT Phe ACC Thr CTA Leu ATT Ile AAT GTC	Glu TTG Leu TCT Ser CTC Leu TTG Leu CAG Gln CAT His CAG Gln CTG Leu CAG Gln CTG D Leu	Leu AGC Ser GGA Gly AAA Lys GCT Ala GCT Ala GCC Ala AAC Asn GCC Ala GCC Ala	His ATC Ile TAT Tyr TAC Tyr GAG Glu CTA Leu CTA Leu GGA Gly TTC Cys GAT Asp TGG	Tyr TAT Tyr CTT Leuu TTC Phe ATG Met TAT Tyr TAT Tyr TAT Tyr TAT Tyr TAT Tor Fhe CCC Pro TGG	Leu TTG Leu GTT Val TGT Cys AAG Lys GTG Cya GGA Gly TTA Leu GGT Gly Lys GGG	Tyr TAC Tyr TAT Tyr GAG Glu ATA Ile GTG Val TTC Phe GTT Val TAT Tyr TTG Leu TTT	Asp ATT Ile GAC Asp TTA Leu CAC His GAT Asp ATG GAT TAC CJY TCC Ser GTT	His CGA Arg TTC Phe CGT Arg Aar Asn GCT Ala CTA Leu CAC His ATA Ile TAT Tyr CGT	Phe TCC Ser TTT Phe CCA Gln CTT Leu GCC Ala CCT Pro TTC Phe CCG Leu CAC	Val TTG Leu ACT Thr GGA Gly AGT Ser TGG Trp TTT Phe ATT Ile CGT Arg AAA Lys CCC	Gln Lys GGA Gly TTA Leu ATG Met ASn Gly GGA Ala AGT Ser GTT Val AAT	Phe GCA Ala CAC His ATT Ile CCA Pro GAG Glu GAT Asp ATT Ile GCA Ala Ala TAC	Ala CCT Pro GAA Glu GGC Gly TCG Ser TCG Ser TCT Ser AAT Asn CCC Pro CTT	Val GAG Glu TTG Leu TGG Trp CTG Leu GCT Ala GTG Val TGG Trp TCT Ser ACT Thr	Ser GAA Glu AAC Asn GTT Val TCA Ser TTP TCG TTP CCG GIn GCG GIn GCG Ala GAT	Ala GAC Asp CCT Pro GTC Val ATG Met TTG Leu GTT Val GTT Val AAA Lys ACT Thr	Ala CTA Leu CGT Arg ATA Ile ATA Ile ATA Thr CCA Pro GCA Ala AAC Asn GGA GIY ATC	Ala GCG Ala ATT Ile AAC Asn CTC Leu ACA Thr TTT Phe GCT Ala AAT ASN AAA Lys ATG	Phe CCA Pro GGC Gly TTG Leu GTG Val ATG Met GTC Val GCA Ala TTCC Phe GGG Gly GCA	Ser GGT Gly AGT Ser GCA Ala AAC Asn TAC Tyr ATT Ile CGA Arg CTT Leu CTT	Met GGA Gly TTT Phe ATG Met AGC Ser ATT Ile AGC Ser ATT Thr AGG ATG CTT Leu GCG
309 991 331 1057 353 1123 375 1189 397 1255 419 1321 441 1387 463 1453 4R5 1519 507 1585 529	Phe GCA Ala AAT Asn GAC Asp CTC Leu TTT Phe ACC Thr CTA Leu ATT Ile AAT Asn GTC Val	Glu TTG Leu TCT Ser CTC Leu TTG Gln CAT His CAG Gln CTG Leu CAG Gln CTG Leu TTG TTG TTG TTG TTG CAG Gln CTG TTG CAG CAG CAG CAG CTG CTG CTC CTC TTG CTC TCT TTG CTC TCT TTG CTC TCT TTG CTC TCT TTG CTC TCT TTG CTC CTC	Leu AGC Ser GGA Gly AAA Lys GCT Ala CTT Leu GAT Asp GCC Asn GCC Asn GCC Asn GCC Asn CTT CTT	His ATC Ile TAT Tyr GAG Glu CTA Leu GGA Glu TTC Cys GAT Asp TGG Trp CCC	Tyr TAT Tyr CTT Leu TTC Phe ATG Met TAT Tyr Tyr Tyr Tyr Tyr Tyr Tyr Tyr Tyr Ty	Leu TTG Leu GTT Val TGT Cys AAG Lys GTG Val CGT GIY Leu GGT GIY AAA Lys GGG GIY GGT	Tyr TAC Tyr TAT Tyr GAG Glu ATA Ile GTG Val TTC Phe GTT Val TAT Tyr TTG Leu TTT Phe TTC	Asp ATT Ile GAC Asp TTA Leu CAC His GAT Asp ATG GAT ATG GIY TCC Ser GTT Val AAT	His CGA Arg TTC Phe CGT Arg Arg Asn Asn Asn Asn Ata Asn Ata Asn CTA Leu CAC CAC CGT CAC CAC	Phe TCC Ser TTT Phe CCA Pro CAA Gln CTT Leu GCC Ala CCT Pro TTC Phe CTG Leu CAC His ATT	Val TTG Leu ACT Thr GGA Gly AGT Ser TGG Trp TTT Phe ATT Ile CGT ATA Lys CCC Pro TTG	Gln AAG Lys GGA Gly TTA Leu ATG Met AAT ASn GGA Gly GCA Ala AGT Ser GTT Val AAT ASn CCA	Phe GCA Ala CAC His Alt Ile CCA Pro GAG Glu GAT Asp ATT Ile GCA Ala Ala ATA TAT TAT	Ala CCT Pro GAA Glu GGU TCG Ser TCG Ser TCG Ser TCG Ser TCC CTT Leu TCC CTT Leu	Val GAG Glu TTG Leu TTG Trp CTG Leu CTG Leu CTG Leu CTG Leu CTG Leu CTG CTG Leu TTT TTT TTT TTT TTT TCT Ser ALTTT TTT TTT TTT TTT TTT TTT TTT TTT T	Ser GAA Glu AAC Asn GTT Val TCA Ser TTP Val TGG Trp Pro CCG Gln GCG Ala GAT Asp	Ala GAC Asp CCT Pro GTC Val ATG Met TTG Leu GTT Val GTT Val AAA Lys ACT Thr ATC Ile ATT	Ala CTA Leu CGT Arg Arg Arg Ile ATA Ile ATA Thr CCA Pro GCA Ala AAC Asn GGA Cly Ile TAT	Ala GCG Ala ATT Ile AAC Asn CTC Leu ACA Thr TTT Phe GCT Ala AAT ASN ALys ATG Met	Phe CCA Pro GGC Gly TTG Leu GTG Val ATG Met CVal GCA Ala GGG Gly GCA Ala ATC	Ser GGT Gly AGT Ser GCA Ala Ala AAC Asn GAC Asn TAC Tyr Ile CGA Arg CTT Leu CTT Leu TGC	Met GGA Gly TTT Phe ATG Met AGC Ser ATT Ile AGC Ser ATT Thr AGG Arg CTT Leu GCG Ala TTG
309 991 331 1057 353 375 1123 375 1189 397 1255 419 1321 441 1387 463 4453 4453 4453 1519 507 1585 529 1651 551	Phe GCA Ala AAT ASN GAC ASP CTC Leu TTT Phe ACC Thr CTA Leu ATT Ile AAT ASN GTC Val TTG GTC CTT	Glu TTG Leu TTG CTC Leu TTG Leu CAG Gln CAT His CAG Gln CTG Leu CAG GIN CTG CTG CTT TTC Ser TTC TTC CTG CTG CTG CTG CTG CTG CTG CTG	Leu AGC Ser GGA Gly AAA Lys GCT Ala GCT Ala CTT Leu GAT ASp GCC Ala AAC ASn GCA Ala CTT LEU CTA LEU	His ATC Ile TAT Tyr GAG Glu CTA Leu CTA Leu GGA Glu TCC Phe TGT TGG TTC Cys GAT TGG Trp CCC CCC CCC	Tyr TAT Tyr CTT Leu TTC Phe ATG Met TAT Tyr TAT Tyr TAT Tyr TAT Tyr TGG Trp TGT Cys GAA	Leu TTG Leu Val TGT Cys AAG Lys GTG Val GGZ GIY Leu GGT GIY GGT GIY GCT	Tyr TAC Tyr TAT Tyr GAG Glu ATA Ile GGG Val TTC Phe GTT Val TAT Tyr TTG Leu TTT Phe CGT	Asp ATT Ile GAC Asp TTA Leu CAC His GAT Asp ATG GAT Tyr TCC Ser GTT Val AAT Asn GAT	His CGA Arg TTC Phe CGT Arg Arg Arg Arg Arg CAT His CGT Arg CAC His CGA AGA	Phe TCC Ser TTT Phe CCA Pro CCA Gln CTT Leu GCC Ala CTT Fro TTC Phe CTG Leu CAT TIC CAT	Val TTG Leu ACT Thr GGA Gly AGT Ser TGG Trp Phe ATT Ile CGT AACG TTG Leu CCC CAC	Gln AAG Lys GGA Gly TTA Leu ATG Met AAT ASn GGA Ala AGT Ser GTT Val AAT ASn CCA Pro TGT	Phe GCA Ala CAC His Tile CCA Pro GAG Glu GAT Asp ATT Ile GCA Ala AITA TIC TYr TAT TYr AAG	Ala CCT Pro GAA Glu GGC Gly TCG Ser GAA Glu TTG Leu TCT Ser AAT Asn CCC CTT Leu TTC Pro	Val GAG Glu TTG Leu TTG Trp CTG Leu CTG Leu GCT Ala GTG Val TCT Ser Thr GGT Gly TAT Tyr AAG	Ser GAA Glu AAC Asn GTT Val TCA Ser Val TCG Trp CCG GIn GCG GIn GCG Ala GAT Asp GTG Val	Ala GAC Asp CCT Pro GTC Val ATG Met TTG Leu GTT Val GTT Val AAA Lys ACT Thr ATC Ile GGT	Ala CTA Leu CGT Arg ATA Ile ATA Ile ATA Thr CCA Pro GCA Ala ASn GGA GGA GIY ATC Ile TAT TYr	Ala GCG Ala ATT Ile AAC Asn CTC Leu ACA Thr TTT Phe GCT Ala AAT Asn AAA Lys ATG Met TTC Phe GCA	Phe CCA Pro GCC Gly TTG Leu GTG Val ATG Met CVal GCA Ala TTCC Phe GGG Gly GCA Ala ATC Ile TTG	Ser GGT Gly AGT Ser GCA Ala AAC Asn GAC Asn TAC Tyr CTT Leu CTT Leu CTT Leu CTT GC AA	Met GGA Gly TTT Phe ATG Met AGC Ser ATT Ile AGC Ser ATT Ile AGC Ser ATT Leu GCG Ala TTG Leu AGG
309 991 331 1057 353 375 1123 375 1129 397 1255 419 1321 441 1387 463 1453 485 1519 507 1585 529 1651 551 1717 573 1783	Phe GCA Ala AAT ASN GAC ASP CTC Leu TTT Phe ACC Thr CTA Leu ATT Ile AAT TGG CTC Leu TTT TTP CTT Leu TAT	Glu TTG Leu TTG Ser CTC Leu TTG Gln CAG Gln CAG Gln CAG Gln CAG CAG Gln CTG CAG CTG Ser TTC TCC CCA GIN TTG TTG TTG TTG TTG TTG TTG TTG TTG Ser	Leu AGC Ser GGA Gly AAA Lys GCT Ala CTT Leu GAT ASp GCC Ala ASn GCC Ala CTT Leu CTT Leu CTT Leu CTT Leu CTT CTT CTT CTT CTT CTT CTT CTT CTT CT	His ATC Ile TAT Tyr GAG Glu CTA Leu CTA Leu GGA Glu TCC Phe TGT TGG TTC Phe Cys GAT TAG Trp CCC SC CCC CCC CCC CCC CCC CCC CCC CCC	Tyr TAT Tyr CTT Leu TTC Phe ATG Met TAT Tyr TAT Tyr TAT Tyr TAT Tyr CCC Fro TGG Trp TGT Cys GAA GIU	Leu TTG Leu Val TGT Cys AAG Lys GTG Val GGA Gly TTA Leu GGT Gly GGT Gly GGT Ala CCA	Tyr TAC Tyr TAT Tyr GAG Glu ATA Ile GTG Val TTC Phe GTT Val TAT Tyr TTG Leu TTT Phe CGT Arg	Asp ATT Ile GAC Asp TTA Leu CAC His GAT Asp ATG GAT Tyr TCC Ser GTT Val AAT Asp ACG	His CGA Arg TTC Phe CGT Arg Arg Arg Arg CGT Ala CTA Leu CAC His TTT Tyr CGT Arg CAC His CAC CAC	Phe TCC Ser TTT Phe CCA Pro CCA Gln CTT Leu GCC Ala CTT Pro TTC Phe CCG His ATT Ile CAT His ATT	Val TTG Leu ACT Thr GGA Gly AGT Ser TGG Trp Phe ATT Ile CGT Arg AAA Lys CCC Pro TTG Leu CAC	Gln AAG Lys GGA Gly TTA Leu ATG Met AAT ASn GGA Gly GGA Ala AGT Ser GTT Val AAT ASn CCA Pro TGT Cys	Phe GCA Ala CAC His Ile CCA Pro GAG Glu GAT Asp ATT Ile GCA Ala ATA Ile TAC Tyr TAT Tyr CAT	Ala CCT Pro GAA Glu GGC Gly TCG Ser TCG Ser TTG Leu TTG Ser CCT Leu TTC Pro CTT Leu CCT Leu CCT TCC CTT CCT CCT CCT CCT CCT CCT CC	Val GAG Glu TTG Leu TTG Trp CTG Leu GCT Ala GTG Val TCT Ser TCT Ser Thr GGT Gly TAT Tyr AAG Lys CTA	Ser GAA Glu AAC Asn GTT Val TCA Ser TCA GTT Val TCG Trp CCG GIn GCG Ala GAT Asp GTG Val TTT Tyr	Ala GAC Asp CCT Pro GTC Val ATG Met TTG Leu GTT Val GTT Val AAA ACT Thr ATC Ile GGT Gly CAC	Ala CTA Leu CGT Arg ATA Ile ATA Ile ATA Thr CCA Pro GCA Ala ASn GGA ASn GGA Cly TTT TTG Leu TCC	Ala GCG Ala ATT Ile AAC Asn CTC Leu ACA Thr TTT Phe GCT Ala Asn AAA Lys ATG Met TTC Phe GCA Ala	Phe CCA Pro GCC Gly TTG Leu GTG Val ATG Met GCA Ala GCA Ala Ala ALC LILE TTGG Trp TAC	Ser GGT Gly AGT Ser GCA Ala AAC Asn GAC Asn TAC Tyr CTT Leu CCTA CCYS GAA Glu CTG	Met GGA Gly TTT Phe ATG Met AGC Ser ATT Ile AGC Ser ATT Leu GCG Ala TTG Leu AGG Arg ATG
309 991 331 1057 353 375 1123 397 1255 419 1321 441 1387 463 441 1387 463 441 1387 507 1585 529 1651 551 1717 573 1783 595	Phe GCA Ala AAT Asn GAC Asp CTC Leu TTT Phe ACC Thr CTA Leu ATT Ile AAT TGG CTT Leu TTG TTT CTT Leu TTT TTT	Glu TTG Leu TTG Ser CTC Leu CAG Gln CAT His CAG Gln CAT His CAG Gln CTG CTG Leu CAG GIN TTG TCT Ser TTC TTC CCA CAG GIN CAG CTC CTC CAG CTC CTC CTC Leu CAG CTC CTC CTC Leu CAG CTC CTC CTC CTC CTC CTC CTC CTC CTC CT	Leu AGC Ser GGA Gly GGT ALA ALYS GCT Ala CTT Leu CTT Asp GCC Ala Asn GCC Ala CTT Leu CTT Leu CTT Leu CTT Leu CTT CTT CTT CTT CTT CTT CTT CTT CTT CT	His ATC Ile TAT Tyr GAG Glu CTA Leu CTA CGA Gly TCC Phe Cys GAT Asp TCC Pro CGA Arg CGT Arg	Tyr TAT Tyr CTT Leu ATG Met TAT Tyr TTT Phe TAT Tyr TAT Tyr CCCC Pro ATT Ile CCCC Pro GAA Glu GTA Val	Leu TTG Leu Val TGT Cys AAG Lys GTG Val Cys Cys Cys Cys Cys Cys Cys Cys Cys Cys	Tyr TAC Tyr Tyr GAG Glu ATA Ile GTU Val TTC Phe GTT Val TAT Tyr TTG Leu TTT Phe CGT Arg TAC Tyr TCC	Asp ATT Ile GAC Asp TTA Leu CAC His GAT Asp ATG GAT Asp TCC Ser GTT TCC Ser GTT Val Asp AASP ACG Thr CAC	His CGA Arg TTC Phe CGT Arg Arg Arg CGT Ala CTA Leu CAC His ATA Tyr CGT Tyr CGT His CAC His CAC CAC His CAC CAC CAC CAC CAC CAC CAC CAC CAC CA	Phe TCC Ser TTT Phe CCA Gln CTT Leu GCC Ala CTT TCC Pro TTC Phe CTG Leu CAT His ATT Ile CAT TIC CAT TTC TTC TTC TTC TTC CAA CTT TTC CAA CTT TTT T	Val TTG Leu ACT Thr GGA Gly AGT TGG Trp TTG TTP Phe ATT Ile CGT AAA Lys CCC Pro CAC His Ser TCC Ser ACT	Gln AAG Lys GGA Gly TTA Leu ATG Met Asn GGA Gly GGA Ala AGT Ser GTT Val AAT Asn CCA Pro TGT Cys TTA Leu TGG	Phe GCA Ala CAC His Tile CCA Pro GAG Glu GAT Asp ATT Ile GAG Ala ATA Ile TAC Tyr TAT Tyr AAG Lys CAT His TCT	Ala CCT Pro GAA Glu GGC Gly TCG Ser CAA CLeu TTG Leu TCT Ser AAT Asn CCCC Pro CTT Leu TTC Phe AAA Lys CTA Leu GTC	Val GAG Glu TTG Trp CTG Trp CTG CTG Leu GCT Ala GTG Val TGG Trp TCT Ser ACT Thr GGT U Lys CTA Leu CTA Leu CTG CTG TTP CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG	Ser GAA Glu AAC Asn GTT Val TCA Ser TTP CCG GIN GCG Ala GAT Asp GTG Val TAT Tyr GAG GIU TAT	Ala GAC Asp CCT Pro GTC Val ATG Met TTG Leu GTT Val GTT Val ATG Lau GTT Val GTT Thr Thr Thr Thr Ile GGT Gly CAC His CCC	Ala CTA Leu CGT Arg ATA Ile ATA Ile ATA Ile ATA Thr CCA Pro GCA Ala ASn GGA Gly TTG Leu TCC Ser GGC	Ala GCG Ala ATT Ile Aan CTC Leu ACA Thr TTT Phe GCT Ala AAA ASN AAA Lys ATG Phe GCA Ala ATT TTT TTT TTT	Phe CCA Pro GCC Gly TTG Caly TTG Val GTG Val GCA Ala GCC Ala TTC Cyal GCA Ala TTC Cleu U TTG GCZ TTG Caly Val GTG CALY CALY CALY CALY CALY CALY CALY CALY	Ser GGT Gly AGT Ser GCA Ala AAC Asp TAC Tyr ATT Ile CGA Arg CTT Leu CTT Leu CTT Cys GAA Glu CTG CLAU	Met GGA Gly TTT Phe ATG Met AGC Ser ATT Ile AGC Ser ATT Ile AGC Ser ATT Thr AGG Arg CTT Leu GCG Ala TTG Leu AGC ATG ATG ATG TTG The ATG AGC Ser ATT AGC Ser ATT AGC Ser ATT AGC Ser ATT AGC Ser ATT AGC Ser ATT AGC Ser ATT AGC Ser ATT Ile AGC Ser ATT AGC Ser ATT AGC Ser ATT AGC Ser ATT AGC Ser ATT AGC Ser ATT AGC Ser ATT AGC Ser ATT AGC Ser ATT AGC Ser ATT AGC Ser ATT AGC Ser ATT AGC Ser ATT AGC Ser ATT AGC Ser ATT AGC AGC Ser ATT AGC ACT ATT AGC AGC ACT ATT AGC ATT AGC ACT ATT AGC ACT ATT AGC ACT ATT AGC ATT ATT AGC ACT ATT AGC ATT ACT ATT AGC ATT ATT AGC ATT ATT AGC ATT ATT AGC ATT ATT AGC ACT ATT AGC ATT ATT ATT AGC ATT ATT ATT ATT AGC ATT ATT ATT ACT ATT ATT ACT ATT ATT AT

Figure 5. Nucleotide sequence of the lamin B receptor cDNA and the deduced amino acid sequence. The amino-terminal methionine residue is indicated as +1 of the amino acid sequence and A of the initiator ATG codon as +1 of the cDNA nucleotide sequence. Regions of the deduced amino acid sequence corresponding to the amino-terminal amino acid sequence and V8 peptide sequences determined in protein sequencing experiments are underlined with thin lines. Hydrophobic amino acid stretches that can form potential membrane spanning domains are underlined with thicker lines. The nucleotide sequence was determined by sequencing four overlapping cDNA clones. Clone DJ-5 was obtained by screening a chick embryonal cDNA library with ³²P-labeled oligonucleotide P3 (shown in Fig. 3). The cDNA insert of clone DJ-5 contained nucleotides -102 through +1760 of the sequence shown. The cDNA insert of clone DJ-5 was ³²P-labeled by random priming and used to screen a chick embryonal erythrocyte cDNA library in lambda gt11, and three shorter overlapping clones (M-3, M-3', and M-7) were characterized. The cDNA inserts of M-3 and M-3' contained nucleotides +1148 to +1987 and that of M-7 contained nucleotides +1655 to +2021. Double-stranded DNA sequencing of both stands of the insert of clone DJ-5 was performed by the dideoxy chain termination method using the Sequenase kit (United States Biochemical Corp.) with both dITP and dGTP. The cDNA inserts of lambda gtl1 clones M-3, M-3' (opposite strand of clone M-3), and M-7 were subcloned into the Eco RI site of M13-mp18 and single-stranded sequencing was performed also using the Sequenase kit. Synthetic oligonucleotide primers were used for sequencing. These sequence data are available from EMBL/Gen-Bank/DDBJ under accession number Y0082.

GGCACGAGCTGAGGA

-102

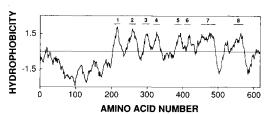


Figure 6. Kyte-Doolittle hydropathy plot for the lamin B receptor. Hydrophobic stretches at least 19 amino acids residues in length with maximal hydrophobic moments of 1.5 or greater are overlined and labeled 1–8. Analysis was performed using the DNA STAR program on an IBM PC-AT computer. The hydrophobic window was set at 19.

Three cross-hybridizing clones were isolated from 400,000 plaques of a lambda Zap chick embryonal cDNA library by screening with ³²P-labeled oligonucleotide P3. The largest of these clones (DJ-5) contained a cDNA insert of 1,862 bp with an open reading frame 1,760-bp long that coded for the lamin B receptor amino terminal sequence and all three V8 peptide fragments. The insert of clone DJ-5 did not contain an in-frame stop codon and was therefore used to screen a chick embryonal erythrocyte cDNA library. Three cross-hybridizing clones were characterized that overlapped DJ-5 and each contained an in-frame stop codon as well as 3' untranslated sequences.

Northern blotting of poly(A⁺)-enriched RNA showed two cross-reactive species of 2.3 and 4.4 kb when probed with the ³²P-labeled cDNA insert of clone DJ-5 (Fig. 4) or oligonucleotide P3 (data not shown). The 2.3 kb RNA is of adequate size to be the mRNA coding for the mature lamin B receptor protein. The 4.4-kb species may represent a differentially polyadenylated mRNA coding for the same protein, or may code for a related protein, perhaps the 96-kD protein that cross-reacts with anti-p58 antibodies (see Fig. 1, *B* and *E*).

Primary Structure of the Lamin B Receptor

A total cDNA sequence of 2,133 bp was obtained from clone DJ-5 and the overlapping clones (Fig. 5). The total sequence contained an open reading frame coding for 637 amino acids containing the amino terminal sequence (starting one amino acid downstream from the deduced initiator methionine) and the sequences of the three V8 protease fragments (Fig. 5, *thin underscoring*). No potential initiation codon is found upstream of the codon for Met (+1) and an in-frame stop codon is located 92 bases upstream from the codon for Met (+1). Thus, the lamin B receptor is synthesized without a cleavable amino terminal signal sequence.

The calculated molecular mass of the lamin B receptor is 73,375 D, which is larger than the apparent molecular mass of 58,000 D estimated from SDS-polyacrylamide gels. This discrepancy is probably due to the known aberrant migration of membrane proteins subjected to SDS-PAGE. However, it cannot be ruled out that the primary translation product is processed in vivo or is degraded during cell fractionation. Both processes would have to be limited to the carboxy terminal region.

Hydropathy analysis (26) of the deduced amino acid sequence shows that the lamin B receptor contains eight hydrophobic segments with membrane spanning potential (Fig. 6). The domains labeled 1-6 in Fig. 6 are typical for transmem-

brane domains in that they are 19-23 amino acids long, have maximal hydropathic indices of 1.5 or greater and are flanked by charged amino acids. The domains labeled 7 and 8 are longer hydrophobic stretches than those typical for membrane spanning domains, and may be perimembranous hydrophobic pockets or could also potentially span the membrane more than once. The finding of multiple potential membrane spanning domains is consistent with the finding that the lamin B receptor is only extractable from membranes with detergents and resistant to extraction with 8 M urea or alkali (Fig. 1 E; also reference 43). The lamin B receptor of turkey erythrocytes does not bind concanavalin A (Worman, H. J., unpublished observation) and therefore does not likely contain Asn-linked oligosaccharides, even though the sequence contains three potential consensus sites for Asnlinked glycosylation at Asn (+123), (+139), and (+405).

Another structural feature of the lamin B receptor sequence is a 204-residue-long amino terminal region from Pro (+2) to Arg (+205) that precedes the first potential transmembrane segment (Figs. 5 and 6). This region contains 79 charged residues and has a calculated pI of 9.89. It also contains two consensus sites (Arg-Arg-X-Ser) for phosphorylation by protein kinase A (13) at Ser (+95) and Ser (+96). Previous studies have shown that the turkey erythrocyte lamin B receptor is phosphorylated by protein kinase A both in vivo and in vitro (3). This amino terminal domain also contains the motif of Ser/Thr-Pro-X-X (at Ser [+71], Ser [+84], and Thr [+190]) that has been proposed to be a DNA-binding unit common to several gene regulatory proteins and histones (39). Two of these Ser/Thr-Pro-X-X sequences (at Ser [+71] and Thr [+190]) also conform to the proposed consensus site for phosphorylation by p34^{cdc2} protein kinase (34); however, it remains to be shown that the lamin B receptor is a substrate for this enzyme. No significant overall sequence identity was seen between the lamin B receptor and other proteins listed in the Dayhoff and translated Genebank protein sequence data bases.

Discussion

We report here the cDNA sequence and deduced primary structure of the chicken lamin B receptor (calculated molecular mass of 73,375 D) and provide further evidence that this protein is located in the inner nuclear membrane. The deduced molecular mass is larger than that previously estimated by electrophoresis (58 kD). In future experiments, a full-length cDNA can be constructed from the overlapping clones isolated in this study and used in cell-free translation assays to determine if this discrepancy is a result of posttranslational processing or aberrant migration on gels. Hydropathy analysis suggests that the lamin B receptor is likely to have several (up to eight) transmembrane segments. The polytopic (7) nature of the lamin B receptor was unexpected as a single transmembrane segment with a single lamin B binding domain exposed to the nucleoplasm would be sufficient to anchor lamin B to the inner nuclear membrane. It therefore remains possible that the lamin B receptor serves an additional function or functions. An example of a multifunctional polytopic integral membrane protein is band 3 of the erythrocyte plasma membrane (24). This protein functions as a Cl⁻/HCO₃⁻ exchanger (8) and also has a cytoplasmically exposed segment that serves as a binding site for the protein ankyrin (5). Whether the lamin B receptor functions in ion or solute transport remains to be determined. Besides functioning as membrane transport proteins, polytopic integral membrane proteins have also been shown to function in signal transduction mediated by GTP-binding proteins (11) and as the enzyme catalyzing the rate-limiting reaction of cholesterol biosynthesis (21, 29).

Besides its polytopic nature, the lamin B receptor contains a 204-amino acid residue long amino terminal region that precedes the first potential transmembrane segment. This region is particularly charged and contains two sequence motifs with potentially important implications and functions. The first motif is the consensus site for phosphorylation by the cAMP-dependent protein kinase A (Arg-Arg-X-Ser) (13), and two such sites overlap each other in this domain of the molecule. As the lamin B receptor has previously been shown to be phosphorylated by protein kinase A in vitro and in vivo (3), it is likely that this entire amino terminal domain is exposed on the nucleoplasmic side (rather than the perinuclear cisternal side) of the inner nuclear membrane where it would be accessible to protein kinase A. Furthermore, the protein kinase A catalyzed phosphorylation of the lamin B receptor has been shown to affect lamin B binding (3), also making it more likely that these potential phosphorylation sites are located on the nucleoplasmic side of the inner membrane in the vicinity of lamin B. If the lamin B receptor has an even number of transmembrane domains (such as the eight predicted), the carboxy terminal segment of the protein would also be located on the nucleoplasmic side of the inner nuclear membrane.

The second functional motif found in the 204-amino acid amino terminal domain of the lamin B receptor is the sequence Ser/Thr-Pro-X-X, which occurs three times. This motif has been proposed to be a DNA-binding unit and occurs in histones and more frequently in gene regulatory proteins that bind DNA (39). When the first X is polar and the second X basic, as occurs twice in the lamin B receptor, this motif may also serve as a consensus site for phosphorylation by p34^{cdc2} protein kinase, which is involved in the initiation of mitosis (34). Thus, because of these potential properties as well as its ability to bind to the lamina, the lamin B receptor may function in both the regulation of gene expression and cell division.

We are indebted to Dr. D. Atherton of The Rockefeller University/Howard Hughes Medical Institute Biopolymer Facility (New York) for protein sequencing and oligonucleotide synthesis. We also thank C. Bennett, J. Newman, and C. Rosen for technical assistance; Dr. N. Chaudhary (Rockefeller University) for considerable help and advice; and Drs. D. Johnson, R. Moon, and A. Dannenberg for supplying essential reagents. The help of Drs. J. Aris, J.-C. Courvalin, S. Georgatos, R. Henriquez, M. Schindler, D. Schnell, G. Shelness, J. Sukagawa, and R. Wozniak (Rockefeller University) is also greatly appreciated.

Dr. H. J. Worman is a recipient of a Physician-Scientist Award (grant No. DK01790) from the National Institutes of Health.

Received for publication 30 April 1990 and in revised form 15 June 1990.

References

- 1. Aebi, U., J. Cohn, L. Buhle, and L. Gerace. 1986. The nuclear lamina is a meshwork of intermediate-type filaments. Nature (Lond.). 323:560-564.
- 2. Amar-Costesec, A., H. Beaufay, M. Wibo, D. Thinès-Sempoux, E. Feytmans, M. Robbi, and J. Berthet. 1974. Analytical study of microsomes and isolated subcellular membranes from rat liver. II. Preparation and

composition of the microsomal fraction. J. Cell Biol. 61:201-212.

- Appelbaum, J., G. Blobel, and S. D. Georgatos. 1990. In vivo phosphorylation of the lamin B receptor. Binding of lamin B to its nuclear membrane receptor is affected by phosphorylation. J. Biol. Chem. 265:4181-4184.
- 4. Aviv, H., and P. Leder. 1972. Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. Proc. Natl. Acad. Sci. USA. 69:1408-1421
- 5. Bennett, V., and P. J. Stenbuck. 1980. Association between ankyrin and the cytoplasmic domain of band 3 isolated from the human erythrocyte membrane. J. Biol. Chem. 255:6424-6432.
- 6. Bergmann, J., and J. Singer. 1983. Immunoelectron microscopic studies of the intracellular transport of the membrane glycoprotein (G) of vesicular stomatitis virus in infected Chinese hamster ovary cells. J. Cell Biol. 97:1777-1787
- 7. Blobel, G. 1980. Intracellular protein topogenesis. Proc. Natl. Acad. Sci. USA. 77:1496-1500.
- 8. Cabantchik, Z. I., P. A. Knauf, and A. Rothstein. 1978. The anion transport system of the red blood cell: the role of membrane protein evaluated by the use of "probes." *Biochim. Biophys. Acta.* 515:239-302.
- Chirgwin, J. M., A. E. Przybyla, R. J. MacDonald, and W. J. Rutter. 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry*. 18:5294-5299. 10. Cleveland, D. W., S. G. Fischer, M. W. Kirschner, and U. K. Laemmli.
- 1977. Peptide mapping by limited proteolysis in sodium dodecyl sulfate
- and analysis by gel electrophoresis. J. Biol. Chem. 252:1102-1106. 11. Dohlman, H. G., M. G. Caron, and R. J. Lefkowitz. 1987. A family of receptors coupled to guanine nucleotide regulatory proteins. Biochemis try. 26:2657-2664. 12. Fawcett, D. W. 1966. On the occurrence of a fibrous lamina on the inner
- aspect of the nuclear envelope in certain cells of vertebrates. Am. J. Anat. 119:129-146
- 13. Feramisco, J. R., D. B. Glass, and E. G. Krebs. 1980. Optimal spatial requirements for the location of basic residues in peptide substrates for the cyclic-AMP dependent protein kinase. J. Biol. Chem. 255:4240-4245.
- 14. Fisher, D. Z., N. Chaudhary, and G. Blobel. 1986. cDNA sequencing of nuclear lamins A and C reveals primary and secondary structural homology to intermediate filament proteins. Proc. Natl. Acad. Sci. USA. 83:6450-6454
- 15. Franke, W. W., U. Scheer, G. Krohne, and E. Jarasch. 1981. The nuclear envelope and the architecture of the nuclear periphery. J. Cell Biol. 91:39s-50s.
- 16. Fry, D. J., and G. J. Wishart. 1976. Apparent induction by phenobarbital of uridine diphosphate glucuronyltransferase activity in nuclear envelopes of embryonic chick liver. Biochem. Soc. Trans. 4:265-266.
- 17. Georgatos, S. D., and G. Blobel. 1987. Lamin B constitutes an intermediate filament attachment site at the nuclear envelope. J. Cell Biol. 105:117-125
- 18. Gerace, L., and G. Blobel. 1982. Nuclear lamina and the structural organization of the nuclear envelope. Cold Spring Harbor Symp. Quant. Biol. 46:967-978
- 19. Gerace, L., Y. Ottaviano, and C. Kondor-Koch. 1982. Identification of a major polypeptide of the nuclear pore complex. J. Cell Biol. 95:826-837.
- 20. Girault, J.-A., F. S. Gorelick, and P. Greengard. 1989. Improving the quality of immunoblots by chromatography of polyclonal antisera on keratin affinity columns. Anal. Biochem. 182:193-196.
- 21. Goldstein, J. L., and M. S. Brown. 1990. Regulation of the mevalonate pathway. Nature (Lond.). 343:425-430. 22. Hunkapiller, M. W., R. M. Hewick, W. J. Dreyer, and L. E. Hood. 1983.
- High-sensitivity sequencing with a gas-phase sequenator. Methods Enzymol. 91:339-413
- 23. Hunkapiller, M. W., E. Lujan, F. Ostrander, and L. E. Hood. 1983. Isolation of microgram quantities of proteins from polyacrylamide gels for amino acid sequence analysis. Methods Enzymol. 91:227-236.
- 24. Kopito, R. R., and H. F. Lodish. 1985. Primary structure and transmembrane orientation of the murine anion exchange protein. Nature (Lond.). 316:234-238.
- 25. Krohne, G., S. L. Wolin, F. D. McKeon, W. W. Franke, and M. W. Kirschner. 1987. Nuclear lamin L1 of Xenopus laevis: cDNA cloning, amino acid sequence and binding specificity of a member of the lamin B subfamily. *EMBO (Eur. Mol. Biol. Organ.) J.* 6:3801-3808.
- 26. Kyte, J., and R. F. Doolittle. 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157:105-132.
- 27. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the heads of bacteriophage T4. Nature (Lond.). 227:680-685. 28. Lee, P. L., D. E. Johnson, L. S. Cousens, V. A. Fried, and L. T. Williams.
- 1989. Purification and complementary DNA cloning of a receptor for basic fibroblast growth factor. Science (Wash. DC). 245:57-60.
- 29. Liscum, L., J. Finer-Moore, R. M. Stroud, K. Luskey, M. S. Brown, and J. L. Goldstein. 1985. Domain structure of 3-hydroxy-3-methylglutaryl coenzyme A reductase, a glycoprotein of the endoplasmic reticulum. J. Biol. Chem. 260:522-530.
- Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. 545 pp. 31. Matsudaira, P. 1987. Sequence of picomole quantities of proteins electro-

blotted onto polyvinylidine difluoride membranes. J. Biol. Chem. 262: 10035-10038.

- 32. McKeon, F. D., M. W. Kirschner, and D. Caput. 1986. Homologies in both primary and secondary structure between nuclear envelope and intermediate filament proteins. Nature (Lond.). 319:463-468.
- 33. Moon, R. T., J. Ngai, B. J. Wold, and E. Lazarides. 1985. Tissue-specific expression of distinct spectrin and ankyrin transcripts in erythroid and nonerythroid cells. J. Cell Biol. 100:152-160.
- 34. Moreno, S., and P. Nurse. 1990. Substrates for p34^{cdc2}: in vivo veritas? Cell. 61:549-551.
- 35. Pathak, R. K., K. L. Luskey, and R. G. W. Anderson. 1986. Biogenesis of the crystalloid endoplasmic reticulum in UT-1 cells: evidence that newly formed endoplasmic reticulum emerges from the nuclear envelope. J. Cell Biol. 102:2158-2168.
- 36. Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich. 1987. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science (Wash. DC). 239:487-491.
- 37. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA. 74:5463-5467.
 38. Senior, A., and L. Gerace. 1988. Integral membrane proteins specific to

the inner nuclear membrane and associated with the nuclear lamina. J. Cell Biol. 107:2029-2036.

- 39. Suzuki, M. 1989. SPXX, a frequent sequence motif in gene regulatory proteins. J. Mol. Biol. 207:61-84. 40. Thomas, P. S. 1980. Hybridization of denatured RNA and small DNA frag-
- ments transferred to nitrocellulose. Proc. Natl. Acad. Sci. USA. 77: 5201-5205
- 41. Torrisi, M., and S. Bonnatti. 1985. Immunocytochemical study of the partition and distribution of Sindbis virus glycoproteins in freeze-fractured membranes of infected baby hamster kidney cells. J. Cell Biol. 101: 1300-1306.
- 42. Torrisi, M., L. Lotti, A. Pavan, G. Migliaccio, and S. Bonatti. 1987. Free diffusion to and from the plasma membrane of newly synthesized plasma membrane glycoproteins. J. Cell Biol. 104:733-737.
 43. Worman, H. J., J. Yuan, G. Blobel, and S. D. Georgatos. 1988. A lamin
- B receptor in the nuclear envelope. Proc. Natl. Acad. Sci. USA. 85:8531-8534.
- 44. Wozniak, R. W., E. Bartnik, and G. Blobel. 1989. Primary structure analysis of an integral membrane glycoprotein of the nuclear pore. J. Cell Biol. 108:2083-2092.