INVESTIGATION OF THE REGULATION OF NUCLEAR TRANSLOCATION OF THE TRANSCRIPTION FACTOR MESODERM INDUCTION-EARLY RESPONSE 1 (MI-ER1) DURING EMBRYONIC DEVELOPMENT OF *Xenopus laevis*

ING. JANINE NICOLE POST


CENTRE FOR NFLD. STUDIES
MEMORIAL UNIVERSITY
OF NEWFOUNDLAND

Your file Votre r□f□rence
ISBN: 0-612-84071-9
Our file Notre r□f□rence
ISBN: 0-612-84071-9

Investigation of the regulation of nuclear translocation of the transcription factor *mesoderm induction-early response 1 (mi-er1)* during embryonic development of *Xenopus laevis*

By

Ing. Janine Nicole Post

A dissertation submitted to the School of Graduate Studies in partial fulfillment of the requirement for the degree of Doctor of Philosophy

Department of Basic Medical Sciences

Faculty of Medicine

Memorial University of Newfoundland

April, 2002

Table of Contents

Abstract.....i

Acknowledgements.....ii

List of Figures.....iii

List of Tables.....viii

Abbreviations.....x

Chapter 1 Introduction **1**

1.1 Regulation of Transcription in *Xenopus laevis* 2
- 1.1.1 *Xenopus laevis* embryonic development 3
- 1.1.2 Regulation of the Early Cell Cycle in *Xenopus laevis* 9
- 1.1.3 Transcription of Embryonic Genes Starts at the Mid-Blastula Transition 11
  - 1.1.3.1 Molecular Mechanisms of the MBT 11
  - 1.1.3.2 Regulation of Nuclear import of transcription factors at MBT 12
- 1.1.4 Regulation of nuclear import and activation of transcription factors is dependent on phosphorylation 14

1.2 FGF Signal Transduction 15
- 1.2.1 FGFs 15
- 1.2.2 FGF receptors 18
- 1.2.3 FGF receptor dimerization and activation 18
  - 1.2.3.1 The SHC/FRS2-RAF/MAPkinase pathway 21
  - 1.2.3.2 The PLCγ signaling pathway 25
- 1.2.4 Immediate Early Genes 26
- 1.2.5 Transcription factors 28

1.3 Nuclear transport of macromolecules 30
- 1.3.1 The Nuclear Pore Complex 31
- 1.3.2 Import machinery 35
- 1.3.3 Signals for transport 37
- 1.3.4 Mechanisms of transport 40
- 1.3.5 Regulation of nuclear transport 43
  - 1.3.5.1 Protein context and competing signals 47
  - 1.3.5.2 Piggy backing into the nucleus 47
  - 1.3.5.3 Regulated or conditional nuclear translocation 48

1.3.5.4 Phosphorylation enhancing NLS dependent transport 50
1.3.5.5 Developmentally regulated NLSs 50
1.3.5.6 CcN motif and phosphorylation regulated NLS 51
1.3.5.7 Regulation of the soluble transport machinery 53
1.3.5.8 Regulation of the Nuclear Pore Complex 54
1.4 Mesoderm induction early response 1, *mi-er1* 56
1.4.1 *Xmi-er1* isolation 56
1.4.2 Expression of *Xmi-er*1 mRNA 57
1.4.3 XMI-ER1 protein expression 59
1.4.4 Protein-domains in XMI-ER1 and their possible function 60
1.5 Objectives 71

Chapter 2 Materials & Methods 75
2.1 Generation of constructs 75
2.1.1 Generation of EGFP-ER1 fusion constructs 81
2.1.1.1 PCR 81
2.1.1.2 Ligation of PCR products into PCR 2.1 82
2.1.1.3 Digestion and purification of inserts from PCR2.1 82
2.1.1.4 Generation of small fusion constructs 83
2.1.1.5 Ligation of purified inserts into EGFP-C2 83
2.1.1.6 Plasmid isolation and purification 84
2.1.2 Deletion or substitution of single or multiple amino acids 84
2.1.3 Generation of gfp-er1 fusions into pT7Ts 84
2.1.4 Generation of β-galactosidase-er1 fusion constructs 85
2.1.5 Sequencing 86
2.2 Cell culture 86
2.2.1 Materials 86
2.2.2 Methods 86
2.3 Transient transfection of NIH 3T3 cells 87
2.3.1 Microscopic examination of transfected cells 88
2.3.1.1 Detection of fluorescent proteins 88
2.3.1.2. Immunocytochemistry 89
2.3.1.3 X-gal staining of transfected NIH 3T3 cells 90
2.3.2 Cell lysis for direct Western Blot 91
2.3.3 Immunoprecipitation of proteins 91
2.4 Western Blot 92
2.4.1 Materials 92
2.4.2 Methods 93
2.5 Fertilization of *Xenopus laevis* eggs 94
2.5.1 Materials 94
2.5.2 Methods 94

2.6 Micro-injection of RNA into *Xenopus laevis* embryos 95
2.6.1 Preparation of RNA for injection 95
2.6.2 Injection of RNA into *Xenopus* embryos 96
2.6.3 Staging of embryos according to Nieuwkoop and Faber 97
2.6.4 Detection of proteins in microinjected embryos 97
2.6.4.1 Detection by fluorescence microscopy 97
2.6.4.2 Detection by whole mount staining 98
2.6.4.3 X-gal staining of whole embryos 99

Chapter 3 Identification of nuclear localization signals in XMI-ER1 100
3.1 Introduction 100
3.2 Sequence analysis of basic regions in XMI-ER1 101
3.3 Identification of functional NLS was performed using different fusion proteins 108
3.4 Identification of a functional NLS in the C-terminus of XMI-ER1 117
3.5 NLS4 is necessary and sufficient for the nuclear transport of XMI-ER1 128
3.5.1 NLS4 is able to direct EGFP and β-GAL to the nuclei of NIH 3T3 cells 128
3.5.2 Deletion of the NLS4 results in cytoplasmic localization of XMI-ER1 134
3.6 Additional weak NLSs in XMI-ER1 138
3.7 Deletion of NLS1 and/or NLS3 does not interfere with the nuclear translocation of XMI-ER1 144
3.8 Analysis of the putative NLS for XMI-ER1 in *Xenopus laevis* 149
3.9 Conclusion 158

Chapter 4 Regulation of nuclear transport of MI-ER1 during *Xenopus laevis* development 160
4.1 Introduction 160
4.2 Nuclear Localization of MI-ER1 during *Xenopus* development 161
4.3 β-GAL-NLS4 translocates to the nucleus prematurely 165
4.4 Cytoplasmic retention of MI-ER1 is not dependent on cdc2 phosphorylation of its CcN motif 171
4.5 MI-ER1 contains a cytoplasmic retention domain, which prevents nuclear translocation 187
4.6 MI-ER1 has a retention domain in its N-terminal half 196
4.7 Attempts to identify the minimal region necessary for retention of MI-ER1 during *Xenopus* development 203
4.8 Conclusion 215

Chapter 5 General Discussion 217
5.1 Analysis of NLS in XMI-ER1 218
5.2 Temporal regulation of nuclear import of MI-ER1 in *Xenopus*

embryos 229
5.3 Model of regulation of nuclear import of MI-ER1 during *Xenopus* development 245
5.4 Future work 249

References **250**

Appendix A1 272
Appendix A2 275
Appendix B 279

Abstract

Temporal and spatial regulation of nuclear transport of transcription factors is important for regulating the function of many cellular and developmental processes. Here the regulation of nuclear transport of a *Xenopus* transcription factor, mesoderm induction-early response 1, MI-ER1, was investigated.

It was shown that MI-ER1 is actively transported to the cell nuclei of transfected NIH 3T3 cells as well as in *Xenopus* embryos. The active import of MI-ER1 into the cell nuclei is dependent on the presence of an intact nuclear localization signal (NLS) located in the C-terminus of the protein. The core of this NLS consists of amino acids ^{463}RPIKRQRMD^{471} but we show that addition of amino acids flanking this region results in more efficient import into the cell nuclei. It was also shown that an additional, albeit weak, NLS is present in the N-terminus of MI-ER1.

During embryonic development in *Xenopus laevis* MI-ER1 is localized in the cytoplasm by cytoplasmic retention, but localizes to the nucleus during the mid-blastula stages. Investigation into the regulation of cytoplasmic retention of MI-ER1 in *Xenopus laevis* embryos, showed that the NLS (ER457-475) of MI-ER1 is able to direct β-GAL to the nucleus prematurely, which suggests that cytoplasmic retention of MI-ER1 is not dependent on NLS recognition and binding by the import machinery.

Investigation of a phosphorylation site in the vicinity of this NLS shows, that the regulation of nuclear transport of MI-ER1 in *Xenopus laevis* is not regulated by phosphorylation of this cdc2 kinase/ PKA site. Furthermore, it is shown that cytoplasmic retention of MI-ER1 in embryos appears to be regulated by binding of MI-ER1 to an "anchor" molecule in the cytoplasm. It was shown that the region important for cytoplasmic retention is localized between amino acid residues 1 and 282.

Acknowledgements

I would like to express my deepest feeling of gratitude to my supervisor, Dr. Gary Paterno, for providing me with the opportunity to pursue a dream which I did not think would come true. Throughout the last 5 years Dr. Paterno has helped me jump many obstacles and I admire the patience and perseverance with which he helped me tackle these obstacles. I sincerely appreciate all the time, inspiration and encouragement that Dr. Paterno has given me.

I would also like to express my sincere appreciation for Dr. Laura Gillespie and Dr. Ken Kao, who were always willing to share their expert knowledge, time and encouragement. I appreciate everything Dr. Gillespie and Dr. Kao have done to help and inspire me.

Many thanks go out to Paula Tucker and Yuan Lew for helping me generate some of the many constructs and for friendship and advice. Many thanks also to Artee, Blue, Corinne, Ding, Gord, Kelly, Lori-Ann, Marianne, Mark, Olga, Rebecca and Yoella for the friendship and for providing a great working environment. Also sincere thanks to Dr. Penny Moody-Corbett and staff at the office of Graduate Studies for all their help.

Last but not least I would like to say a special thank you to all my family and friends from home who always believed in me.

And thank you to Rolf, who helped me believe in myself.

List of Figures

Figure 1.1.1 Schematic illustrating events during *Xenopus laevis* early embryonic development 6

Figure 1.2.1 Schematic illustration of a FGFR protein 20

Figure 1.2.2 Schematic illustration of some of the FGF signaling Pathways 24

Figure 1.3.1 Model of the NPC spanning the nuclear envelope 34

Figure 1.3.2 Schematic representation of the classical import pathway 42

Figure 1.3.3 Model of different mechanisms that directly modulate transcription factor localization 46

Figure 1.4.1 Amino acid sequence of *Xenopus* MI-ER1, indicating the motifs found by computer analysis of the sequence 62

Figure 2.1 Amino acid sequence of *Xenopus laevis* MI-ER1 77

Figure 3.1 Identification of potential NLS in XMI-ER1 103

Figure 3.2 Amino acid sequence comparison of XMI-ER1 NLS-4 to the NLS of the c-MYC protein, and the NLS in the human orthologue of XMI-ER1, hMI-ER1β 107

Figure 3.3 Fusion of EGFP to XMI-ER1 does not interfere with the nuclear translocation of XMI-ER1 110

Figure 3.4 Addition of β-galactosidase to the N-terminus of MI-ER1 does not interfere with the nuclear localization 113

Figure 3.5 Definition of subcellular localization of fusion proteins in NIH 3T3 cells 116

Figure 3.6 The nuclear targeting signal of MI-ER1 is located near the C-terminus 119

Figure 3.7 Western Blot analysis of cells expressing different EGFP-ER1 fusion proteins 123

Figure 3.8 NIH 3T3 cells transfected with β-gal-ER339-493 display exclusive nuclear staining, whereas cells transfected with β-GAL-ER1-137 or β-GAL-ER1-143 do not 127

Figure 3.9 NLS4 is necessary and sufficient for the nuclear translocation of MI-ER1 in NIH 3T3 cells 130

Figure 3.10 Western Blot analysis of EGFP-ER1 fusions used to identify the minimal NLS4 region 133

Figure 3.11 The NLS4 basic cluster and flanking regions are necessary for the nuclear translocation of MI-ER1 in NIH 3T3 cells 137

Figure 3.12 NLS1 functions as a weak NLS when fused to EGFP, while NLS2 and NLS3 do not 141

Figure 3.13 Western Blot analysis showing the level of protein expression of EGFP-NLS1, 2 and 3 143

Figure 3.14 Nuclear translocation of MI-ER1 is not dependent on co-operation of NLS4 with NLS1, NLS3 or NLS1 and 3 146

Figure 3.15 Western Blot analysis of EGFP-ER1 fusion proteins after immuno-precipitation using anti-MI-ER1 show similar levels of expression in NIH 3T3 cells 148

Figure 3.16 EGFP-ER339-493, containing NLS4, is translocated to the nuclei of developing *Xenopus laevis* embryos at stage 13 153

Figure 3.17 NLS4 is necessary and sufficient for the nuclear translocation of MI-ER1 in stage 13 *Xenopus* embryos 157

Figure 4.1 Nuclear localization of MI-ER1 during *Xenopus* embryonic development 164

Figure 4.2 Injection of MI-ER457-475 (NLS) into *Xenopus* embryos results in premature nuclear translocation 168

Figure 4.3 Subcellular localization of β-GAL-ER457-475 during different stages of *Xenopus* development 170

Figure 4.4 Amino acid sequence of XMI-ER1 indicating a putative CcN motif 173

Figure 4.5 Substitution of ^{472}serine to alanine (A), asparagine (N) or aspartic acid (D) has no effect on the nuclear localization of EGFP-ER1-493 in NIH 3T3 cells 176

Figure 4.6 Substitution of ^{472}serine to alanine (A), asparagine (N) or aspartic acid (D) has no effect on the nuclear localization of EGFP-ER339-493 in NIH 3T3 cells 178

Figure 4.7 Substitution of ^{472}serine to alanine (A) or aspartic acid (D) has no effect on the nuclear localization of EGFP-ER457-475 in NIH 3T3 cells 180

Figure 4.8 Substitution of ^{472}serine with alanine (A) or aspartic acid (D) in EGFP-ER1-493 does not influence the temporal pattern of nuclear import of MI-ER1 during *Xenopus* embryonic development 186

Figure 4.9 MI-ER1 is retained in the cytoplasm by binding to a cytoplasmic anchor 191

Figure 4.10 Subcellular localization of β-GAL-ER1, β-GAL-XNF7NLS and β-GAL-XNF7NLS-ER1 in *Xenopus laevis* embryos 193

Figure 4.11 Subcellular localization of β-GAL- ER1, β-GAL-XNF7NLS and β-GAL-XNF7NLS-ER1 during different stages of *Xenopus* development 195

Figure 4.12i MI-ER1 possesses a cytoplasmic retention domain in its N-terminal half 198

Figure4.12ii MI-ER1 possesses a cytoplasmic retention domain in its N-terminal half 200

Figure 4.13 Subcellular localization of β-GAL-XNF7NLS-ER1-282 and β-GAL-XNF7NLS-ER267-493 during different stages of *Xenopus* development, compared to β-GAL-XNF7NLS-ER1 and β-GAL-XNF7NLS at these stages.....202

Figure 4.14 Addition of the XNF7NLS to ER1-143, ER145-250, ER251-340 and ER339-493 results in pre-mature nuclear translocation....206

Figure 4.15 Addition of the XNF7NLS to ER1-143, ER145-250, ER251-340 or ER339-493 results in pre-mature nuclear translocation.....208

Figure 4.16 Subcellular localization of XNF7NLS-ER1-143 and XNF7NLS-ER145-250, compared to XNF7NLS-ER1 and XNF7NLS.....210

Figure 4.17 Subcellular localization of XNF7NLS-ER251-340 and XNF7NLS-ER339-493, compared to XNF7NLS-ER1 and XNF7NLS.....212

Figure 4.18 Schematic illustrating different β-GAL fusions used to identify the mechanism of regulation of nuclear transport of MI-ER1 during *Xenopus* embryonic development.....214

Figure 5.1 Schematic illustrating a possible mechanism of regulation of cytoplasmic retention of XMI-ER1.....248

List of Tables

Table 1.4.1 Example of putative phosphorylation sites predicted by pbase_predict and NetPhos 70

Table 2.1 Amino Acid sequences of all constructs used 78

Table 2.2 Antibodies used for ICC, IPT and WB 93

Table 3.1 Subcellular localization of EGFP-XMI-ER1 fusion proteins after transfection into NIH 3T3 cells 120

Table 3.2 Predicted molecular weight of the EGFP-fusion proteins and the probability of diffusion of the EGFP-ER1 fusion proteins through the nuclear pores 121

Table 3.3 Subcellular localization of β-gal fusion constructs in NIH 3T3 cells 124

Table 3.4 Subcellular localization of EGFP-ER1-461 and EGFP-ER1-473 after transfection into NIH 3T3 cells 128

Table 3.5 Subcellular localization of EGFP-ER461-471, EGFP-ER457-475, EGFP-ERΔPIKR, β-GAL-ER461-471, β-GAL-ER457-475 and β-GAL-ΔNLS4 after transfection into NIH 3T3 cells 135

Table 3.6 Percentage nuclear staining in NIH 3T3 cells transfected with EGFP-NLS1, 2, 3 and

EGFP-ΔNLS1 and 3 139

Table 3.7 Subcellular localization different EGFP constructs after injection into *Xenopus* embryos 154

Table 4.1 Temporal pattern of nuclear translocation of XMI-ER1 162

Table 4.2 Transfection of cdc2 kinase/ PKA phosphorylation site mutations in different ER1 constructs into NIH 3T3 cells 181

List of abbreviations

aFGF	acidic fibroblast growth factor, FGF-1
AP1	activator protein 1
ATP	adenosine triphosphate
bp	base pair
bFGF	basic fibroblast growth factor, FGF-2
BMP-4	bone morphogenetic protein 4
C	cytoplasmic
CcN	Casein kinase II, cdc2 kinase, NLS
cDNA	complementary deoxyribonucleic acid
CKII	casein kinase II
CRD	cytoplasmic retention domain
cRNA	complementary ribonucleic acid
C-terminus	carboxy-terminus
DAG	diacylglycerol
DBD	DNA binding domain
DEPC	di-ethyl-pyro-carbonate
DNA	deoxyribonucleic acid
drNLS	developmentally regulated nuclear localization signal
eFGF	embryonic fibroblast growth factor
EGFP	enhanced green fluorescent protein
ELM2	EGL-27 and MTA like motif
ER1	early response 1
FGF	fibroblast growth factor
FGFR	fibroblast growth factor receptor
FHF	FGF homologous factor
FNK	FGF inducible kinase
FRS-2	FGFR substrate 2
GDP	guanosine diphosphate
GTP	guanosine triphosphate
HAT	histone acetylase
HDAC	histone deacetylase
hnRNP	heterogeneous nuclear ribonucleoprotein
HLGAG	heparin-like glycosaminoglycans
Hu*mi-er1*α	human *mi-er1* α
IBB	importin beta binding site
Ig	immunoglobin
IP_3	phosphatidylinositol
kDa	kilo Dalton
MAPK	mitogen activated protein kinase

MBT	mid-blastula transition
MCS	multiple cloning site
Min.	minutes
MI-ER1	mesoderm-induction early response-1, protein
mi-er1	mesoderm-induction early response-1, mRNA
mRNA	messenger ribonucleic acid
N	nuclear
N>C	nuclear greater than cytoplasmic
N=C	nuclear equal to cytoplasmic
N/C	nuclear cytoplasmic ratio
NE	nuclear envelope
NES	nuclear export signal
ng	nanograms
NLS	nuclear localization signal
nm	nanometer
NPC	nuclear pore complex
NRD	nuclear retention domain
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PI3'K	phosphoinositide 3' kinase
PIP_2	phosphatidyl-inositol-4,5-biphosphate
PKA	protein kinase A
PKC	protein kinase C
PLCγ	phospholipase gamma
prNLS	phosphorylation regulated NLS
RNA	ribonucleic acid
rRNA	ribosomal RNA
SH2	Src homology 2
SH3	Src homology 3
Sos	son of sevenless
SR-proteins	serine-argenine rich proteins
tRNA	transfer ribonucleic acid
µg/ ml	microgram per milliliter
Xbra	*Xenopus* Brachyury
XFD	dominant-negative *Xenopus* FGF receptor
XMI-ER1	*Xenopus* MI-ER1
Xvent	*Xenopus* ventral inducing protein

Chapter 1 Introduction

Transcription factors are essential for the initiation of RNA synthesis. Initiation of transcription occurs when transcription factors bind to the promoter and recruit other proteins of the transcription machinery, including RNA polymerases. Temporal and spatial regulation of nuclear transport of transcription factors is important for regulating the function of many cellular and developmental processes (Nothias *et al.*, 1995). During *Xenopus laevis* embryonic development, the genome is transcriptionally silent until a specific stage in development, the mid-blastula transition, around the 12^{th} cleavage (Newport and Kirschner, 1982b; Shiokawa *et al.*, 1994). At this time transcription factors are imported into the nucleus and transcription starts. Regulating the timing of nuclear import of transcription factors is crucial for the development of many organisms (Nothias *et al.*, 1995).

Regulation of import can be modulated by phosphorylation of the transported protein, masking of the nuclear localization signal(s), cytoplasmic retention by binding to a cytoplasmic anchoring protein, modulation of the import machinery, and possibly interplay between these mechanisms (Dingwall and Laskey, 1986; Dingwall and Laskey, 1998; Gorlich and Kutay, 1999; Jans, 1995; Jans and Hubner, 1996; Vandromme *et al.*, 1996).

Here, I investigate the regulation of nuclear transport of a *Xenopus* transcription factor, mesoderm induction-early response 1, MI-ER1, an

immediate early gene activated by FGF signaling (Paterno *et al.*, 1997). The investigation of regulation of nuclear import of MI-ER1 in NIH 3T3 cells as well as in *Xenopus laevis* embryos will be described in this thesis.

1.1 Regulation of Transcription in *Xenopus laevis*

Early stages of *Xenopus* development rely completely on maternal stores of RNA and protein, which are produced during oogenesis (Smith *et al.*, 1991). Transcription does not start until the so-called mid-blastula transition, about 7 hours after fertilization and when there are approximately 4000 cells (Newport and Kirschner, 1982a). The early development of vertebrate and many invertebrate species is characterized by a period during which the embryonic genome is transcriptionally silent (Almouzni and Wolffe, 1995; Newport and Kirschner, 1982a; Veenstra *et al.*, 1999). During this period, many proteins are expressed in the embryo. This includes proteins that are stored during oogenesis as well as proteins translated from mRNAs that are stored during oogenesis. The developmental stage at which the major transcriptional activity starts ranges from the 2–cell stage in the mouse to the 9-14 cycle in *Drosophila*. In nematodes and amphibians the major activation of the embryonic genome starts after 7 cycles (in *C. elegans*) and 12 cycles, respectively (Nothias *et al.*, 1995; Veenstra *et al.*, 1999; Yasuda and Schubiger, 1992).

1.1.1 *Xenopus laevis* embryonic development

The *Xenopus* egg is 1.2-1.4 mm in diameter and consists of a darkly pigmented animal hemisphere and a lighter yolky 'vegetal hemisphere'. When laid, the eggs are oriented randomly with respect to gravity and held in position by a transparent vitelline membrane inside a jelly coat (Jones and Smith, 1999). After fertilization, granules located just below the surface of the egg fuse with the plasma membrane and release their contents into the space between the vitelline membrane and the egg, allowing the egg to rotate so that the less dense animal hemisphere is on top. This rotation normally occurs within 20 min. after fertilization (Jones and Smith, 1999). During fertilization, the sperm enters the egg in the animal hemisphere. The position of sperm entry defines the future dorsal-ventral axis of the embryo (Jones and Smith, 1999).

The earliest developmental stage of *Xenopus* embryogenesis consists of a period of rapid cleavage divisions. These divisions are nearly synchronous for all cells within the embryo and occur approximately every 35 minutes. This rapid cleavage program ends abruptly after about 12 cleavages and is followed by the activation of several new cell activities, called the mid-blastula transition (Newport and Kirschner, 1982a). These include the desynchronization of the embryonic cell cycles, elongation of the cell cycle as a result of the inclusion of G1 and G2 phases and elongation of both S and M phases; the onset of cell motility and the activation of RNA transcription in all the cells of the embryo (Newport and Kirschner, 1982a).

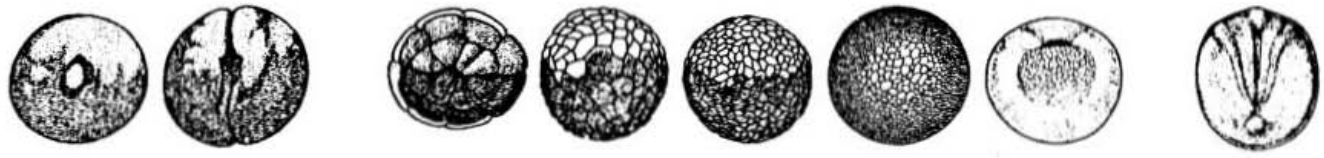
The first *Xenopus* cell cycle occupies approximately 90 min at 21 °C; subsequent cell cycles last about 30 min. (figure 1.1.1). The first cleavage in *Xenopus* initiates at the animal pole and separates the future left and right-hand sides of the embryo. The second cleavage, 30 min. later, occurs at right angles to the first and separates the future dorsal and ventral halves. The third cleavage is orthogonal and separates the animal and vegetal poles. In contrast to the first two divisions, which divide the cells into roughly equal pieces, the plane of the third cleavage is well above the equator of the embryo. After the fourth and fifth cleavages the embryo comprises 32 cells arranged in four tiers of eight cells, which corresponds to stage 6 according to Nieuwkoop and Faber (figure 1.1.1; (Nieuwkoop and Faber, 1967).

During the early cleavages a small space forms between the blastomeres that becomes larger as cleavages proceed and eventually becomes the blastocoel. At this time the embryo is referred to as blastula (Jones and Smith, 1999), a hollow ball of cells. At the twelfth cell division, the cleavage rate slows down and cell division becomes asynchronous. This point, the mid-blastula transition (MBT) is also marked by the onset of cell motility and zygotic transcription (Masui and Wang, 1998; Newport and Kirschner, 1982b).

Figure 1.1.1 *Xenopus laevis* early embryonic development is characterized by many events.

A schematic illustrating *Xenopus* embryos during different stages of development, as described by Nieuwkoop and Faber (1967). The animal pole is shown for embryos at stages 1, 2, 6, 7, 8 and 13; the vegetal pole is shown for stages 9 and 10 since identification of these stages is easier determined from the vegetal pole.

Different developmental milestones are indicated. Arrow indicates that MBT is an event that cannot be determined by morphology and takes place during stage 8-8.5.



Stages	1	2	6	7	8	9	10	13
Time after Fertilization (hrs)	0	1.5	3	4	5	7	9	15
Developmental Milestones			morula	large cell blastula	MBT →	fine cell blastula	gastrula	slit-blastopore

The next important step in *Xenopus* development after MBT is gastrulation. During gastrulation, the blastula is converted into a three-layered structure with a central midline and bilateral symmetry. The first sign of gastrulation is the appearance of a pigmented depression in the dorsal-vegetal quadrant of the embryo, the dorsal lip of the blastopore. This is the first sign of a series of migration of cells resulting in the formation of three germ layers, ectoderm, mesoderm and endoderm. The ectodermal layer will later form epidermis, nervous system and sensory organs. The mesodermal layer develops into muscle, connective tissue, skeletal tissue and the circulatory and excretory systems. The endodermal layer eventually differentiates into the gut and associated organs, such as the liver and the pancreas (Jones and Smith, 1999).

The next vital event in development is neurulation at stage 16. Early during neurulation, the posterior neural plate undergoes convergence and extension, assisting in blastopore closure. The first sign of neurulation in *Xenopus* is the thickening of the inner layer of dorsal ectoderm. This is followed by the formation of a dark pigment line along the dorsal midline of the embryo, the neural groove. As neurulation proceeds the neural groove deepens and the lateral neural folds converge on the dorsal midline, where they eventually fuse to form the neural tube (Jones and Smith, 1999).

During and after neurulation the mesoderm becomes subdivided into different tissues along the dorsal-ventral axis. The most dorsal cell type is the

notochord. Lateral to the notochord are the cells of the somites, which, in *Xenopus,* predominantly form muscle. Lateral and ventral to the somites are cells of the pronephros. The lateral mesoderm goes on to form structures such as the limbs, and the most ventral mesoderm forms blood (Jones and Smith, 1999).

Organogenesis is the process by which organs are formed. The development of most organs begins with aggregates of precursor cells, which then migrate in response to orientation and positional cues. Once in their defined positions, extensive variations in cell adhesion, proliferation and apoptosis occurs which helps to shape the organ. Organogenesis in *Xenopus* has not been extensively studied, but there are some recent reports that focus on the molecular aspects of organogenesis (for reviews see (Capdevila and Izpisua Belmonte, 2000; Ramsdell and Yost, 1998). Organogenesis leads to the formation of a free-swimming tadpole (stage 45) at four days after fertilization (Wolpert *et al.*, 1998). The final event in the development of the adult *Xenopus laevis* is metamorphosis. This is a post-embryonic process that involves the transformation of a free-swimming tadpole into an adult organism. Indeed, during this transformation many structural changes occur, but the most dramatic and obvious are the growth and differentiation of limbs and the death and resorption of the tadpole's tail. Metamorphosis in *Xenopus* is triggered by environmental cues such as nutrition, light and temperature and is controlled by Thyroid hormone (for review see (Tata, 1996; Tata, 1999). Metamorphosis to form an adult frog occurs around 60 days after fertilization (Wolpert *et al.*, 1998).

1.1.2 Regulation of the Early Cell Cycle in *Xenopus laevis*

Adult cells do not enter mitosis until they have grown to double their masses, but also do not grow more than double without dividing (Mitchison and Creanor, 1971). Thus the cell maintains a constant ratio of the nucleus to the cytoplasm, the nucleocytoplasmic (N/C) ratio characteristic of the cell type. The exceptions to this rule are animal oocytes and zygotes. The oocyte continuously grows during oogenesis to increase the cell mass enormously without cell division. Therefore, the N/C ratio of the zygote at fertilization is very low. Conversely, the zygote divides without cell growth during the cleavage period to diminish cell volume and increase the N/C ratio exponentially as cleavage progresses (Masui and Wang, 1998).

Cell cycle durations of blastomeres remain constant and synchronous up to the blastula stage, and then rapidly and variably lengthen (Graham and Morgan, 1966). In the *Xenopus* embryo as a whole, cleavage synchrony ends at the 9^{th} cycle, when blastomeres in the vegetal pole begin asynchronous cleavage, but the blastomeres near the animal pole cleave synchronously up to the 12^{th} cell cycle and then cleavage becomes asynchronous (Masui and Wang, 1998).

In order to maintain genetic integrity throughout a cell line, it is necessary for proliferating cells to have mechanisms that ensure accurate DNA replication and exact delivery of chromosomes to daughter cells. One of these mechanisms

is referred to as 'checkpoint' control (Hartwell and Weinert, 1989). It operates by arresting the cell cycle at S phase until DNA replication is complete, or at M phase until chromosomes are correctly arranged on the equatorial plate of the spindle. Therefore, most cells are arrested in S phase if DNA synthesis is inhibited or in G_2 phase if DNA is damaged, or in M phase if the mitotic apparatus is disassembled by microtubule inhibitors (Masui and Wang, 1998). In *Xenopus* embryos, experiments have shown that the critical N/C ratio is responsible for the development of the S phase checkpoint mechanism and not the chronological age (Kimelman *et al.*, 1987). However, the development of the M phase checkpoint is dependent on the chronological age of the cell (reviewed by (Masui and Wang, 1998).

In the early development of *Xenopus laevis*, fundamental changes in cell cycle regulation occur during MBT or from the 12^{th} to 13^{th} cell cycle (Newport and Kirschner, 1982b). At this stage, unlike at pre-MBT stages, cell cycles become highly dependent on cell surface activities of blastomeres. The cleavage of blastomeres also changes from synchronous to asynchronous division, G_1 and G_2 phases first appear in the cell cycle and cell cycle durations suddenly increase. As well, blastomeres develop checkpoints at S and M phases. MBT is also the time when zygotic transcription is initiated (Newport and Kirschner, 1982b).

1.1.3 Transcription of Embryonic Genes Starts at the Mid-Blastula Transition

MBT is defined by an interval, during which completion of the rapid cleavage period is followed by the immediate activation of a new developmental program, which includes start of zygotic transcription (Meehan and Stancheva, 2001; Newport and Kirschner, 1982a; Newport and Kirschner, 1982b; Stancheva *et al.*, 2002). This indicates that MBT cannot be observed visually, although it is known to occur around cleavage 12, during stage 8.

1.1.3.1 Signaling Mechanisms of the MBT

Once an egg is fertilized, intracellular Ca^{2+} is elevated, cyclin B degradation is stimulated and cytostatic factor, CSF, activity declines in concert with loss of MAPK and RSK activity (Maller *et al.*, 2001). The first 12 cleavage cycles display only M and S phase, except for the first cycle, which is longer and has a detectable G_2 phase (Masui and Wang, 1998). Despite these short cycles, the enzymes that control cdc2-tyrosine15 phosphorylation, wee1 and cdc25A, are rapidly synthesized after fertilization and new cyclin B/cdc2 complexes appear to undergo a rapid cycle of tyrosine15 phosphorylation/ dephosphorylation in every cell cycle. Since no G_2 phase is evident in these cells, the function of tyrosine-15 phosphorylation is not clear at this time (Maller *et al.*, 2001). In contrast, Hartley *et al.* have shown that tyrosine phosphorylation of cdc2 is absent until loss of Cyclin E at the MBT (Hartley *et al.*, 1996). The

events at MBT fall into two classes: those that occur at a fixed time after fertilization and those that require a threshold N/C ratio (Masui and Wang, 1998). For example, the degradation of maternal cyclin E happens at a specific time after fertilization, while the start of zygotic transcription is dependent on the N/C ratio (Hartley *et al.*, 1997; Newport and Kirschner, 1982b).

The introduction of the DNA replication checkpoint at MBT is dependent on the N/C ratio. The block to mitotic entry might be a result of failure to dephosphorylate and activate cyclin B/cdc2. It is implied that cdc25 activation is inhibited by phosphorylation on serine287 which results in the creation of a docking site for 14-3-3 proteins (Kumagai *et al.*, 1998; Peng *et al.*, 1997). The 14-3-3 binding is believed to block nuclear translocation and/ or interaction with cyclin B/cdc2, thus preventing enzyme/substrate interaction. A role for the Plx/cdc25C pathway in the DNA replication checkpoint has not yet been established (Maller *et al.*, 2001).

1.1.3.2 Regulation of Nuclear import of transcription factors at MBT

At the time of zygotic gene activation, transcription is under control of maternally inherited factors (Newport and Kirschner, 1982b). In order to prevent aberrant expression of genes normally transcribed at or soon after the start of transcription, the activity of these maternal factors must be tightly regulated. Although a repressive chromatin structure and a repressed basal transcription machinery have been implicated in repressing transcription before MBT

(Almouzni and Wolffe, 1993a; Almouzni and Wolffe, 1993b; Almouzni and Wolffe, 1995), it is not clear how maternally derived transcription factors are prevented from activating transcription despite the high levels present in the early embryo. These transcription factors and the RNAs encoding these factors are produced and stored in the oocyte. During early development, the proteins present in the nucleus segregate over an increased number of nuclei (Veenstra *et al.*, 1999). Therefore, the total amount of maternal proteins present, are divided over many nuclei, resulting in a low concentration of protein per nucleus. Transcription factors, which are stored in the oocyte, would reach high local concentrations if translocated to the small number of nuclei in pre-MBT embryos. These high local concentrations might counteract transcriptional repression. Therefore, cytoplasmic sequestration of maternally inherited transcription factors may provide a mechanism by which the transcriptional machinery is restrained (Veenstra *et al.*, 1999). Examples of this regulation are provided by the CCAAT box transcription factor (CBTF), which is retained in the cytoplasm by binding to maternal RNA transcripts. At MBT, these maternal RNA transcripts undergo degradation and this results in the release of CBTF and subsequent nuclear translocation (Brzostowski *et al.*, 2000). Another example of a protein translocated at MBT is Xnf7, which is hyperphosphorylated before MBT, which results in cytoplasmic retention. Dephosphorylation at MBT causes release from a cytoplasmic anchoring protein and subsequent import into the nucleus (Li *et al.*, 1994a; Shou *et al.*, 1996).

1.1.4 Regulation of nuclear import and activation of transcription factors is dependent upon phosphorylation

Nuclear transport of transcription factors as well as the activation of transcription factors can be regulated by phosphorylation. Proteins responsible for activating other proteins by phosphorylation and dephosphorylation are called kinases and phosphatases, respectively. Changes in phosphorylation state of a protein reflect changes in the relative balance between the protein kinase(s) and protein phosphatase(s) acting on the protein (Maller *et al.*, 2001). The balance between the action of kinases and phosphatases can be regulated by signal transduction.

Signal transduction is the path that a signal can travel from the surface of the cell to the cell nucleus where the message gets translated into action. A receptor on the cell surface can bind a ligand and can subsequently activate intracellular proteins. This results in a cascade of proteins being activated until the signal reaches the cell nucleus, where subsequent activation or inhibition of transcription can occur. Activation of proteins occurs by phosphorylation and/or dephosphorylation of specific sites in the protein. The cascade of events triggered after binding of a ligand to a cell surface receptor is called a signaling pathway. There are many signaling pathways in a cell (reviewed in (Fiorini *et al.*, 2001; Powers *et al.*, 2000).

The transcription factor MI-ER1 was isolated as an immediate early gene following activation of a fibroblast growth factor (FGF) signal transduction pathway (Paterno *et al.*, 1997). The FGF signaling pathways are well-investigated pathways, found to be very diverse. FGFs and the FGF signaling pathways appear to play significant roles in cell survival, apoptosis, proliferation, differentiation, matrix composition, chemotaxis, cell adhesion, migration and growth of cell processes. Different cell types may display alternate, sometimes even opposite responses to FGFs, depending on the state of differentiation, biochemical status and the cellular, physical and chemical environment of the cell (reviewed in (Powers *et al.*, 2000; Szebenyi and Fallon, 1999). Clearly, misregulation of FGF pathways can lead to disastrous events like neoplasia.

The following section will describe FGF in more detail as well as some well-known FGF pathways, the Ras/Raf/MAPK pathway and the PLC-γ pathway.

1.2 FGF Signal Transduction

1.2.1 FGFs

FGFs were first isolated in the 1970s from bovine brain extracts based on their mitogenic activities. Subsequent research established that FGFs form a family of structurally related polypeptide growth factors. FGFs appear to play significant roles not only in normal development and wound healing, but also in

tumor development and progression (reviewed in (Szebenyi and Fallon, 1999) and (Powers *et al.*, 2000).

Defining features of the FGF family are a strong affinity for heparin and heparin-like glycosaminoglycans (HLGAGs) (Burgess and Maciag, 1989) as well as a central core of 140 amino acid residues that is highly homologous between different family members. Structure and not specificity of growth promoting activity is the defining feature of the FGF family, although the nomenclature of the first isolated FGFs was based on their biological activity (Baird and Klagsbrun, 1991).

Twenty-two distinct FGFs have been discovered to date, numbered consecutively from 1 to 22, which have a molecular weight between 17 and 34 kDa (reviewed in (Ornitz and Itoh, 2001). FGF-11 to 22 are more correctly called FGF homologous factors, because they share less than 30% identity to other FGFs, whereas they show between 50 and 80% homology to each other. They may therefore constitute a different but related family (Schoorlemmer and Goldfarb, 2001; Smallwood *et al.*, 1996). FGFs induce mitogenic, chemotactic and angiogenic activity in cells of mesodermal and neuroectodermal origin (Ornitz and Itoh, 2001). FGF-1 and –2 were the first FGFs to be isolated and are the prototypes of the FGF family.

Both FGF-1 and FGF-2 were initially isolated from bovine pituitary extracts based on their stimulation of [^{3}H]thymidine incorporation in 3T3 fibroblasts (Armelin, 1973; Gospodarowicz, 1974). In humans, FGF-1 is a 155 amino acid

protein and the *fgf-1* open reading frame is flanked by stop codons resulting in only one protein form (Jaye *et al.*, 1986). FGF-1 does not have a signal peptide for channeling through the classical secretory pathway (Jaye *et al.*, 1986), however it does possess a nuclear localization signal (Imamura *et al.*, 1990) and has been found associated with the nucleus (Sano *et al.*, 1990; Speir *et al.*, 1991). The presence of this nuclear localization signal appears to be important for FGF-1 induced mitogenesis and removal has been shown to abrogate FGF-1's mitogenic effect (Imamura *et al.*, 1990). FGF-1 has been shown to stimulate DNA synthesis without signaling through a cell surface receptor (Wiedlocha *et al.*, 1994), suggesting that the nuclear localization signal allows FGF-1 to act through an intracrine mechanism.

The 18 kilo-Dalton (kDa) form of FGF-2 has a 55% sequence homology with FGF-1 (Bohlen *et al.*, 1985; Gimenez-Gallego *et al.*, 1985). Four different FGF-2 polypeptides can be translated from the one *fgf*-2 gene by use of upstream CUG start codons: in addition to the 18-kDa form, 22.5, 23.1 and 24.2-kDa forms have been identified (Florkiewicz and Sommer, 1989).

Like FGF-1, FGF-2 does not contain a signal for secretion, but also contains a nuclear localization signal in the N-terminus of the larger isoforms (Bugler *et al.*, 1991). The role of nuclear localization in the activity of FGF-2 remains unclear.

1.2.2 FGF receptors

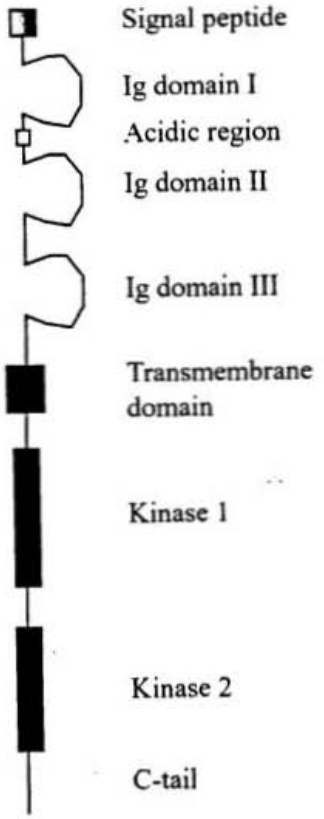
FGFs usually produce their mitogenic and angiogenic effects in target cells by signaling through transmembrane tyrosine kinase receptors. There are four FGF receptors, designated FGFR1, 2, 3 and 4 which share between 55% and 72% identity at the protein level (Johnson and Williams, 1993). There are different isoforms for each of the four FGFRs, generated by alternative splicing of the mRNA (Johnson and Williams, 1993). FGFs can bind to a distinct subset of FGFRs, creating the possibility for distinct intracellular signaling. Figure 1.2.1 illustrates the sequence motifs found in FGFRs. Note that the immunoglobin (Ig) domain I is not found in all FGFR isoforms (for review, see (Powers *et al.*, 2000)). The highest sequence similarity between the four FGFRs is found in the Ig domain III, the Kinase I and the Kinase II regions.

1.2.3 FGF receptor dimerization and activation

Following ligand binding and dimerization, the receptors become capable of phosphorylating specific tyrosine residues on their own and each others cytoplasmic domains (Lemmon and Schlessinger, 1994). The FGFRs can transphosphorylate in homo- and heterodimers (Bellot *et al.*, 1991), allowing for additional complexity in FGF signaling. Phosphorylated tyrosine residues recruit other signaling molecules to the activated receptors and propagate a signal through many possible transduction pathways (Pawson, 1995). The key step from the extracellular to the intracellular pathways is receptor dimerization.

Figure 1.2.1 Schematic illustration of a FGFR protein.

(Modified from (Powers *et al.*, 2000).


Signal peptide
Ig domain I
Acidic region
Ig domain II
Ig domain III
Transmembrane domain
Kinase 1
Kinase 2
C-tail

The activated tyrosine kinase receptor recruits target proteins of the signaling cascade to its cytoplasmic tail and modifies them by phosphorylation. One way these recruited target proteins may be localized to the activated receptor is through interaction between their Src-homology 2 (SH2) domains and specific phosphotyrosine residues on the activated receptor (Pawson, 1995). These SH2 containing proteins may be substrates for receptor mediated phosphorylation or they may function as adaptor proteins to recruit other target proteins (Powers *et al.*, 2000). FGFRs share signaling pathways with other tyrosine kinases. The biological outcome of FGF stimulation depends on the exact concentration and combinations of ligand, receptors and signaling intermediates present in a particular cell. There is accumulating evidence that effects on different cellular functions are mediated by distinct combinations of signaling pathways, some of which may involve nuclear forms of FGFs and FGFRs (Szebenyi and Fallon, 1999). Two extensively studied pathways are the Ras/Raf/MAPK pathway and the PLCγ pathway, which are widely used in a variety of cells.

1.2.3.1 The SHC/FRS2-RAF/MAPkinase pathway

This pathway was shown to be important for growth factor induced cell cycle progression (Szebenyi and Fallon, 1999). Independent of activation of the receptors, an SH2 (Src homology 2) protein, SNT-1 (Wang *et al.*, 1996), better

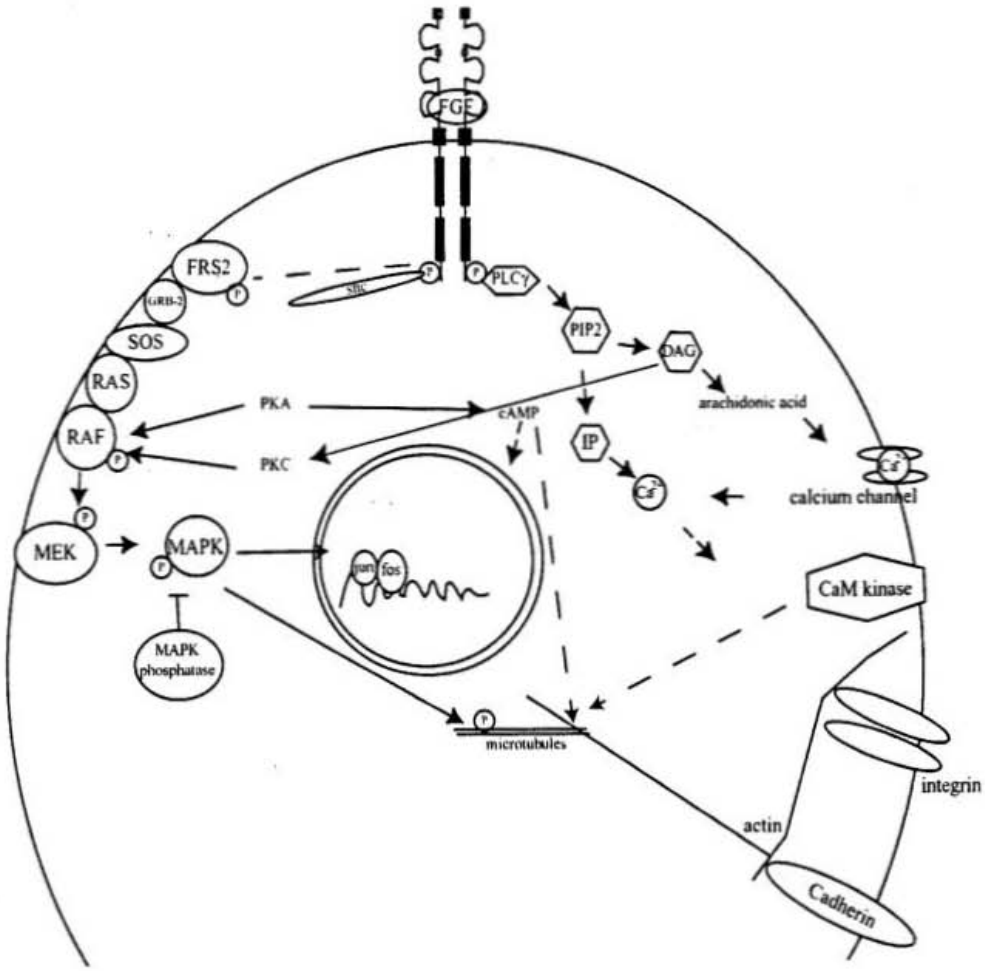
known under the name FRS2 (FGFR substrate 2) (Kouhara *et al.*, 1997), is recruited to the receptor (Ong *et al.*, 2000). Following ligand binding and dimerization of the receptor, FRS2 will be activated. Activation of SNT-1/FRS2 recruits the adaptor protein Grb-2/Sos. This recruits Ras to the FGFR complex, where Ras gets activated (Kouhara *et al.*, 1997). Then membrane associated Ras activates Raf-1, a serine-threonine MAP kinase kinase kinase (MAPKKK), an activator of MAPKK (MEK). In turn, MAPKK activates MAPK (also known as ERK1 (=p44) and ERK2 (=p42)). MAPK can signal directly to the nucleus by phosphorylating transcription factors such as JUN, FOS and the ribosomal S6 kinase, which results in activation of different responses, like proliferation, differentiation, migration, survival, etc.

In addition, each of the kinases in the MAPK pathway can be regulated by multiple signals; for example, PKC and PKA are both regulators of Raf-1 kinase. This results in regulation at different levels of the pathway to prevent misregulation of cellular events (Szebenyi and Fallon, 1999). MAPKs act in concert with other cell signaling systems. Therefore, cross-talk between pathways is crucial to the coordinated responses of cells. MAPKs may act antagonistically in cells undergoing apoptosis or in supporting cell proliferation (Robinson and Cobb, 1997); (Powers *et al.*, 2000; Szebenyi and Fallon, 1999).

Figure 1.2.2 Schematic illustration of some of the FGF signaling pathways

(modified from (Szebenyi and Fallon, 1999).

The diagram illustrates the cytoplasm and the nucleus. Some of the well-known FGF signaling pathways are indicated: the Ras/ Raf/ MAPK pathway and the PLCγ pathway. Molecules indicated are described in the text. After FGF binding to the FGF receptors, the receptors dimerize and are activated by phosphorylation. Solid arrows indicate direct activation of proteins. Dashed arrows indicate activation of which is not known whether it is direct or indirect *via* activation of other proteins.


FGF
FRS2
GRB-2
SOS
RAS
RAF
MEK
MAPK
MAPK phosphatase
SHC
PLCγ
PIP2
DAG
arachidonic acid
PKA
PKC
cAMP
IP
calcium channel
CaM kinase
jun
fos
microtubules
integrin
actin
Cadherin

1.2.3.2 The PLCγ signaling pathway

This pathway is implicated in responses like neurite extension, and modulating the cytoskeletal dynamics of neurons (reviewed in (Szebenyi and Fallon, 1999). Activated PLCγ hydrolyses phosphatidyl-inositol-4,5-biphosphate (PIP2) to inositol-1,4,5-trisphosphate (IP_3) and diacylglycerol (DAG). In turn, IP_3 facilitates the release of calcium stores from the endoplasmic reticulum while DAG and calcium activate PKC, a serine/threonine specific kinase (Merle *et al.*, 1995; Shitaka *et al.*, 1996). PLCγ was identified as a 150-kDa phosphoprotein associated with FGFR following ligand dependent receptor activation (Burgess and Maciag, 1989) and this association is due to the binding between the SH2 domain of PLCγ and tyrosine^{766} of FGFR-1 (Mohammadi *et al.*, 1991). Mutation of tyrosine^{766} did not affect FGFR-mediated mitogenesis, neuronal differentiation, or mesoderm induction in *Xenopus* animal cap assays, indicating that this pathway is not required for these regulatory mechanisms (reviewed in (Johnson and Williams, 1993; Powers *et al.*, 2000).

FGF is involved in many more signal transduction pathways, some of which involve cross-talking of signals between pathways activated by different receptors (Fambrough *et al.*, 1999; Fanger, 1999; Ferrell, 1998; Karin, 1994; Klint and Claesson-Welsh, 1999; Pawson and Saxton, 1999; Powers *et al.*, 2000; Szebenyi and Fallon, 1999). However, in all signal transduction pathways, the first genes to be activated are called immediate early genes. Therefore, these

genes are considered to be master regulators. To date, very few FGF immediate early genes have been isolated and described.

1.2.4 Immediate Early Genes

Immediate early genes (or early response genes) are the first genes to be transcribed differentially in response to a stimulus, for example in response to growth factor treatment of a cellular population. Transcription of immediate early genes is not dependent on *de novo* protein synthesis and is normally rapid, within thirty minutes of binding of the ligand to the receptor. The majority of early response genes are transcription factors; other immediate early genes include receptor molecules, transmembrane proteins or secretory molecules. The transcription factors induced as immediate early genes in turn initiate, or terminate, transcription or activation of new genes and proteins, which carry out the desired cellular response. For example, the *Xenopus* homeobox transcription factor, *XVent1* and *XVent2* are BMP-4 (Bone morphogenetic protein) induced immediate early genes important for ventralization of *Xenopus laevis* embryos (Henningfeld *et al.*, 2002). The *XVent* transcription factors have been identified as pivotal mediators of BMP function in Xenopus. Consistent with the *XVents* as downstream effectors of BMP, these genes are expressed in the ventral and lateral marginal zones and are excluded from the dorsal most region of early *Xenopus* embryos (Henningfeld *et al.*, 2002). Ectopic expression of the *XVents* in the dorsal blastomeres mimics the BMP induced ventralization of *Xenopus*

embryos and results in activation of BMP induced genes (Henningfeld *et al.*, 2002).

Understanding gene regulation via the signal transduction cascades is important for elucidating the mechanisms whereby FGF elicits different cellular actions in different cells and how misregulation can lead to cancer.

C-jun is an immediate early gene involved in mesoderm induction in *Xenopus laevis* by FGF activation. C-jun is an oncogene and transcription factor, which forms homodimers and can heterodimerize with c-fos. The dimer arrangement composes activator protein-1 (AP1) (Dong *et al.*, 1996).

The *Xenopus* homologue of the Brachyury (Xbra) gene is a DNA binding protein upregulated by FGF. Xbra is able to induce mesoderm formation; ventral mesoderm at low concentrations and intermediate mesoderm at higher concentrations (Cunliffe and Smith, 1994). Xbra and embryonic FGF (eFGF) are able to activate transcription of each other and are thought to act in an autocatalytic manner during mesoderm induction (Isaacs *et al.*, 1994).

FGF inducible kinase, *fnk*, is another early response gene upregulated by FGF, More specifically it was shown that induction of *fnk* in NIH 3T3 cells occurred after activation by FGF-1 and FGF-2 (Donohue *et al.*, 1995).

Another immediate early gene, cloned in Gillespie's/ Paterno's laboratory, that is upregulated in response to FGF treatment of *Xenopus* is *mi-er1*, for mesoderm induction early response 1, previously called *er1* (Paterno *et al.*, 1997). MI-ER1 is a transcription factor, whose nuclear translocation is

developmentally regulated in *Xenopus laevis* development. Its human homologue is shown to be present in many cancer cells but not in normal cells of the same tissue. It is likely that regulation of nuclear localization of MI-ER1 is important for development of *Xenopus* embryos.

1.2.5 Transcription factors

Transcription factors are proteins that allow the cell to respond to environmental cues (reviewed in (Marks, 1996). There are many parameters that can influence transcription factor activity, like developmental stage, tissue type, health and nutritional state, stress and many more. Misregulation of transcription factors is most profound in cell transformation and oncogenesis.

Almost all eukaryotic transcription factors are regulated by phosphorylation (Karin, 1994). There are several features that make protein phosphorylation an attractive device for mediating signal dependent regulation: I) It is rapid, mediating responses within minutes. II) The phosphorylation-mediated activation of transcription factors is reversible. III) Molecular information that is transduced by protein phosphorylation can be amplified; many proteins can be phosphorylated in a very short period. IV) Because a phosphorylation site on a protein is usually defined by a short peptide sequence, the identity of a protein as a substrate for a certain kinase or phosphatase can be rapid (Marks, 1996).

Transcription in eukaryotic nuclei is catalyzed by three enzymes called RNA polymerase I, II and III. These RNA polymerases differ in their target

specificity; RNA polymerase I has 1 target gene: the precursor for 28S, 16S, and 5.8S ribosomal RNA. RNA polymerase III transcribes tRNAs, 5S rRNA, and a number of small RNAs including U6 snRNA. RNA polymerase II has the largest number of target genes and transcribes all protein-encoding genes (Marks, 1996). None of the eukaryotic RNA polymerases can recognize promoters or transcription initiation sites, nor initiate transcription on their own, they need transcription factors (Marks, 1996).

There are two classes of transcription factors: general transcription factors, which are required for all gene transcription with a specific RNA polymerase, and promoter-selective transcription factors, which are required for the transcription of a subset of genes transcribed by RNA polymerases II or III. These transcription factors are sequence specific DNA binding proteins (Marks, 1996). The majority of promoter-selective transcription factors function by recognition and binding to *cis*-acting transcription control elements on the DNA. These *cis*-acting elements are often part of larger regulatory entities called promoters or enhancers. The process of transcriptional activation by promoter-selective transcription factors involves a series of biochemically discernable steps:

a) The transcription factor has to move to the site of transcription, the nucleus.

b) In many cases, the transcription factor must homo- or heterodimerize in order to specifically recognize and bind to their cognate promoter element.

c) The factors must stimulate or repress the transcription of their target gene.

Transcriptional activation of immediate-early genes represents the primary nuclear response to signal reception. This process is independent of *de novo* protein synthesis. It is mediated by pre-existing transcription factors, which become activated in response to the incoming signal (reviewed in (Marks, 1996).

Regulation of transcription factors is important for ensuring correct cellular responses to signal transduction pathways. The transcription factor MI-ER1 is an early response gene to FGF activated signaling in *Xenopus*. The nuclear import of MI-ER1 into the nuclei of *Xenopus laevis* embryos is developmentally regulated. In the next section the principles of nuclear transport as well as the different mechanisms of regulation of nuclear import are discussed.

1.3 Nuclear transport of macromolecules

Eukaryotic cells are characterized by distinct nuclear and cytoplasmic compartments separated by the nuclear envelope (NE), a double membrane that is continuous with the endoplasmic reticulum. The nuclear envelope is penetrated by nuclear pore complexes (NPCs), which allow transport of molecules between the 2 compartments. During mitosis the nuclear envelope breaks down and the nuclear and cytoplasmic compartments can mix. However, during interphase all nucleocytoplasmic transport occurs through the NPC (Gorlich and Kutay, 1999). Amongst the molecules transported are nuclear proteins such as histones and transcription factors, which are transported from

the cytoplasm to the nucleus, but also RNAs like transfer RNA (tRNA), messenger RNA (mRNA) and ribosomal RNA (rRNA), which are synthesized by transcription in the nucleus and need to be transported to the cytoplasm where they function in translation (Gorlich and Kutay, 1999).

There are a few reasons for the complexity of the nuclear compartment found in eukaryotic cells compared to prokaryotes. First, the existence of a cell nucleus enables the eukaryotic cell to handle large amounts of genetic information. Second, the compartmentation allows regulation of key cellular events, for example by controlling the access of transcriptional regulators to chromatin, thereby controlling transcription. A third reason is the fact that eukaryotic genes are comprised of introns and exons, which requires splicing of primary transcripts before translation may occur. Translation of unspliced transcripts would result in proteins that are not only non-functional, but may act as dominant-negative inhibitors.

1.3.1 The Nuclear Pore Complex

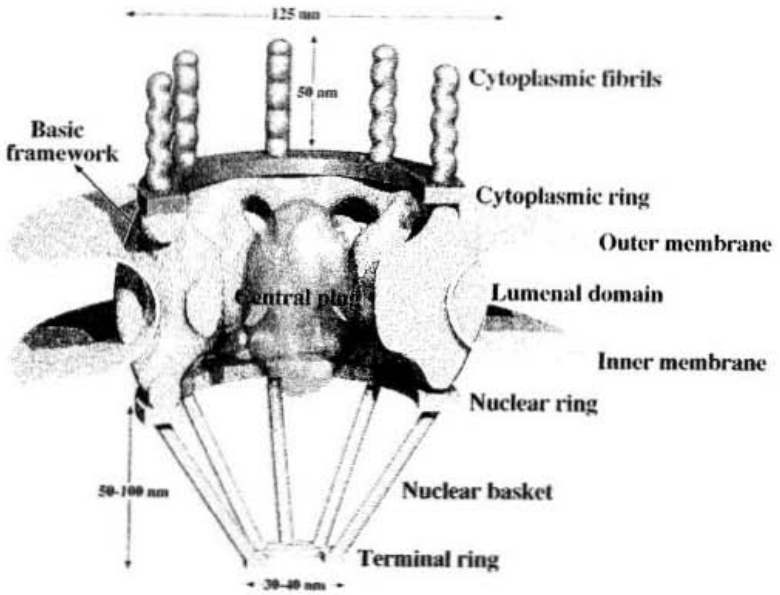
The Nuclear Pore Complex (NPC) consists of a membrane embedded basic framework made up of 8 multidomain spokes with two rings on each face. The ring facing the cytoplasm has eight 50nm fibrils extending into the cytosol, and the nuclear ring is capped with a basket-like assembly of eight thin, 50-100 nm filaments joint distally with a 30-40 nm diameter terminal ring (see figure 1.3.1, (Izaurralde *et al.*, 1999). The center of the basic framework harbors the

gated transport central plug, which is involved in the bi-directional transport of macromolecules bearing a transport signal.

The number of NPCs per cell appears to depend on the demand for nuclear transport and varies with cell size and synthetic and proliferative activity (Gorlich and Kutay, 1999). There are 189 NPCs in a yeast cell (Rout and Blobel, 1993), about 3000-5000 in a proliferating human cell, and about $5x10^7$ in a mature *Xenopus* oocyte (Cordes *et al.*, 1995).

NPCs are aqueous channels through which all nuclear transport occurs (Richardson *et al.*, 1988), for review, see (Bagley *et al.*, 2000; Rout and Aitchison, 2001; Stewart *et al.*, 2001). Passive diffusion through the pore is believed to be possible for molecules smaller than 9 nm in diameter, which corresponds to a protein with a molecular weight of about 45 kDa. However, it is shown that diffusion is a slow and obviously passive process and that even small proteins or RNAs smaller than 20-30 kDa, like histones (Jakel *et al.*, 1999) and tRNAs (Kutay *et al.*, 1998) normally cross the membrane by active transport, thus providing a mechanism for regulated transport.

Figure 1.3.1 Model of the NPC spanning the nuclear envelope, modified from (Izaurralde *et al.*, 1999).


125 nm
50 nm
Cytoplasmic fibrils
Basic framework
Cytoplasmic ring
Outer membrane
Lumenal domain
Inner membrane
Nuclear ring
50-100 nm
Nuclear basket
Terminal ring
30-40 nm

1.3.2 Import machinery

Passive diffusion does not require any specific interactions between the diffusing species and components of the NPC; it is fast for small molecules, but becomes restricted and inefficient as the diffusing objects approach a size limit of 40-60 kDa (Paine *et al.*, 1975); (Ribbeck and Gorlich, 2001). Facilitated transport requires specific interactions between the translocating molecule and the NPC and is therefore a highly selective process. Only some proteins, called importins and exportins (both also referred to as karyopherins) have the ability to directly interact with proteins on the surface of the NPC. Importins are the proteins that have the ability to direct cargo to the nucleus, whereas exportins export molecules from the nucleus to the cytoplasm. The proteins in the NPC are collectively called nucleoporins (Nigg, 1997). Many nucleoporins contain highly repetitive motifs conforming to either the consensus FXFG and/or GLFG (single letter amino acid code is used with X indicating any amino acid). These FG repeats are shown to interact with the transport receptors *in vitro* (Rexach and Blobel, 1995); reviewed in (Nigg, 1997).

The first transport receptor identified was importin-β that functions together with an adaptor, importin-α (Gorlich *et al.*, 1994; Gorlich *et al.*, 1995a; Gorlich and Kutay, 1999). Importin-α binds to the protein to be imported, whereas importin-β binds to the NPC (Adam and Adam, 1994; Gorlich *et al.*, 1995b; Imamoto *et al.*, 1995; Moroianu *et al.*, 1995a; Moroianu *et al.*, 1995b; Rexach and Blobel, 1995; Weis *et al.*, 1995). Both importins contain several hydrophobic

42 amino acid repeats, known as 'arm' repeats, named after the *Drosophila* gene *armadillo* (Nigg, 1997). There are at least four human α-importins, and homologues have been characterized in other eukaryotes (Moroianu, 1999). The importin-β family consists of about 10 homologues isolated from yeast and human (Moroianu, 1999).

Other important components of the import machinery are a small GTPase Ran and the RanGDP-interacting protein, p10/NTF-2. Ran was found to be an essential factor for almost all known nuclear import and export pathways. Ran is an abundant protein that is mainly localized in the nucleus and switches between GDP- and GTP-bound states. Ran has low intrinsic activity and requires regulators for efficient GTP hydrolysis or exchange of nucleotides (Gorlich and Mattaj, 1996; Moroianu, 1998; Moroianu, 1999; Nigg, 1997).

A significant feature of the RanGTPase system is its asymmetric distribution. The main GTPase-activating protein RanGAP (Bischoff *et al.*, 1995) and a Ran binding protein RanBP1 (Coutavas *et al.*, 1993), which together can achieve maximal GTPase activity, are mainly cytoplasmic (Richards *et al.*, 1996; Richards *et al.*, 1997). In contrast, the nucleotide exchange factor of Ran, RCC1, is mainly nuclear and bound to chromatin. This leads to the asymmetric distribution of Ran being converted to the GDP form in the cytoplasm, whereas in the nucleus Ran is mainly in the GTP-bound form through the action of RCC1. The result is a gradient of RanGTP across the NPC with low concentration in the cytoplasm and high concentration in the nucleus. Disruption of this gradient

results in blockage of nucleocytoplasmic transport (Izaurralde *et al.*, 1997; Richards *et al.*, 1997).

1.3.3 Signals for transport

Active transport of macromolecules is based on recognition of specific signals in the molecule that is transported. This was first found through a study of nucleoplasmin (Dingwall *et al.*, 1982). Nucleoplasmin is a large 165 kDa pentameric, thermostable protein found in *Xenopus laevis* oocytes and consists of subunits that each contain a protease resistant "core" and a protease sensitive "tail". Intact nucleoplasmin quickly enters the nucleus after injection in the cytoplasm of *Xenopus* oocytes. However, when all the tails are removed from the core by protease treatment, the residual core remains a pentamere but fails to localize to the nucleus. However, the tails show rapid entry into the nucleus. This experiment showed that the tails of nucleoplasmin contain a signal for nuclear entry (Dingwall *et al.*, 1982).

In 1984, the first nuclear import signal was identified by Kalderon *et al.*, (Kalderon *et al.*, 1984a; Kalderon *et al.*, 1984b) in the simian virus large T-antigen. A region necessary for the nuclear import of this protein was identified. The region identified had a stretch of five basic amino acids ^{127}PKKKRKV^{133} and it was shown that the substitution of the ^{128}Lys for Thr completely abolished nuclear translocation of the SV40 T-Ag (Kalderon *et al.*, 1984a). The nuclear import signal of nucleoplasmin was identified in 1991 and consists of two

interdependent regions of basic amino acids that are spaced by a linker of ten intervening amino acids (Robbins *et al.*, 1991). The import signals from nucleoplasmin and the SV40 T antigen are the prototypes of the classical nuclear localization signal (NLS); bipartite NLS and monopartite NLS respectively (Kalderon *et al.*, 1984a; Lanford and Butel, 1984; Robbins *et al.*, 1991).

Since the discovery of the SV40 T antigen and the nucleoplasmin NLS, many NLS have been described. Boulikas has described the requirements of a functional NLS, which are based on the comparison of hundreds of NLS that have been shown to be functional (Boulikas, 1994; Boulikas, 1997). In short, a core NLS is often a hexapeptide in which four amino acids are arginine or lysine. It does not contain any bulky, acidic or hydrophobic residues in the core or in the flanking regions. The core NLS is often flanked by proline or glycine, which function as helix breakers, for better exposure of the signal to the importins. There are three classes of classic NLS: The monopartite and bipartite NLS comprise the first 2 classes. The third class is exemplified by an NLS found in the human protein c-MYC. In this NLS only 3 out of 6 amino acids are positively charged and acidic residues in the flanking region are shown to increase the binding to import receptors (Hodel *et al.*, 2001; Makkerh *et al.*, 1996).

Some proteins are actively exported from the nucleus because their presence in the nucleus can be harmful for the cell survival or can interfere with cell control. Therefore these proteins carry a Nuclear Export Signal, NES. The prototypic NES is hydrophobic with high leucine content. NES of this type are

found in PKI (rat, (Wen *et al.*, 1995)), viral proteins like Rev (HIV-1, (Fischer *et al.*, 1995) and TFIIIA (*Xenopus*) (reviewed by (Nigg, 1997). NES are also found on regulators of transcription factors, like IκB, which can bind NF-κB in the nucleus and relocates to the cytoplasm as the complex NF-κB/I-κB, where it is located until re-activation of the NF-κB pathway. If I-κB did not possess a NES, NF-κB function is inhibited in the nucleus by I-κB, resulting in apoptosis (Henderson and Eleftheriou, 2000; Huang and Miyamoto, 2001). A NES is also found on MAPKK as localization of MAPKK in the nucleus leads to constitutive activation of the MAPK pathway. It is therefore important for MAPKK to be localized in the cytoplasm (Fukuda *et al.*, 1996; Fukuda *et al.*, 1997b; Fukuda *et al.*, 1997a).

There are also specific signals for nucleocytoplasmic transport found on shuttling proteins like the mRNA binding protein heterogeneous nuclear riboprotein, hnRNP A1. Shuttling proteins are proteins that have transport signals that can function in both import and export and thus import as well as export are strictly regulated. The shuttling protein hnRNP A1 contains a sequence that can function in import and export, the M9 motif. This approximately 40 amino acid motif is rich in glycine and aromatic amino acids and was shown to be important for the transport of mRNAs (Michael *et al.*, 1997; Siomi and Dreyfuss, 1995; Weighardt *et al.*, 1995).

Other signals for transport have been reported, like the hnRNP K bi-directional signal called hnRNP K nuclear signal, KNS (Michael *et al.*, 1997), SR-

proteins (Kataoka *et al.*, 1999)) and more (see (Nigg, 1997) and (Jans *et al.*, 2000)).

1.3.4 Mechanisms of transport

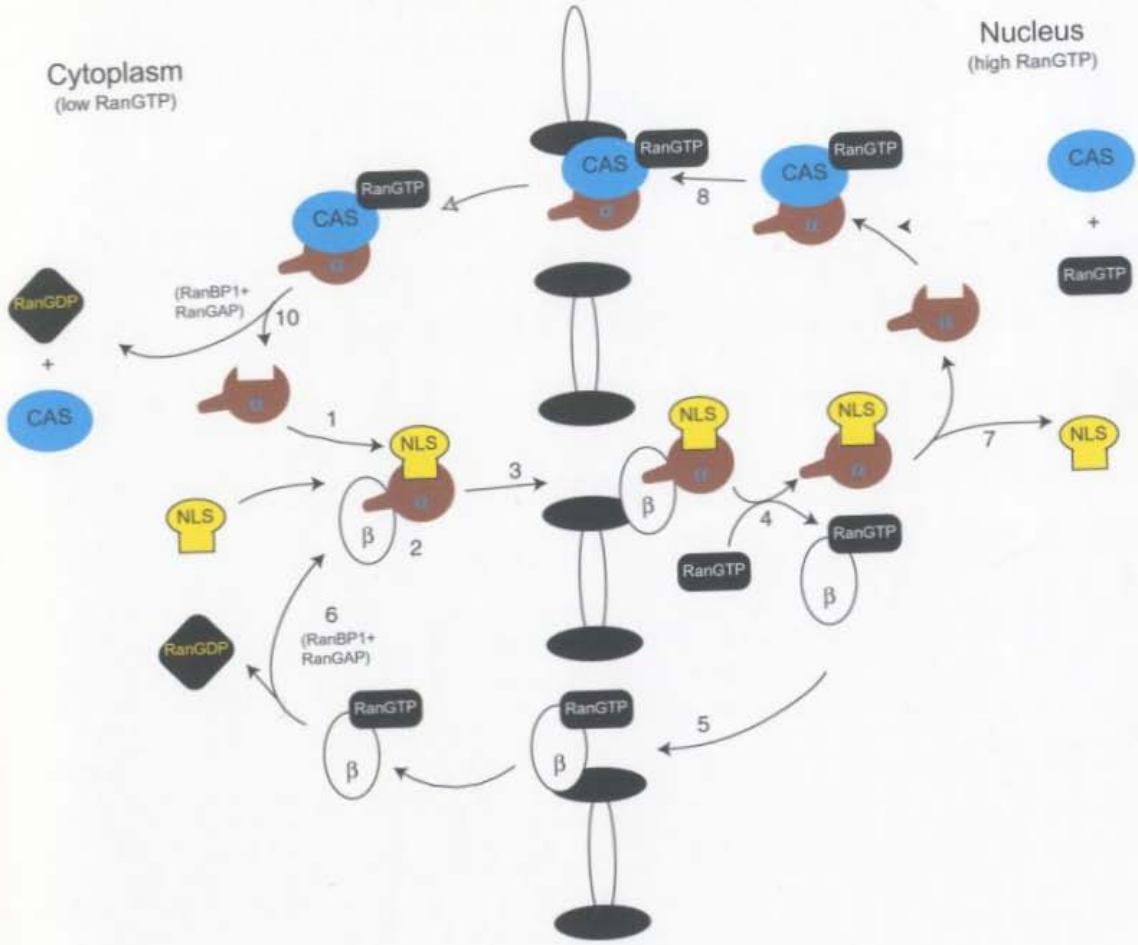
There are several different nuclear trafficking pathways which share a number of common features (reviewed in (Gorlich and Kutay, 1999; Gorlich and Mattaj, 1996; Stewart *et al.*, 2001). Generally, substrates do not directly interact with the NPC, but are transported bound to soluble transport receptors, which are recycled back to their original compartment after the translocation. The focus of my research is based on import of macromolecules into the nucleus; therefore this process will be described here. However, the general principles involving nuclear translocation are thought to be similar for other nuclear trafficking pathways (Gorlich and Kutay, 1999).

Translocation through the NPC can be divided into two steps: I) Docking of the substrate on the cytoplasmic face of the NPC. This requires the presence of an NLS, but is independent on metabolic energy. II) Translocation of the complex, which is shown to be dependent on GTP hydrolysis (Bischoff and Gorlich, 1997; Gorlich *et al.*, 1996; Izaurralde *et al.*, 1997; Melchior *et al.*, 1995). It was shown that in the absence of ATP, cargo is located on the cytoplasmic surface of the NPC, indicating that either the translocation or the release of the cargo on the other side of the NPC is energy dependent.

Figure 1.3.2 Schematic representation of the classical import pathway.

(Modified from (Gorlich and Kutay, 1999). For details see main text. In summary: the protein to be transported (NLS) binds to importin-α (α) and importin-β (β). The complex binds to the NPC and translocates to the nucleus. In the nucleus RanGTP binds to importin-β, releasing it from importin-α and the NLS containing protein. Importin-α binds to its export receptor CAS and RanGTP, thereby releasing the NLS bearing protein in the nucleus. The importin-β/ RanGTP complex binds to the NPC and translocates to the cytoplasm directly. Importin-α is exported by its export receptor CAS. GTP hydrolysis causes the release of the import receptors from RanGTP (in the case of importin-β) and CAS/RanGTP (in the case of importin-α). Importin-α and β are now able to bind other NLS-bearing proteins for subsequent rounds of import.

Abbreviations: α and β, importin-α and -β respectively; NLS indicates the import substrate.


Cytoplasm
(low RanGTP)
Nucleus
(high RanGTP)
CAS
RanGTP
RanGDP
(RanBP1+
RanGAP)
NLS
α
β
1
2
3
4
5
6
7
8
10

The processes of nuclear protein import are schematically shown in figure 1.3.2. In short: importin-α binds an NLS-bearing protein (1) and is bound by importin-β (2). This complex docks at the cytoplasmic side of the NPC and subsequently is transported through the NPC (3). Exactly how this translocation occurs is to date unknown (see (Gorlich and Kutay, 1999; Jans *et al.*, 2000; Stewart *et al.*, 2001). Binding of RanGTP to the importin-β binding site (IBB) on Importin-α causes release of Importin-β (4) from the complex and results in export of importin-β to the cytoplasm (5). In the cytoplasm RanBP1 and RanGAP cause ATP hydrolysis resulting in release of the RanGDP from importin-β (6). In turn, binding of RanGTP to importin-α in the region of the NLS-binding side causes release of the cargo in the nucleus (7). The RanGTP-Importin-α complex is exported by the transport receptor CAS (8). In the cytoplasm RanBP1 and RanGAP cause ATP hydrolysis resulting in release of CAS and the RanGDP from Importin-α, which can be used for subsequent rounds of import (9) (Gorlich and Kutay, 1999).

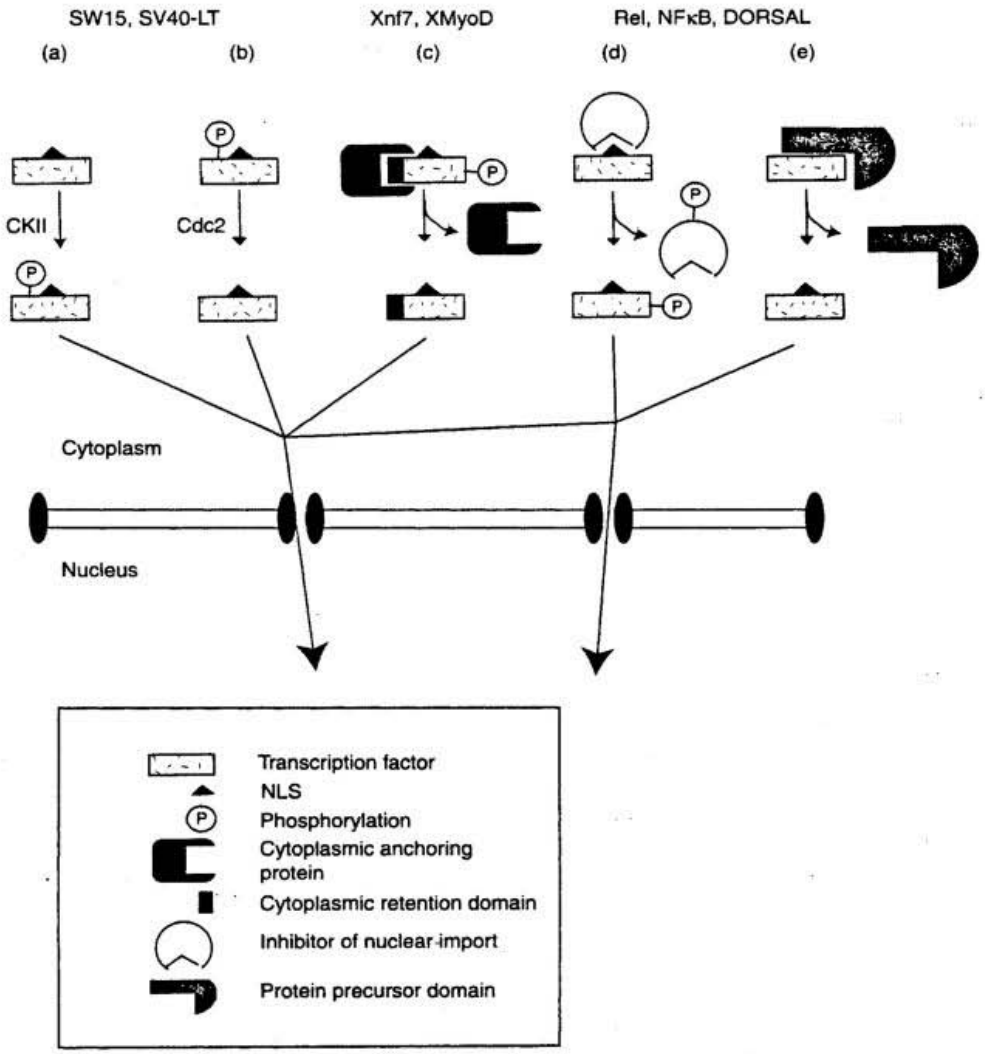
1.3.5 Regulation of nuclear transport

Nuclear import of proteins is not just dependent on the presence of a functional NLS. Most proteins containing a functional NLS are not constitutively translocated to the nucleus, while there are proteins larger than 45 kDa that have no functional NLS but are actively imported into the nucleus. This is because

nuclear translocation is tightly regulated. Cells use regulated nuclear transport in at least three ways. (I) Cells can use regulated transport as a molecular switch to control transcription of genes or to regulate transitions in cell cycle. (II) Regulated transport can also be used as a device to provide a delay between synthesis of a protein and its ability to act in the nucleus. (III) Regulated nuclear transport is used to ensure proper maturation of many biological activities as in the case of mRNA or tRNA or the assembly of ribosomal particles (Izaurralde and Adam, 1998).

Several steps in the movement of proteins across the nuclear envelope could be targeted for regulation (for review see (Kaffman and O'Shea, 1999; Vandromme *et al.*, 1996). (a) The cargo itself can be modified in a way that affects its ability to bind the import or export receptors. (b) The cargo-receptor complex can be tethered to an insoluble (not freely moving) cellular component thereby preventing it from binding the NPC. (c) The activity of the soluble transport machinery can be regulated. (d) The NPC itself might be modified in a way that affects its transport properties (Kaffman and O'Shea, 1999; Vandromme *et al.*, 1996) and references herein). Examples of each of the mechanisms of regulation are described below.

Figure 1.3.3 Model of different mechanisms that directly modulate transcription factor localization. Phosphorylation of sites adjacent to the NLS can influence the efficiency of NLS driven import, this can be either (a) positively by casein kinase II, or (b) negatively by cdc2 kinase. In (c), cytoplasmic retention of the transcription factor is achieved via a cytoplasmic retention domain in the factor, which binds to a cytoplasmic anchoring protein. In (d) and (e), the NLS is masked by either intermolecular interaction with an inhibitor (d), or intra-molecular masking with a subsequently proteolysed carboxyterminal portion of the molecule (e). In most of these cases phosphorylation or dephosphorylation (indicated by a circled P) is the event triggering the conformational changes that will allow nuclear transport of the protein. (Figure is modified from (Vandromme *et al.*, 1996).


SW15, SV40-LT
Xnf7, XMyoD
Rel, NFκB, DORSAL
(a)
(b)
(c)
(d)
(e)
CKII
Cdc2
P
Cytoplasm
Nucleus
Transcription factor
NLS
Phosphorylation
Cytoplasmic anchoring protein
Cytoplasmic retention domain
Inhibitor of nuclear import
Protein precursor domain

1.3.5.1 Protein context and competing signals

Possession of an NLS may not be sufficient for the translocation of the protein into the nucleus: translocation is also dependent on the position of the NLS within the protein. The NLS has to be exposed on the surface of the proteins to be accessible by import receptors. An example of regulation of nuclear transport by controlling accessability of the NLS is the NF-κB p50/105 protein, whereby the p50 comprises the N-terminal NLS-containing region of the p105 protein. Whereas the p50 protein is mainly nuclear, the p105 precursor is exclusively cytoplasmic due to intramolecular masking of the NLS (figure 1.3.3e) (Blank *et al.*, 1991; Henkel *et al.*, 1992). This is supported by the observation that monoclonal antibodies to the NLS recognize native p50, but fail to recognize p105, unless it has been denatured (Henkel *et al.*, 1992).

1.3.5.2 Piggy backing into the nucleus

There are examples of proteins larger than 45 kDa, which do not possess an NLS but are still able to translocate to the nucleus. This is caused by 'co-transport' with NLS-bearing proteins, so-called 'piggy back'. For example the non-histone nuclear high-mobility group protein, HMG1, is able to co-transport a specifically reacting 170-kDa monoclonal antibody to the nucleus. Also, the predominant localization of the NLS-deficient cyclin dependent kinase cdc2 in the nucleus is believed to be through complex formation with Cyclin B, which possesses three putative NLS. (reviewed in (Jans and Hubner, 1996).

1.3.5.3 Regulated or conditional nuclear translocation

Advantages for cytoplasmic localization of a transcription factor include the potential to control its nuclear activity and direct accessibility to cytoplasmic signal transducing factors. One mechanism regulating nuclear protein import is cytoplasmic retention, whereby a cytoplasmically localized "anchor" protein or retention factor specifically binds an NLS containing protein and prevents it from migrating to the nucleus, as is the case for *Xenopus* Zinc-finger protein XNF7 and its homologue in the newt, Pwa33 (figure 1.3.3c). The maternally expressed transcription factor XNF7 is first detected in the oocyte nucleus and then released into the cytoplasm after maturation. After fertilization the protein is retained in the cytoplasm until it re-enters the nucleus at MBT (Li *et al.*, 1994a). The state of XNF7 phosphorylation has been shown to regulate its nuclear/ cytoplasmic distribution. The cytoplasmic form of XNF7 is hyperphosphorylated, whereas the nuclear entry at MBT is coincident with its dephosphorylation, which results in release from its cytoplasmic anchor and subsequent nuclear import (Li *et al.*, 1994b). Cytoplasmic retention has also been described for c-fos, *Xenopus* MyoD, glucocorticoid receptor, and B-type cyclins (reviewed in (Jans and Hubner, 1996; Vandromme *et al.*, 1996). Cytoplasmic retention is dominant over the function of the NLS, even in the presence of an additional NLS.

Another example of regulated transport is NLS masking (figure 1.3.3d,e). The definition of NLS masking is when the NLS is not exposed on the surface of the

protein so it cannot be recognized and bound by the import receptors. This can be a result of different mechanisms:

- Intra- and intermolecular masking

An example of intramolecular masking is the masking of an NLS by the C-terminus of p105, a precursor of NF-κB p50 (Blank *et al.*, 1991; Henkel *et al.*, 1992) as described previously.

The active form of NF-κB is composed of the p50 and p65 (Rel A) protein components. The NF-KB-binding inhibitor protein IκB functions in retaining the NLS-carrying NF-κB p65 subunit in an inactive complex in the cytoplasm, by masking its NLS. Release from retention is effected by proteolytic degradation of IκB, which is triggered by IκB phosphorylation, and results in unmasking of the NF-κB p65-NLS (Zabel *et al.*, 1993). This is an example of intermolecular masking.

- NLS masking by phosphorylation

Phosphorylation sites in the vicinity of an NLS can influence the binding of the NLS to importins by changing the conformation or the charge. An example of NLS masking by phosphorylation is the cell cycle dependent exclusion of the *S. cerevisiae* TF SWI5. Nuclear translocation of SWI5 is affected by phosphorylation by cdc28 (Moll *et al.*, 1991), reviewed in (Jans and Hubner, 1996; Kaffman and O'Shea, 1999).

1.3.5.4 Phosphorylation enhancing NLS dependent transport

A well-described example of phosphorylation enhancing nuclear import comes from the casein kinase 2 (CKII) site close to the NLS (figure 1.3.3a). Phosphorylation on this site greatly influences the import kinetics; it increases the rate of import such that there is a maximum accumulation within 20 minutes compared to an import rate of close to 10 hours when phosphorylation is prevented. Aspartic acid can simulate phosphorylation at the CKII site in terms of accelerating the rate of nuclear import, implying that the negative charge provided by the aspartic acid residue at the CKII site is mechanistically important in this process (Jans and Jans, 1994).

In c-Rel, phosphorylation on a site 22 amino acids N-terminal of the NLS, by protein kinase A (PKA) has shown to enhance the import. Substitution for alanine at this site abolishes nuclear localization of the c-Rel proto-oncogene, while substitution with an aspartic acid residue stimulated nuclear transport. The 22 amino acid distance seems to be critical for the function of this PKA site. The PKA site is conserved between other members of the Rel family (Kidd, 1992).

1.3.5.5 Developmentally regulated NLSs

Alternative NLS, different from the classic NLS have been described. One example is developmentally regulated NLS, which are only functional during specific stages of embryonic development (Standiford and Richter, 1992). A developmentally regulated NLS (drNLS) has been reported for a 45 amino acid

region of the adenovirus E1a protein (Standiford and Richter, 1992), which is distinct from the constitutive E1a NLS ^{285}KRPRP^{289}. This drNLS is functional in targeting a carrier protein to the nucleus up to the neurula stage of *Xenopus* embryonic development, when it is activated in a hierarchical fashion among the embryonic germ layers (Standiford and Richter, 1992). Significantly, a sequence homologous to the drNLS of E1a has been identified in the rat and mouse and human glucocorticoid receptor (GR) hormone-binding domain (Picard *et al.*, 1988).

1.3.5.6 CcN motif and phosphorylation regulated NLS

There appears to be a consensus of phosphorylation sites close to the NLS, which modify the action of these NLSs. The phosphorylation sites, together with the NLS constitute phosphorylation mediated regulatory modules for nuclear localization, referred to as phosphorylation-regulated NLS (prNLS) (Jans, 1995). Phosphorylation can inhibit or enhance nuclear translocation (Figure 1.3.3a, b).

An example of phosphorylation-regulated NLS is found in the SV40 large T-antigen. The SV40 large T-antigen contains a specified prNLS, consisting of both enhancing and inhibitory phosphorylation sites. The CKII site increases the rate of NLS-dependent nuclear import, whereas phosphorylation at the cdc2 site adjacent to the NLS inhibits transport, markedly reducing the level of maximum nuclear accumulation (Figure 1.3.3a, b) (Jans *et al.*, 1991). The CKII and cdc2

sites appear to function completely independent of one another. This regulatory module is called a CcN motif, for CKII, cdc2 and NLS (Jans *et al.*, 1991).

Phosphorylation of the cdc2 kinase site in the vicinity of the NLS in SV40 large T-Ag inhibited nuclear translocation by about 70% in transfected HTC (Rat hepatoma) cells (Jans, 1995). Substitution of this phosphorylation site to alanine mimiced a dephosphorylated amino acid and resulted in increased nuclear translocation of the SV40 large T-Ag. However, substitution with aspartic acid mimiced a phosphorylated amino acid because of the presence of a negative charge and resulted in cytoplasmic localization of the SV40 large T-Ag.

The casein kinase 2 (CKII) site in the CcN motif is known to be important for the rate of nuclear import (Jans and Jans, 1994). Phosphorylation on this site greatly influences the import kinetics; it increased the rate of import such that there is a maximum accumulation within 20 minutes compared to an import rate of close to 10 hours when phosphorylation is prevented. Aspartic acid can simulate phosphorylation at the CKII site in terms of accelerating the rate of nuclear import, implying that the negative charge provided by the aspartic acid residue at the CKII site is mechanistically important in this process (Jans and Jans, 1994).

A variety of proteins possess putative CcN motifs, including oncogene products, viral proteins as well as transcription factors. More specifically, CcN motifs have been described for mSTI1 (Longshaw *et al.*, 2000), p53 (Bischoff *et al.*, 1990), lamin (Loewinger and McKeon, 1988), nucleoplasmin (Dingwall *et al.*,

1988; Robbins *et al.*, 1991), SWI5 (Jans *et al.*, 1995) and nucleolin (Schwab and Dreyer, 1997).

Other kinases have been implicated in the regulation of protein import, in particular the cAMP-dependent kinase (PKA). Direct phosphorylation by PKA is shown to regulate nuclear import of the Rel family of nuclear factors (Mosialos *et al.*, 1991). This PKA site is not in the vicinity of the NLS, showing that regulation of nuclear import by phosphorylation can occur at any site within the protein. An advantage of using phosphorylation as a mechanism to regulate nuclear import is that phosphorylation occurs quickly and it is a reversible process, ensuring rapid regulation of cellular processes, which are dependent on the import of these molecules (Jans and Hubner, 1996; Kaffman and O'Shea, 1999).

1.3.5.7 Regulation of the soluble transport machinery

The soluble transport machinery includes the importins, exportins and components of the RanGTPase cycle. These factors could easily be targeted for regulation, but not a lot of information is available to date. However, evidence has shown that import receptors can be tissue specific and have different binding efficiencies for different NLS (Gorlich and Kutay, 1999; Kohler *et al.*, 1999). Also the change in localization of importin-α in response to cell cycle position is shown in *Drosophila*. Pendulin (an importin-α orthologue in *D. melanogaster*) enters the nucleus at the onset of mitosis (Kussel and Frasch, 1995; Torok *et al.*, 1995). The mechanism for this change in localization is not known, and it might be

related to the ability of importin-α to be phosphorylated (Torok *et al.*, 1995). However, this change in localization suggests that the relative import- or export rates of Pendulin are subject to cell cycle regulation, possibly providing a mechanism to block import of a large number of cargoes during mitosis (Kaffman and O'Shea, 1999).

1.3.5.8 Regulation of the Nuclear Pore Complex

The amount of NPCs in the nuclear membrane is dependent on the proliferative state of the cell and thus the need for nuclear transport. A transcriptionally quiescent cell has less NPCs than a cell which is transcriptionally active (Cordes *et al.*, 1995; Gorlich and Kutay, 1999; Rout and Blobel, 1993).

The specific regulation of NPC has been studied using a unique system: *Tetrahema thermophilia*, a unicellular organism that has two nuclei. These nuclei, the macronucleus and the micronucleus share the same cytoplasm (Gorovsky, 1973). The two nuclei serve distinct cellular functions and share some proteins, but each also contains a unique set of proteins (Gorovsky, 1973). Because these two nuclei share the same soluble transport system, it was proposed that the asymmetric distribution of proteins is caused by different transport properties of the NPC in the nuclei (White *et al.*, 1989).

Another example of regulation of NPC is given by experiments that show that quiescent cells show a slower rate of nuclear import of large gold particles

coated with nucleoplasmin, compared to proliferating cells. However, import rates of smaller gold particles were not different between these cells, indicating that the difference in import rates was size dependent (Feldherr and Akin, 1990; Feldherr and Akin, 1991). Quiescent and proliferating cells were fused together to investigate whether this difference in import rates was caused by differences in the soluble import machinery or the NPC. Injection of large gold particles into the cytoplasm of this heterokaryon showed that import into the nucleus of proliferating cells was more efficient compared with import into the quiescent nucleus (Feldherr and Akin, 1993). This supports the notion that activity of the NPC might be regulated (Kaffman and O'Shea, 1999).

It is clear that the regulation of NLS-dependent nuclear import is affected at different levels, which provides an excellent mechanism of control of diverse biological processes.

1.4 Mesoderm induction early response 1, *mi-er1*

1.4.1 *mi-er1* isolation

Xenopus laevis

The *mi-er1* gene was first discovered by Paterno *et al.* (1997) using a technique called differential display (Liang and Pardee, 1992), to identify and characterize transcripts that were expressed early during cellular response to FGF-2 (Paterno *et al.*, 1997). The cDNA sequence isolated from the differential display was used to isolate a full-length 2.3 kilo-base-pair (kb) cDNA from a *Xenopus* blastula library. The cDNA contains a predicted 1497 base pair (bp) open reading frame (ORF), a 214 bp 5' region containing several stopcodons in all three frames and a 626 bp 3' untranslated region (UTR). The start codon is shown to be within a Kozak consensus sequence containing a purine at position –3 and a guanine at position +4. The ORF gives rise to a protein of 493 amino acids, with a predicted size of 55 kDa, which contains 3 regions of similarity to the metastasis associated protein, mta-1 (Paterno *et al.*, 1997).

Human

There are 6 isoforms of human MI-ER1 isolated to date, which are generated by alternative splicing of a single gene. The first form of human *er1* was isolated by PCR from a human testis library using *Xenopus mi-er1* (*Xmi-er1*) specific primers and was designated h*mi-er1α*. A full-length cDNA containing a

1296 bp ORF surrounded by a 68 bp 5' and a 210 bp 3' UTR was isolated. The protein of 432 amino acids was shown to be 91% similar at the amino acid level to XMI-ER1 (Paterno *et al.*, 1998). The second human isoform, h*mi-er1*β, was isolated by PCR from MDA-468, a human breast cancer carcinoma cell line. This isoform is 1536bp in length. h*mi-er1*α and β are identical in the N-terminus, but differ at the C-terminus due to alternative splicing of the RNA transcript. Also, both *mi-er1*α and β transcripts are formed by use of two transcription start sites, which have corresponding promoter sites. Inclusion of an alternate exon in either *mi-er1*α or β results in two more isoforms, making a total of six isoforms by alternative splicing of the primary transcript (Gillespie and Paterno, unpublished data). It was shown that transcription of different h*mi-er1*α and β transcripts is derived from at least two different specific promoter regions, which results in alternative isoforms (Ding and Paterno *et al.*, unpublished data).

1.4.2 Expression of *mi-er1* mRNA

Xenopus

Investigation of *mi-er1* showed it is an immediate early gene. Its response is rapid and not dependent on *de novo* protein synthesis. The FGF induced increase in *mi-er1* levels was measured in the presence or absence of cycloheximide, a known inhibitor of protein synthesis. Cycloheximide inhibited 90% of ^{35}S-methionine incorporation into trichloroacetic acid (TCA) precipitated

material, but did not prevent the FGF-2 induced increase in MI-ER1 levels, demonstrating *mi-er1* is an immediate early gene.

Northern analysis of RNA extracted from *Xenopus laevis* embryos at different stages during development showed the presence of *mi-er1* during initial cleavage stages, prior to the start of zygotic transcription, indicating that *mi-er1* is maternally derived. The levels of mRNA expression were constant during early cleavage (stages 2, 6 and 7), increased slightly at blastula stage and decreased 6-fold during gastrula, neurula and tailbud stages (stages 12, 17 and 22). The levels then remained below detectable levels during subsequent development (stages 30 and 41) (Paterno *et al.*, 1997).

FGF expression was shown to be important for mesoderm induction in embryonic tissue at blastula stages (Slack *et al.*, 1987), therefore the expression of *mi-er1* at blastula and gastrula stages was examined at 1 hour intervals using quantitative reverse transcription- polymerase chain reaction (RT-PCR). It was shown that expression levels of *mi-er1* increased two-fold from early blastula (stages 7 and 8) to late blastula (stages 8 and 9) followed by a five fold decrease at gastrulation (stage 10) (Paterno *et al.*, 1997).

Human

The expression of h*mi-er1*α in normal and neoplastic tissue has been investigated and described previously (Paterno *et al.*, 1998). Expression of h*mi-er1* was barely detectable by Northern blotting of total RNA from eight human

tissues. Dot blot analysis of messenger RNA confirmed that h*mi-er1* was expressed at very low levels in 50 normal human tissues. RT-PCR analysis of breast carcinoma cell lines showed *mi-er1* expression, while no *mi-er1* was detected in three normal breast cell lines. Varying levels of *mi-er1* were detected in breast cancer tissues examined.

1.4.3 XMI-ER1 protein expression

In vitro translation of *Xmi-er1* followed by immunoprecipitation using an anti-MI-ER1 antibody and separation of the protein by SDS-PAGE, showed that although the predicted size is 55 kDa, the protein runs around 78 kDa, much higher than expected. This is probably due to the presence of acidic domains in the N-terminus of the protein, which may alter the net charge of the protein even when separated on denaturing gels (Bell and Bell, 1988). When transfected into NIH 3T3 cells, translated MI-ER1 is exclusively translocated to the nuclei of these cells, indicating that in these cells MI-ER1 is constitutively targeted to the nucleus (Paterno *et al.*, 1997).

The subcellular localization of MI-ER1 during *Xenopus laevis* development has been well investigated. MI-ER1 is expressed during early stages of development, but only localizes to the nucleus around the mid-blastula transition. Using an anti-MI-ER1 antibody against the C-terminus of MI-ER1, protein can be detected during different stages of development by whole mount staining. Whole mount staining of albino embryos has shown that MI-ER1 is localized in the

cytoplasm during early stages, is found to be nuclear in marginal zone cells of stage 8 blastulae and that, over time, more and more nuclei become stained. By late blastula stage (stage 9) MI-ER1 is found in all nuclei in the animal hemisphere and by early gastrula (stage 10) ubiquitous nuclear staining is observed. Subsequently, during tailbud stages, nuclear staining decreases and later in development during tadpole stages, the only nuclear staining detected is in some endodermal tissue (Luchman *et al.*, 1999).

1.4.4 Protein-domains in XMI-ER1 and their possible function

Putative domains and motifs found in XMI-ER1 were identified by computer analysis using computer analysis, like MOTIF (www.motif.genome.ad.jp), PSORT (psort.nibb.ac.jp) and Domain Architecture Retrieval Tool (DART; www.ncbi.nlm.nih.gov/Structure/lexington/lexington.cgi). The motifs and domains found are indicated in figure 1.4.1. Since XMI-ER1 and the hMI-ER1 isoforms are about 90% similar at the amino acid level, most motifs are found on all predicted proteins. In this section some of the motifs found in XMI-ER1 will be described.

Figure 1.4.1 Amino acid sequence of *Xenopus* MI-ER1, indicating the motifs found by computer analysis of the sequence. The acidic domains, ELM2 domain, SANT domain, proline rich region and putative NLS are indicated, see boxed legend.

1 MAEPSLRTAS PGGSAASDDH EFEPSADMLV HEFDDEQTLE EEEMLEGEVN

51 FTSEIEHLER ESEMPIDELL RLYGYGSTVP LPGEEDEEDM DNDCNSGCSG

101 EIKDEAIKDS SGQEDETQSS NDDPTPSFTC RDVREVIRPR RCKYFDTNHE

151 IEEESEDDED YVPSEDWKKE IMVGSMFQAE IPVGICKYRE TEKVYENDDQ

201 LLWNPEYVME ERVIDFLNEA SRRTCEERGL DAIPEGSHIK DNEQALYEHV

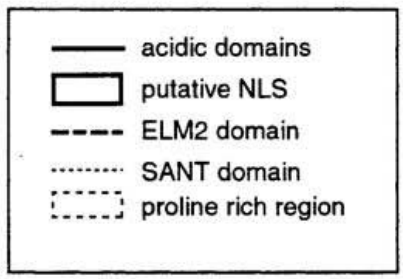
251 KCNFDTEEAL RRLRFNVKAA REELSVWTEE ECRNFEQGLK AYGKDFHLIQ

301 ANKVRTRSVG ECVAFYYMWK KSERYDFFAQ QTRFGKKKYN LHPGVTDYMD

351 RLLDESESAT SSRAPSPPPT TSNSNTSQSE KEDCTASNNT QNGVSVNGPC

401 AITAYKDEAK QGVHLNGPTI SSSDPSSNET DTNGYNRENV TDDSRFSHTS

451 GKTDTNPDDT NERPIKRQRM DSPGKESTGS SEFSQEVFSH GEV*



a. Acidic domains

MI-ER1 has four stretches rich in acidic residues, underlined with a solid line in figure 1.4.1. Stretches of acidic residues have previously been described to be characteristic for acidic transcription activation domains. Acidic activator domains were first described in yeast Gal4 and GCN4 activator proteins (Blair *et al.*, 1994; Hope and Struhl, 1986; Ma and Ptashne, 1987). In the Gal4 protein, transcriptional activation is mediated by acidic amino acid residues and activation closely correlates with net negative charge. This indicated that acidic amino acids are crucial for transactivation. However, several studies with acidic activation domains indicated the importance of hydrophobic amino acids for activation (Blair *et al.*, 1994; Cress and Triezenberg, 1991; Lin *et al.*, 1994). Further studies of acidic domains will help identify conserved amino acids that are important for the activation function of these proteins (Kistanova *et al.*, 2001).

The possibility of MI-ER1 functioning as a transcription factor was investigated by testing the transactivation potential of various regions of the MI-ER1 protein. Constructs containing different regions of *MI-ER1* fused to the Gal4 DNA binding domain were used along with a chloramphenicol acetyltransferase (CAT) reporter plasmid in transient transfections (Paterno *et al.*, 1997). It was shown that although full length MI-ER1 did not activate transcription, the N-terminal region (MI-ER1-175) containing all four acidic stretches stimulated transcription 10-fold. MI-ER1-98, containing only the first three acidic domains

stimulated transcription 80-fold. Fusions with other parts of MI-ER1 did not show any transcriptional activation. These results indicate that MI-ER1 has the potential to function as a transcriptional activator. However, HMI-ER1α and HMI-ER1β were recently found to have potent transcriptional repression activity (Ding *et al.*, unpublished data).

b. SANT domain

The SANT domain is similar to the DNA binding domain (DBD) of the oncoprotein MYB (Aasland *et al.*, 1996). The SANT domain was first found in SWI3 (a yeast component of the SWI-SNF complex (Peterson and Tamkun, 1995)), ADA2 (a component of a histone acetylase complex; (Horiuchi *et al.*, 1995)), transcriptional co-repressor N-CoR (Horlein *et al.*, 1995) and TFIIIB (a basal polymerase III transcription factor in yeast; (Kassavetis *et al.*, 1995), hence the name SANT domain. SANT-myb DBD alignments show a number of highly conserved residues, including the three tryptophan aromatic residues, which are shown to be important for formation of the hydrophobic core of the myb repeats. The similarity of the SANT domain to myb DBDs is further supported by the SANT secondary structure prediction, resembling the homeodomain helix-loop-helix motif (Aasland *et al.*, 1996). To date, there are 123 aligned sequences as found by a conserved domain (cd)-search (RPS-BLAST 2.2.1, NCBI; http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi).

The SANT domain of MI-ER1 does not seem to bind DNA directly, and a function of this domain has not been found in MI-ER1 (Paterno *et al.*, unpublished results). In CoRest, a transcriptional regulator containing two SANT domains, the N-terminus, which contains the SANT domain, is shown to be important for the co-repression function of the protein, due to interaction of the SANT domain with the histone deacetylases HDAC1/2. This interaction of the SANT domain with HDAC1/2 is also shown in MTA1 and MTA2 (Humphrey *et al.*, 2001; You *et al.*, 2001). More recently, Sterner *et al.*, have shown that the N-terminal region of the yeast ADA-2 SANT domain is responsible for binding histone acetylase, HAT (Sterner *et al.*, 2002). Histone deacetylases and histone acetylases are important regulators of transcription, because of their ability to influence chromatin structure. By removing acetyl groups from histones, access of the transcription machinery is inhibited and transcription cannot occur. However, when acetyl groups are added to histones by histone acetylase (HAT) transcription can occur. Therefore binding of a transcription factor like MI-ER1 to regulators of transcription, like HDAC and HAT implicates tight regulation of transcription.

c. ELM2 domain

The ELM2 domain is a sequence similar to that of a domain found in the EGL-27 protein as well as in MTA1 (Solari *et al.*, 1999). The ELM2 domain of human MI-ER1 is found to be responsible for complexing with HDAC1. This is related to the repressor function of hMI-ER1 (Ding and Paterno, unpublished data).

Interestingly, CoRest contains an ELM2 domain upstream of the N-terminal SANT domain, which was included in the constructs used in the experiments for the HDAC1/2 interactions. Therefore, it is possible that in CoRest the ELM2 domain and not the SANT domain is responsible for this interaction.

Most proteins that contain a ELM2 domain also contain one or more putative SANT domains, as predicted by BLAST search (RPS-BLAST 2.2.1, NCBI; http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). Presence of a putative HDAC binding domain and a HAT binding domain within a protein suggests co-operation between these domains in regulating accessibility of the transcription machinery to the chromatin.

d. Proline rich region

Proline rich motifs are characterized by the presence of the consensus PXXP tetrapeptide, in which the P represent prolines and X represents any amino acid residue. This consensus sequence is found in all proline rich proteins identified (Alexandropoulos *et al.*, 1995). It is known that SH3 (Src homology 3) domains (*vide infra*) recognize proline rich sequences and that all known SH3 binding proteins contain proline rich regions with at least one PXXP motif (Alexandropoulos *et al.*, 1995; Cohen *et al.*, 1995; Lowenstein *et al.*, 1992; Scorilas *et al.*, 2001; Sparks *et al.*, 1996; Wang *et al.*, 2000). PXXP motifs have been identified in a number of diverse proteins such as epidermal growth factors (EGF), phosphatidylinositol-3-kinase (PI3'K), and, more recently the small

GTPase R-Ras protein and members of the R-Ras superfamily such as the TC21 protein (Lowenstein *et al.*, 1992; Wang *et al.*, 2000).

Proline rich regions display a variety of functions, based on the specific and unique structure of proline in which the side-chain is cyclized onto the backbone nitrogen atom. This results in a restriction of the conformation of the proline and its proceeding residue (MacArthur and Thornton, 1991). Due to the relative rigidity of the polyproline region, the binding with other proteins is weak, which has great advantages for fast regulation of binding to these regions. Changes in the sequence of proline rich region or its binding domain, either by sequence changes or by covalent modification such as phosphorylation, can alter the strength of the interaction (Kay *et al.*, 2000). These changes account for the large variety of proline motifs in proteins.

Proline rich regions are commonly found in situations requiring the rapid recruitment or interchange of several proteins, such as during initiation of transcription, signaling cascades, and cytoskeletal rearrangements (Kay *et al.*, 2000). For example, the proline rich region on son of sevenless (Sos1) binds to the two SH3 domains of Grb2, thereby bringing Sos1 to the cell membrane following receptor activation, where Sos1 in turn then activates the Ras pathway (Williamson, 1994).

A variety of proteins involved in intracellular signal transduction pathways contain Src homology 3 (SH3) domains. Originally described in the protein tyrosine kinase Src, these domains consist of about 50 amino acids (Koch *et al.*,

1991) and bind to proline-rich regions in other intracellular proteins. While SH3 domains from different proteins have poorly conserved sequences, they share a similar overall structure. SH3-containing kinases such as *Lyn* and *Hck*, both members of the Src family, may use these domains to bind substrates or interact with regulatory proteins. Other proteins, termed adaptors, are non-catalytic; yet contain several SH3 and/or SH2 domains. These domains allow them to interact with multiple protein partners to form cytoplasmic signaling complexes. One such adaptor protein is Grb2, which links son-of-sevenless (Sos) and GDP/GTP exchanger, and the activated tyrosine kinase receptor during initiation of the Ras signaling pathway. Additional processes that may be dependent upon SH3 function include vesicle trafficking, cytoskeletal organization, and the targeting of proteins to subcellular locations. These diverse roles make SH3 domains the subject of intense effort to further elucidate their complex interactions and functions (for review see Klint and Claesson-Welsh, 1999 and Kay *et al.* 2000).

e. Nuclear Localization Signals

NLS are the signals necessary for import of large proteins into the nucleus. There are many proteins with different NLS, mostly existing of a region of basic amino acids, arginines and lysines. Initially two putative NLS were identified in MI-ER1 as identified by computer analysis using PSORT (http://psort.nibb.ac.jp) (Nakai and Horton, 1999). Analysis of the basic amino acids and comparison of

groups of basic amino acids to NLS described in the literature showed that MI-ER1 has 2 additional putative NLS (Post *et al.*, 2001).

The analysis of the function of each of these NLS and the regulation of nuclear transport during *Xenopus laevis* development will be described in this thesis.

f. Phosphorylation sites

An important feature of signaling molecules is the ability to be phosphorylated for activation. Putative phosphorylation sites were identified by computer analysis of the protein sequence using pbase_predict (www.cbs.dtu.dk/databases/Phospho-Base/predict/predform.html) and NetPhos (www.cbs.dtu.dk/services/NetPhos) (Blom *et al.*, 1999). MI-ER1 has multiple putative phosphorylation sites, suggesting MI-ER1 may be phosphorylated (Table 1.4.1). Whether these phosphorylation sites are functional in MI-ER1 has yet to be investigated.

MI-ER1 is an immediate early gene in response to FGF. The presence of different conserved motifs in MI-ER1 implicates an important role in signal transduction: It has potent transcriptional regulatory activity and is translocated to the nucleus, indicating a role as a transcriptional regulator. The presence of a proline rich region, which might function as an SH3 binding domain, and the early response to FGF implies a role in the FGF signaling cascade. The presence of the ELM2 domain, which functions to bind HDAC1, and a SANT domain, which

may bind HAT, indicates an important function in modulation of chromatin state and subsequent trans-activation or repression of transcription.

Table 1.4.1 Example of putative phosphorylation sites predicted by pbase_predict and NetPhos.

Phosphorylation site	Kinase	NetPhos 2.0 prediction
S-10 LRTA**S**PGGS	CaMII, G3K3, PKA	0.993
T-38 DDEQ**T**LEEE	CKII	0.977
S-53 VNFT**S**EIEH	CKII	0.804
S-62 LERE**S**EMPI	PKA	0.942
S-111 IKDS**S**GQED	CKII	0.991
T-191 KYRE**T**EKVY	PKC	0.972
Y-247 EQAL**Y**EHVK	?	0.709
S-308 VRTR**S**VGEC	CaMII, CKII, p70s6k,	0.995
S-366 SRAP**S**PPPT	CaMII, CKI, PKA	0.996
S-472 QRMD**S**PGKE	CaMII, p34cdc2, p70s6k, PKA.	0.997
S-489 QEVF**S**HGEV	CKII	0.973

1.5 Objectives

Regulation of transcription factors is important for ensuring correct cellular responses to signal transduction pathways. One step in the regulation of transcription factors is regulation of nuclear import. Regulation of import can be modulated by phosphorylation/ dephosphorylation of the transported protein, masking of the nuclear localization signal(s), cytoplasmic retention by binding to a cytoplasmic anchoring protein, modulation of the import machinery, and possibly interplay between these mechanisms.

In this study, the regulation of nuclear import of an FGF immediate early gene in *Xenopus*: the transcription factor mesoderm-induction early response 1, MI-ER1, was examined. MI-ER1 is localized in the cytoplasm of developing *Xenopus laevis* embryos and translocates to the nucleus beginning stage 8. Investigating the type of regulation of nuclear transport will bring us a step closer in identifying the role MI-ER1 plays after FGF activation of the signal transduction pathway.

Objective 1: To identify and characterize NLS sequences found in XMI-ER1.

Analysis of the XMI-ER1 sequence identified four putative NLS, which might be responsible for the nuclear translocation of XMI-ER1 in cells and during embryonic development. My hypothesis is that XMI-ER1 is translocated to the nucleus by binding of the import receptors to one or more of the putative NLS in MI-ER1, and thus is translocated by active transport.

A functional NLS is defined as being able to carry a non-nuclear protein into the nucleus. As well, deletion of the NLS results in cytoplasmic localization of the protein. According to this definition, the efficiency of each of the four putative NLS can be determined by addition of a non-nuclear protein, like enhanced green fluorescent protein (EGFP) or β-galactosidase, to each of the NLS regions and measurement of the ability to translocate to the nucleus. If a region containing a functional NLS is localized in the nucleus, the necessity of the NLS can be determined by deletion of the NLS (or part thereof) region in the context of the full-length protein.

To identify functional NLS in MI-ER1 the protein will be divided into 6 regions. The regions are chosen based on the position of putative NLS in the sequence. EGFP or β-GAL will be fused to these regions and the fusion constructs will be transfected into NIH 3T3 cells. Subcellular localization of the fusion proteins can be determined by microscopy. After identification of a region containing a functional NLS, the NLS will be further analyzed by deletion of the NLS as well as fusion of the smallest NLS region to EGFP or β-GAL. This will be followed by transfection into NIH 3T3 cells and determination of the subcellular localization of the overexpressed protein by microscopy.

Objective2: Elucidation of the regulation of XMI-ER1 nuclear translocation during embryonic development.

Maternal transcription factors are present in high concentrations in the developing embryo. If these factors could translocate to the nucleus before MBT, the high levels present in the nuclei could result in an early start of transcription. Therefore, regulation of nuclear import of these transcription factors is an important means by which the cell can ensure that transcription occurs at a specific time during development.

There are different levels at which nuclear translocation can be regulated, and the most direct regulation is modification of the transcription factor itself whereby nuclear import is blocked for this factor only. An important means of regulation of transcription factors is phosphorylation. Phosphorylation of proteins is fast and reversible and can enhance or inhibit nuclear translocation by interfering with the conformation of the protein. Conformational changes can influence the NLS directly by masking. MI-ER1 contains numerous phosphorylation sites, including a putative CcN motif as described for the SV40 large-T antigen. Investigation of the function of this site will be performed by mutational analysis of a putative cdc2 kinase site in the vicinity of the functional NLS.

Also, conformational changes can influence protein-protein interaction sites, influencing the binding to cytoplasmic anchoring proteins or proteins that mask the NLS, as in XNF7 and NF-κB, respectively.

My hypothesis is that XMI-ER1 translocation during embryonic development is regulated by inhibition of the NLS function. This could occur

either by NLS masking (as in the case of IκB binding to NF-κB) or by binding to an anchoring protein in the cytoplasm (as in the case of XNF7). Investigation into the regulation of cytoplasmic retention can be performed by addition of an NLS to the N-terminus of MI-ER1 and injection into *Xenopus* embryos followed by staining of the protein to observe subcellular localization. If addition of the NLS causes premature nuclear localization of MI-ER1 (before stage 8.5-9) the import of MI-ER1 is regulated by NLS-masking. Conversely, if MI-ER1 is retained in the cytoplasm by cytoplasmic anchoring, addition of the NLS will not interfere with the normal timing of nuclear localization. Moreover, different regions of MI-ER1 can be fused to the additional NLS to investigate the exact region of regulation.

For my experiments, the NLS of XNF7 will be fused to MI-ER1 to test whether MI-ER1 is localized in the cytoplasm by binding to a cytoplasmic anchor. The XNF7 NLS was chosen based on an observation of Li *et al.* (1994), that the XNF7 NLS can translocate to the nuclei of *Xenopus* embryos before the mid-blastula stages (Li *et al.*, 1994a). In addition, to test if binding to a cytoplasmic anchor is responsible for the cytoplasmic localization during early *Xenopus* development, I will attempt to identify the smallest region in MI-ER1 responsible for binding the cytoplasmic anchor. If however, localization of MI-ER1 in the cytoplasm is due to NLS masking, I will investigate whether masking is intra- or intermolecular.

Chapter 2 Materials & Methods

2.1 Generation of constructs

Constructs were generated for use in cell culture as well as for use in *Xenopus laevis* embryos. Amino acid residues are numbered from the start of translation (1) to the stop codon (493), as shown in figure 2.1.

Figure 2.1 Amino acid sequence of *Xenopus laevis* MI-ER1.

The four putative NLS are indicated by solid lines, the dashed line indicates NLS4, including the regions flanking the basic core of the NLS.

1 MAEPSLRTAS PGGSAASDDH EFEPSADMLV HEFDDEQTLE EEEMLEGEVN

51 FTSEIEHLER ESEMPIDELL RLYGYGSTVP LPGEEDEEDM DNDCNSGCSG

101 EIKDEAIKDS SGQEDETQSS NDDPTPSFTC RDVREVIRPR RCKYFDTNHE

151 IEEESEDDED YVPSEDWKKE IMVGSMFQAE IPVGICKYRE TEKVYENDDQ

201 LLWNPEYVME ERVIDFLNEA SRRTCEERGL DAIPEGSHIK DNEQALYEHV

251 KCNFDTEEAL RRLRFNVKAA REELSVWTEE ECRNFEQGLK AYGKDFHLIQ

301 ANKVRTRSVG ECVAFYYMWK KSERYDFFAQ QTRFGKKKYN LHPGVTDYMD

351 RLLDESESAT SSRAPSPPPT TSNSNTSQSE KEDCTASNNT QNGVSVNGPC

401 AITAYKDEAK QGVHLNGPTI SSSDPSSNET DTNGYNRENV TDDSRFSHTS

451 GKTDTNPDDT NERPIKRQRM DSPGKESTGS SEFSQEVFSH GEV*

1 = NLS basic region
2 = NLS basic region
3 = NLS basic region
4 = NLS4 + flanking regions

Table 2.1 Amino Acid sequences of all constructs used.

Construct name	Amino acids	Sequence
EGFP	N/A	
EGFP-ER1	1-493	MAEP...GEV*
EGFP-ER1-143	1-143	MAEP...RRCK
EGFP-ER1-137	1-137	MAEP...REVI
EGFP-ER145-250	145-250	FDTN...YEHV
EGFP-ER251-319	251-319	KCNF...YYMW
EGFP-ER251-340	251-340	KNCF...KKYN
EGFP-ER339-493	339-493	YNLH...GEV*
EGFP-ER1-461	1-461	MAEP...DDTN
EGFP-ER1-473	1-473	MAEP...MDSP
EGFP-ΔNLS4	Δ464-467	ΔPIKR
EGFP-NLS1	136-143	VIRPRRCK
EGFP-NLS2	259-273	ALRRLRFNVKAAREE
EGFP-NLS3	317-341	MWKKSERYDFFAQQTRFGKKKYN
EGFP-NLS4a or EGFP-ER461-471	461-471	NERPIKRQRMD
EGFP-NLS4b or EGFP-ER457-475	457-475	PDDTNERPIKRQRMDSPGK
EGFP-ΔNLS1	Δ138-141	ΔRPRR

EGFP-ΔNLS3	Δ320-338	ΔKKSERYDFFAQQTRFGKKK
EGFP-ΔNLS1+ ΔNLS3	Δ138-141 + ^{320}KK = AA + Δ336-338	ΔRPRR + ^{320}KK + ΔKKK
EGFP-ER1 SPGK → APGK	1-493, ^{472}A	...RMD**A**PGK...
EGFP-ER1 SPGK → NPGK	1-493, ^{472}N	...RMD**N**PGK...
EGFP-ER1 SPGK → DPGK	1-493, ^{472}D	...RMD**D**PGK...
EGFP-ER339-493 SPGK → APGK	339-493, ^{472}A	...RMD**A**PGK...
EGFP-ER339-493 SPGK → NPGK	339-493, ^{472}N	...RMD**N**PGK...
EGFP-ER339-493 SPGK → DPGK	339-493, ^{472}D	...RMD**D**PGK...
EGFP-ER457-475 SPGK → APGK	457-475, ^{472}A	...RMD**A**PGK...
EGFP-ER457-475 SPGK → DPGK	457-475, ^{472}D	...RMD**D**PGK...
β-GAL or c-βGAL	N/A	

nβ-GAL/ or β-GAL-NLS	SV40 NLS	PKKKRKV
β-GAL-XNF7NLS	XNF7 NLS	KTPQKRKIEEPEPEKKAK
β-GAL-ER1	1-493	MAEP...GEV*
β-GAL-ER1-137	1-137	MAEP...REVI
β-GAL-ER1-143	1-143	MAEP...RRCK
β-GAL-ER145-250	145-250	FDTN...YEHV
β-GAL-ER251-340	251-340	KNCF...KKYN
β-GAL-ER339-493	339-493	YNLH...GEV*
β-GAL-ER461-471	461-471	NERPIKRQRMD
β-GAL-ER457-475	457-475	PDDTNERPIKRQRMDSPGK
β-GAL-ΔNLS4	Δ464-467	ΔPIKR
β-GAL-XNF7NLS-ER1-282	1-282	MAEP...EEEC
β-GAL-XNF7NLS-ER267-493	267-493	VKAA...GEV*

2.1.1 Generation of EGFP-ER1 fusion constructs

2.1.1.1. PCR

For all constructs, full length *Xenopus laevis er*1 cDNA in PCR2.1 was used (described in (Paterno *et al.*, 1997). For the generation of EGFP-ER1 fusions, different regions of *er*1 were amplified by PCR with *er*1 specific primers, see Appendix B: EGFP-ER1 1-493: er14 and er13; EGFP-ER1 1-130: er14 and er26; EGFP-ER1 1-144: er14 and er28; EGFP-ER1 145-250: er29 and er30; EGFP-ER1 251-320: er31 and er25; EGFP-ER1 251-340: er31 and er32; EGFP-ER1 339-493: er27 and er13; EGFP-ER1 1-461: er14 and er34; EGFP-ER1 1-473: er14 and er35; EGFP-ER1 339-473: er27 and er35; EGFP-NLS-3: er24 and er25; EGFP-ER1 457-475 (NLS4 b): er38 and er39; EGFP-ER1 457-493: er38 and er13; EGFP-ER1 446-493: er50 and er13. The primers contained *Bgl II* (agatct) or *Sal I* (gtcgac) specific regions to facilitate cloning (underlined in primer sequence). The PCR was performed with *Pfu Taq* (Promega Co.) for accurate amplification. The ΔNLS1 and ΔNLS3 were generated by three separate PCRs, as follows: 2 complementary primers were designed to contain the deletion and span the flanking regions. The first PCR was performed with primer Xen1as (ΔRPRR) or NLS2as (Δbipartite) and primer er14. The second PCR was performed with primer Xen1s (ΔRPRR) or NLS2s (Δbipartite) and primer er13. The PCR products for ΔRPRR or Δbipartite were then combined and heated to 95°C for 10 minutes and then slowly cooled to RT. A third PCR was then

performed with primers er13 and er14, creating a full length *er*1, lacking RPRR or KKSERYDFFAQQTRFGKKK region, which was then cloned into PCR2.1 (TA cloning, Invitrogen). After plasmid extraction (miniprep Wizard SV, Promega corporation), the DNA was digested with *Bgl* II (Invitrogen), and the insert was ligated into EGFP-C2 (Clontech laboratories Inc.) and sequenced (Sequenase, version 2, United States Biochemical).

2.1.1.2 Ligation of PCR products into PCR 2.1

The PCR products were treated with *Taq* polymerase to create a TA-overhang (Invitrogen Laboratories Co.) and ligated into PCR 2.1 according to the manufacturer's protocol (TA cloning or TOPO cloning, Invitrogen Laboratories Co.). After transformation, a single colony was picked and grown in LB containing 100 μg/ ml ampicillin o/n at 37° C at 225 rpm. The plasmid DNA was then isolated as described in 2.1.1.6.

2.1.1.3 Digestion and purification of inserts from PCR2.1

PCR2.1 containing the desired insert were digested with restriction enzymes (Invitrogen Laboratories Co.) of which the recognition sites were incorporated in the primers. The digested inserts were gel purified from a 1% agarose gel in modified TAE (40 mM Tris-acetate pH8.0, 0.1 mM Na_2EDTA) using Ultrafree DA filters (Millipore corporation).

2.1.1.4 Generation of small fusion constructs.

For the constructs EGFP-NLS1, EGFP-NLS2 and EGFP-NLS4a, two complementary oligonucleotides (Oligo's Etc.) of the region of the putative NLS were annealed by heating the complementary oligonucleotides (see below) together to 95°C and slowly cooling down to RT. EGFP-NLS1: er23K and er33; EGFP-NLS2: er36s and er36a; EGFP-NLS4: er37s and er37as. After cooling down, the strands were treated with T4 DNA Kinase (Invitrogen Inc.) according to the manufacturers protocol. The enzyme was then removed from the reaction mixture with the enzyme removal filters (Millipore Corporation). The DNA was ligated into Egfp-C2 (Clontech Laboratories Inc.) and transformed into *XL1-Blue* cells.

2.1.1.5. Ligation of purified inserts into EGFP-C2.

After purification, the inserts were directly ligated into the *Bgl II* site of the EGFP-C2 expression vector (Clontech Laboratories Inc.), using T4 DNA ligase according to the manufacturers protocol (Invitrogen Inc.). After transformation into chemically competent *XL1-Blue* cells, and selection on LB plates containing 30 μg/ml Kanamycin, the colonies were screened by PCR with one *mi-er*1 specific 3' primer and a 5' vector primer PCZ-5'.

2.1.1.6. Plasmid isolation and purification

For transfection, transformed *XL1-Blue* cells were grown o/n in 200 ml LB containing 30 μg/ml Kanamycin or 50 mg/ml ampicillin and plasmid DNA was isolated by large or midi scale plasmid preparation kits (maxi-prep: Concert, Invitrogen, midi-prep: high speed plasmid preparation, Qiagen Inc.). Before use, all constructs were sequenced (Sequenase Version 2.0, United States Biochemical), according to the manufacturer's protocol.

2.1.2 Deletion or substitution of single or multiple amino acids

For the generation of constructs that only differed from the wild type sequence by 1-4 amino acids, the Quick-Change mutagenesis™ kit was used (Stratagene Inc.), according to the manufacturer's protocol, with the following primers: for the cdc2/ PKA phosphorylation site in ER1-493 and ER339-493: S → N: er40K and er40R; S→A: er41K and er41R; S→ D: er53K and er53R; in ER457-475: S→A: er45K and er45R; S → D: er46K and er46R. The ΔPIKR constructs were generated using the primers er51K and er51R.

2.1.3 Generation of gfp-er1 fusions into pT7Ts

Primers were generated against the EGFP-C2 vector, a sense primer against the start of the *egfp* gene and an anti-sense primer against the stop-codon of the *egfp*-C_2 plasmid which is located after the MCS, thus a PCR product

could be generated spanning the multiple cloning site containing the different *Xer*1 inserts. These primers had *Xbal* restriction recognition sites (tctaga) and were used to subclone the egfp-fusion constructs into the *Xbal* site of the pT7Ts vector (Cleaver *et al.*, 1996). PCR were performed using Pfu Taq (Promega Corporation) and PCR products were TA or TOPO cloned (Invitrogen) before sub cloning into pT7Ts, as described before for the egfp-c2 constructs.

2.1.4 Generation of β-galactosidase-er1 fusion constructs

Inserts for the β-galactosidase-CS2+ vector (Rupp *et al.*, 1994; Turner and Weintraub, 1994) were subcloned by PCR from the egfp-c2 constructs using a sense primer against the MCS (pEGFP-5') 5' of the insert and gfp-as. This was done to create the right reading frame for fusion with β-galactosidase. The PCR products were treated as described above and ligated into the *Xbal* site of the β-gal-CS2+ vector.

β-gal-CS2$^+$-Xnf7NLS was generated by annealing of primers Xnf7s and Xnf7as (Li *et al.*, 1994a), see Appendix B for sequence) after which the dsDNA strands were treated with T4 DNA Kinase (Invitrogen Inc.) according to the manufacturers protocol. The enzyme was then removed from the reaction mixture with the enzyme removal filters (Millipore Corporation). The DNA was ligated into the *Xhol* site of the β-gal-CS2$^+$ vector.

β-gal-xnf7NLS-ER1-282 and β-gal-xnf7NLS-ER267-493 were generated by PCR using primers er54, er55, er56 and er57 and full length *Xenopus laevis*

*er*1 cDNA in PCR2.1 as a template. PCR was performed as described before. After digestion of the PCR products with *XbaI*, the inserts were ligated into the *XbaI* site of β-gal-CS2$^+$ using T4 DNA ligase as described before.

2.1.5 Sequencing

All constructs were sequenced using Sequenase Version 2.0, (United States Biochemical), according to the manufacturers protocol.

2.2 Cell culture

2.2.1 Materials

Most of the materials used for cell culture were supplied by Invitrogen, including DMEM, Penicillin/ Streptomycin, Trypsin, Calf serum, Fetal Bovine serum, and cell culture plates. NaCl (BDH), KCl, KH_2PO_4, $Na_2HPO_4 7H_2O$, $NaHPO_4$, EDTA and Paraformaldehyde were supplied by Fisher Scientific.

2.2.2 Methods

NIH 3T3 cells (American type Culture Collection, ATCC, USA) were grown in DMEM containing 3.7 g/l $NaHCO_3$, 7.5 % Calf serum, 2.5% Fetal Bovine serum and antibiotics 50 U Penicillin, 50µg/ml Streptomycin; at 37 °C, 10% CO_2. Cells were trypsinized in 0.025% trypsin in PBS (137 mM NaCl, 2.7 mM KCl, 1.4 mM KH_2PO_4, 8.1 mM $Na_2HPO_4 \cdot 7H_2O$ pH 7.4 with HCl)/ 1mM EDTA every 3 days

and diluted 1:5 into a fresh 100 mm tissue culture dish. Stocks of cells were frozen at -70°C in Calf serum containing 10% DMSO. $1x10^6$ cells were frozen in a freezing vial (Nalge Nunc International) at a rate of 1°C/ minute in a 2-propanol container. When thawing the cells, the vial was submersed in 37°C water to quickly thaw the cells. The cells were washed in 10 ml DMEM and plated into a 100 mm cell culture dish containing DMEM and 10% serum, as described above.

2.3 Transient transfection of NIH 3T3 cells

Cells were plated at a density of $2x10^5$ cells/well in a 6 well tissue culture dish and left to grow o/n in DMEM containing 10% serum and antibiotics (see under cell culture). The next day, cells were transfected with 0.5-2 µg plasmid DNA. For transfection, LipofectAmine plus (Invitrogen Inc.) was used according to the manufacturer's protocol with the following modifications: 0.5- 2 µg plasmid, 6 µl Plus reagent and 6 µl LipofectAmine were used. Twenty-four hours after transfection cells were trypsinized with 0.025% trypsin (Invitrogen) in PBS/EDTA and plated into a 6 well-plate for analysis of the overexpressed protein and on to chamber slides (Labtek, Nalge Nunc International) for microscopic examination. Forty-eight hours post transfection the cells were lysed for protein assays (see section 2.3.2). Cells grown on the chamber slides were fixed with 4% Paraformaldehyde in PBS (pH7.2) and treated as described below.

All analysis of transiently expressed proteins was performed 48 hrs post-transfection because the highest level of protein expression was detected at this

time, using fluorescence microscopy and Western Blot analysis (as described in section 2.4).

2.3.1 Microscopic examination of transfected cells

2.3.1.1 Detection of fluorescent proteins

Cells transfected with EGFP-C2 vectors, expressing enhanced green fluorescent protein fused to the protein of interest, were washed twice in PBS and fixed in 4% Paraformaldehyde for 20 minutes at RT. After fixation, the cells were washed twice in PBS and were stained by incubation with 10 ng/ml Hoechst 33342 in PBS for 10 min. After washing in PBS, the cells were mounted in 10% glycerol in PBS, viewed by combined fluorescence/ brightfield microscopy on an Olympus BH2 microscope using DplanApo/UV objectives and the images were captured with a Coolsnap digital camera. Localization of the nuclei within the cells was determined by staining of the DNA with Hoechst 33342, which resulted in a blue stain when viewed by fluorescence microscopy using a DAPI filter (406 nm). Hoechst 33342 binds AT pairs in the DNA and can enter cells that are fixed as well as live (Holmquist, 1975; Shapiro, 1981).

Cells were scored as either N (exclusive nuclear staining), N>C (nuclear fluorescence was greater in intensity than cytoplasmic), or WC (whole cell fluorescence, in which nuclear fluorescence was equal in intensity to cytoplasmic staining). No nuclear exclusion of the overexpressed protein was observed after transfection of these constructs, since EGFP was able to diffuse into the nucleus.

All transfections were performed in NIH 3T3 cells. EGFP-fusions generally displayed transfection efficiencies between 10 and 40 %, meaning that 10-40% of the total cells expressed EGFP. Cells transfected with β-gal fusions displayed transfection efficiencies between 50 and 70%. For each construct that was transfected at least 500 intact cells containing transient protein were counted. Cells that showed abnormal morphology and cells where the nuclei could not be distinguished were excluded from the count.

2.3.1.2. Immunocytochemistry

Transfected cells were grown on chamberslides as described above and MI-ER1 protein was detected by immunocytochemistry using an anti-rabbit detection system (Santa Cruz Biotechnology ABC staining system) using an anti-ER1 antibody. The anti-MI-ER1 antibody was generated by injection of the peptide ^{464}PIKRQRMDSPGKEST^{478} into rabbits. The polyclonal antibody was used directly from the diluted serum of the rabbits. The staining procedure was as follows: after fixation in 4% Paraformaldehyde in PBS, the slides were washed in PBS, then incubated for 10 min in 0.2% Triton X-100 in PBS. The slides were then incubated for 20 min in 1.5% normal goat serum in PBS (normal blocking serum) and then directly transferred to a solution containing the primary antibody (against the C-terminus of XER1, see table in section 2.4; a 1:1000 dilution) in normal blocking serum and incubated at RT for 2 hrs. The slides were then washed in 0.2% Triton X-100 in PBS twice for 10 min each. The slides were

incubated in normal blocking serum containing a secondary antibody (biotinylated goat-anti rabbit IgG; 1:200) for 30 min at RT followed by washing in 0.1% Triton X-100 in PBS for 3 min and PBS for 3 min. An avidin-biotin-HRP complex was added to the slides for 30 min followed by washing in PBS for 3 min. Detection was performed by DAB staining (Sigma-Aldrich) for 4-7 min or until color developed. The gasket was then removed from the slides and the slides were mounted in a PBS/ 10% glycerol solution.

2.3.1.3 X-gal staining of transfected NIH 3T3 cells

Cells were grown in chamberslides as described and 24-48 h post-transfection cells were washed in PBS and incubated for 15 min. in fresh 4% paraformaldehyde in PBS. After fixation the cells were washed for 10 min in wash buffer (PBS containing 5 mM EGTA, 2 mM $MgCl_2$, 0.01% (w/v) deoxycholate and 0.02% (v/v) Nonidet P-40) then placed in staining buffer (wash buffer containing 1 mg/ml X-gal (dissolved in DMSO), 10 mM potassium ferrocyanide, 10 mM potassium ferricyanide) and incubated at 37 °C for 30-60 min. or until color developed. After staining the cells were washed in PBS for 20 min. and then fixed in 4% Paraformaldehyde for 20 min. The cells were washed in PBS and then embedded in 10% (v/v) glycerol in PBS and covered with a large coverslip (24x60 mm) before viewing under the light microscope.

2.3.2 Cell lysis for direct Western Blot

Transfected NIH 3T3 cells were plated at $3x10^5$ cells/ well and left to grow o/n. After washing in cold PBS on ice, cells were then lysed in protein sample buffer (0.125 M Tris pH 6.8, 2% SDS, 5% β-mercaptoethanol, 2% glycerol) and kept on ice. A determination of the protein concentration was performed with the Bio-Rad protein assay kit, (Bio-Rad Laboratories). 100-µg total protein was separated on a 10% acrylamide gel (all components of the electrophoresis and transfer apparatus (systems Protean II and III) were purchased from Bio-Rad Laboratories). The separated protein was transferred by electro-blotting on to a nitrocellulose membrane (Hybond™ ECL™, Amersham Pharmacia Biotech).

2.3.3 Immunoprecipitation of proteins

Transfected cells were plated at $3x10^5$ cells/well in a 6-well plate and allowed to grow o/n at 37°C in a 10% CO_2 incubator. After washing the cells in cold PBS the cells were lysed in cold Triton buffer (10 mM Tris pH 7.5, 1% Triton X-100, 10 mM Na_2EDTA, 0.002% sodium-azide) and extracts were left on ice for 20 minutes. The soluble proteins were separated from the insoluble proteins by centrifugation at 12,000xg in a mini centrifuge (Eppendorf). The soluble proteins (supernatant) were transferred to a fresh eppendorf tube and an antibody was added to precipitate a specific protein. The antibody-protein mixture was incubated at 4°C for 1hr to O/N (depending on the antibody) with gentle rotation. After this incubation, 50 µl Protein-A-Sepharose (1 mg/ml, Amersham Pharmacia

Biotech) was added to each immunoprecipitation mixture to bind the antibody and the mixture was incubated at 4°C for 1 hr with gentle rotation. The protein-A-sepharose-antibody-protein complex was washed 4x with Triton-buffer and 4x with 150mM NaCl. After washing, the protein was released from the protein-A-sepharose-antibody complex by adding 30µl protein sample buffer (0.125 M Tris pH 6.8, 2% SDS, 5% β-mercaptoethanol, 2% glycerol) and boiling for 4 minutes. The protein was separated on a 10% acrylamide gel and transferred to a nitrocellulose membrane (Hybond ECL, Amersham Pharmacia) by electro-blotting.

2.4 Western Blot Analysis

2.4.1 Materials

The nitrocellulose membranes were purchased from Amersham Pharmacia Biotech (Nitrocellulose Hybond ECL). Prestained high and low molecular weight markers, protein-concentration estimation dye and polyacrylamide (30%) were purchased from Bio-Rad Laboratories. Detection was performed with chemiluminescence using FemtoLucent™ (Chemicon International Inc.). ECL™-hyperfilm (Amersham Pharmacia Biotech) was used for the detection.

Table 2.2 Antibodies used for immunocytochemistry, immunoprecipitation and Western blot analysis

Name	Epitope	Source	Concentration	Incubation
Rabbit-anti MI-ER1	C-terminus MI-ER1 PIKRQRMDSPGKEST	Paterno/ Gillespie	1:3000	1 hr
Anti-GFP	GFP	Clontech laboratories Inc.	1:1500	O/N
Goat anti Rabbit IgG-HRP	For use with anti-MI-ER1 Ab	Amersham Pharmacia Biotech	1:3000	1 hr
Sheep anti Mouse IgG-HRP	For use with GFP Ab	Amersham Pharmacia Biotech	1:1000	1 hr

2.4.2 Methods

After transfer of the proteins to the nitrocellulose membrane, the membrane was incubated in 5% blocking powder (Chemicon International Inc.) in TBS-T (20 mM Tris, 137 mM NaCl, 1% (v/v) Tween-20, pH 7.6) for 30 min. and then washed in large volumes of TBS-T for 30 min. The membrane was then incubated in TBS-T containing a primary antibody (see table 2.2 for list of antibodies used, concentration and incubation time for each antibody). After the first antibody incubation the membrane was washed in large volumes of TBS-T for 1 hr and then the secondary Ab (see table 2.2) was added in TBS-T and incubated for 1 hr, after which the membrane was washed in TBS-T before detection of the protein with the FemtoLucent™ detection system and exposure on ECL™-Hyperfilm™ (Amersham Pharmacia).

2.5 Fertilization of *Xenopus laevis* eggs

2.5.1 Materials

Xenopus laevis were purchased from Nasco International Inc.. Human Chorionic Gonadotrophin hormone and cysteine were supplied by Sigma-Aldrich. NaCl (BDH), KCl, $Ca(NO_3)_2 4H_2O$, $MgSO_4$, $Na_2HPO_4 7H_2O$, EGTA, EDTA, HEPES were supplied by Fisher Scientific.

NAM (normal amphibian medium; Slack and Forman, 1980) consisted of 110 mM NaCl, 2 mM KCl, 1 mM $Ca(NO_3)_2 4H_2O$, 1 mM $MgSO_4 7H_2O$, 0.1 mM Na_2EDTA, 100 mM HEPES, pH 7.5)

2.5.2 Methods

Mature eggs were isolated and artificially fertilized as per Godsave et al. (Godsave *et al.*, 1988). Approximately 16 hours before eggs were needed, 500 I.U. of HCG hormone was injected subcutaneously into the upper hind leg of a female *Xenopus laevis*, just above the cloaca. Within 12-18 hours, the female began to lay eggs, which were collected on a plastic petri dish.

Eggs were then fertilized with sperm extracted from a male testis. A male frog was sacrificed; testes were removed and stored at 4 °C in a 1x NAM solution (11 mM NaCl, 0.2 mM KCl, 0.1 mM $Ca(NO_3)_2 4H_2O$, 0.1 mM $MgSO_4 7H_2O$, 0.01 mM Na_2EDTA, 10 mM HEPES, pH 7.5, 1 mM $NaHCO_3$, 0.25% gentamycin). Approximately 100 µl of macerated testis were mixed with a large amount of sterile water and added to the eggs. After a period of 10 minutes to allow for

fertilization, the eggs were flooded in water. A visible rotation of zygotes to animal (dark) side up indicated that the eggs had been fertilized.

Approximately 20 minutes after flooding, the eggs were dejellied in a solution of 2% cysteine, pH 7.8-8.1. The eggs were swirled in the solution until they visibly touched and then rinsed thoroughly with 1 l of deionized water. Water was removed and the eggs were submerged in NAM/20 (5.5 mM NaCl, 0.1 mM KCl, 50 μM $Ca(NO_3)_2 4H_2O$, 50 μM $MgSO_4 7H_2O$, 5 μM Na_2EDTA, 5 mM HEPES, pH 7.5, 0.25% gentamycin).

2.6 Micro-injection of RNA into *Xenopus laevis* embryos

2.6.1 Preparation of RNA for injection

EGFP-ER1 fusions were subcloned into pT_7Ts (see paragraph: 2.1.3) and were transcribed using the Ribomax RNA transcription kit (Promega corporation) according to the manufacturers protocol. RNA cap ($m^7G(5')ppp(5')G$; New England Biolabs) was added to the transcription reaction at a concentration of 100 nM. After transcription the RNA was purified by 2 rounds of precipitation, once in sodium acetate (CH_3CO_2Na) and ethanol and once in ammonium acetate ($CH_3CO_2NH_4$) and Ethanol at –70 °C overnight followed by washing of the precipitated RNA with 70 % ethanol in DEPC (diethyl pyrocarbonate) treated water. The RNA was then resuspended in DEPC- H_2O and the RNA concentration was determined by spectrophotometry at 260/280 nm. The RNA

was then diluted to 10 ng in 4.6 nl, unless otherwise noted. This could be directly used for injection of *Xenopus* embryos.

RNAs were translated *in vitro* using TNT® T7 Coupled Reticulocyte Lysate according to the manufacturer's protocol, and ^{35}S-methionine (Mandel). Translation efficiencies and size of proteins were determined by precipitation and SDS-PAGE respectively.

2.6.2 Injection of RNA into *Xenopus* embryos

After dejellying, the embryos were placed in a solution of NAM/2 (5.5 mM NaCl, 0.1 mM KCl, 0.05 mM $Ca(NO_3)_2 \cdot 4H_2O$, 0.05 mM $MgSO_4 \cdot 7H_2O$, 5 μM Na_2EDTA, 5 mM HEPES, pH 7.5, 1 mM $NaHCO_3$, 0.25% Gentamycin), containing 4% Ficoll (Amersham Pharmacia Biotech) in a petri dish containing a plastic grid. The grid provides support to the embryos at four sides and separates individual embryos, so embryos can be injected with RNA. The Ficoll serves to increase the viscosity and osmotic value of the solution so that the vitelline membrane will tightly surround the embryo. For injection a micro-injector apparatus was used with a beveled glass capillary containing the RNA to be injected. A total volume of 4.6 nl RNA was injected per embryo and the embryos were left to develop in 4% Ficoll in NAM/2 to the desired stage. Staging of the embryos was done according to the table of Nieuwkoop and Faber (1967).

During development, the translated EGFP-fusions were visualized by fluorescence microscopy using a 495 nm filter. Embryos injected with β-galactosidase fusions were stained with X-gal.

2.6.3 Staging of embryos according to Nieuwkoop and Faber

Identification of the different stages of *Xenopus* development was performed morphologically according to the table by Nieuwkoop and Faber (Nieuwkoop and Faber, 1967). Identification of stages by morphology might not be the best way of determining the exact stage because of the differences in cell size between embryos of different frogs, and the subtle morphological differences. Also, there might be slight variations in determining the stages of the embryos, depending on the researcher. Besides morphological differences we also use timing after fertilization as a measure of stage, but this is variable because not all eggs are fertilized at the exact same time. In addition small differences occur in rate of development between embryos. Fig 1.1.1 shows morphology of the embryos used in the experiments.

2.6.4 Detection of proteins in microinjected embryos

2.6.4.1 Detection by fluorescence microscopy

Embryos were fixed in MEMFA (0.1 M MOPS, pH7.4, 2mM EGTA pH8.0, 1 mM $MgSO_4$, 3.7 %(w/v) formaldehyde) for at least 1 hr and maximum O/N and then washed in PBS-A (96 mM NaCl, 1.9 mM KCl, 1 mM KH_2PO_4, 5.7 mM

$Na_2HPO_4 \cdot 7H_2O$) and stored in PBS-A/ 0.02% (w/v) sodium azide. The embryos were then directly viewed by fluorescence microscopy on an Olympus BH2 microscope using Dplan Apo/UV objectives (filter I.B. 495 nm)

2.6.4.2 Detection by whole mount staining

Injected albino embryos were staged by comparison with wild-type embryos that were fertilized in the same dish. Embryos were fixed at the desired stage by submersion in MEMFA for 1 hr. and then stored in 100% methanol at –20°C. Before staining embryos, the embryos were incubated in bleaching solution (1 part 30% hydrogen peroxide and 2 parts Dent's fixative: 20% DMSO, 80% methanol (Dent *et al.*, 1989)) for 1-2 hr under white light. The embryos were then washed in maleic acid buffer (MAB; 100 mM maleic acid, 150 mM NaCl pH 7.5 with 10 N NaOH). Embryos were incubated for 1 hr in blocking solution to eliminate background staining (MAB, 2% (w/v) Blocking powder for use with hybridization of nucleic acids (Roche), 5% (v/v) DMSO, 5% (v/v) heat inactivated goat serum, $1x10^{-3}$ % (v/v)/ 24.7 μM thimerosal (=ethylmercurithiosalicylic acid sodium salt). The primary antibody (anti-gfp monoclonal antibody, generated in mouse cells, Clontech laboratories Inc.) was diluted to 1:600 in fresh blocking buffer and the embryos were incubated o/n at 4° C while shaking. After this incubation, the embryos were washed in large volumes of MAB over a period of 5-6 hr while shaking at RT. The MAB was replaced every 30-60 min. The secondary antibody (alkaline phosphatase goat-anti-mouse IgG) was diluted

1:600 in fresh blocking buffer and the embryos were incubated at 4°C o/n. The washing was repeated as before over 6 hr. Detection was performed in alkaline phosphatase (AP) buffer (100 mM Tris pH 9.5, 50 mM $MgCl_2$, 100 mM NaCl, 0.1 % (v/v) Tween-20, 50 mM levamisole) as follows: The embryos were washed in AP buffer twice for 5-10 min. and then incubated in NBT/BCIP in AP buffer (35 µl NBT and 45 µl BCIP, both from Boehringer Mannheim) until color developed. The color reaction was stopped by addition of MEMFA. For long time storage embryos were kept in MAB+ 0.02% (w/v) (3mM) sodium-azide (NaN_3).

2.6.4.3 X-gal staining of whole embryos

Embryos were incubated at room temperature until the desired stage was reached, collected in ice-cold PBS-A and fixed in 4% Paraformaldehyde in PBS-A for 1 hr. After fixation the embryos were washed three times for 15 min each in wash buffer and incubated in staining buffer for 30-60 min. or until color developed (for recipe see X-gal staining of transfected NIH 3T3 cells, paragraph 2.3.1.3). The embryos were then washed in PBS-A for 10 min and incubated 1hr in 4% Paraformaldehyde in PBS to prevent diffusion of the staining. For long-term storage, the embryos were kept in PBS-A + 0.02% (w/v) sodium azide at 4°C.

Chapter 3

The identification of nuclear localization signals in XMI-ER1*

3.1 Introduction

The subcellular localization of XMI-ER1 during *Xenopus laevis* embryonic development has been well investigated (Luchman *et al.*, 1999). XMI-ER1 is expressed during early stages of development, but only localizes to the nucleus during the mid-blastula stages (Luchman *et al.*, 1999). When transfected into NIH 3T3 cells, translated XMI-ER1 is exclusively targeted to the nuclei of these cells indicating that in these cells XMI-ER1 is constitutively localized to the nucleus (Paterno *et al.*, 1997).

XMI-ER1 is a 55kDa protein which is predicted to contain two putative NLS (Paterno *et al.*, 1997). Nuclear localization of proteins with a molecular weight greater than 40-60 kDa is dependent on the presence of a functional NLS. The identification of a functional NLS is based on two criteria:

1) Mutation or deletion of the NLS leads to cytoplasmic localization of the protein.

2) The NLS is active in nuclear targeting of a normally cytoplasmic protein, either as a peptide covalently coupled to the carrier or when encoded in the same reading frame as a fusion protein (Jans and Hubner, 1996).

In this chapter, four putative NLS in XMI-ER1 are analyzed for their ability to direct the nuclear localization of a cytoplasmic protein. A functional NLS was

* Part of the results discussed in this chapter have been published: (Post *et al.*, 2001).

identified in the C-terminal region of XMI-ER1. It was shown that, in addition to the basic core NLS, flanking regions were necessary for its efficiency in translocating a non-nuclear protein to the nucleus. An additional, albeit weak, NLS was identified in the N-terminus. The efficiency of these NLS when expressed in mouse fibroblast cells was similar to their efficiency in *Xenopus* cells.

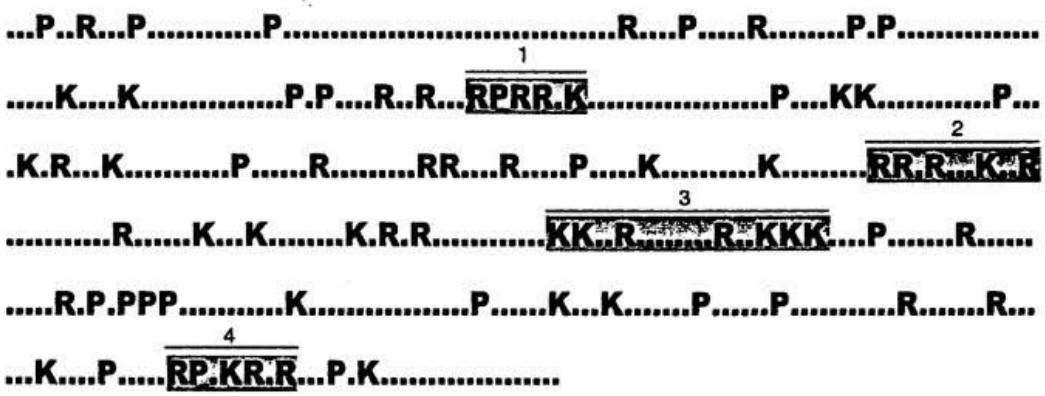
3.2 Sequence analysis of basic regions in XMI-ER1.

Previously, two putative NLS were identified in XMI-ER1 (Paterno *et al.*, 1997) by computer analysis using PSORT (Nakai and Horton, 1999). However, Schwamborn *et al.* (1998) showed that some NLS are not identified by computer analysis but rather by identification of stretches of basic amino acids (Schwamborn *et al.*, 1998). Therefore, I analyzed XMI-ER1 for existence of additional putative NLSs. By analyzing the basic amino acids in XMI-ER1 I found 2 additional putative NLSs (see below).

To determine the regions important for the nuclear translocation of XMI-ER1, a map of XMI-ER1 was assembled in which the basic amino acids (which can bind the import receptors) and prolines (which are often found in the regions flanking the core NLS) were indicated (Figure 3.1). These regions were designated NLS 1, 2, 3 and 4 in the order in which they appeared in the protein (Figure 3.1).

Figure 3.1 Identification of potential NLS in XMI-ER1. Schematic of the XMI-ER1 protein showing all prolines and basic residues within the sequence, with the remaining residues as dots. Four clusters of basic amino acids that have the potential to translocate MI-ER1 to the nucleus were identified (shaded boxes) and called NLS1-4.

NLS1= localized between residues 138 and 143; NLS2= 261-271; NLS3= 320-338; NLS4= 463-472. (Numbers indicate number of amino acid residue as shown in figure 2.1.)


...P..R...P.....P.....R....P.....R.....P.P.....
1
.....K....K.....P.P....R..R...RPRR.K.....P....KK.....P...
2
.K.R...K.....P.....R.....RR....R.....P.....K.....K.....RR.R...K..R
3
.....R.....K...K.....K.R.R.....KK..R.....R..KKK....P.....R.....
.....R.P.PPP.....K.....P.....K...K.....P.....P.....R.....R...
4
...K....P.....RP.KR.R...P.K.....

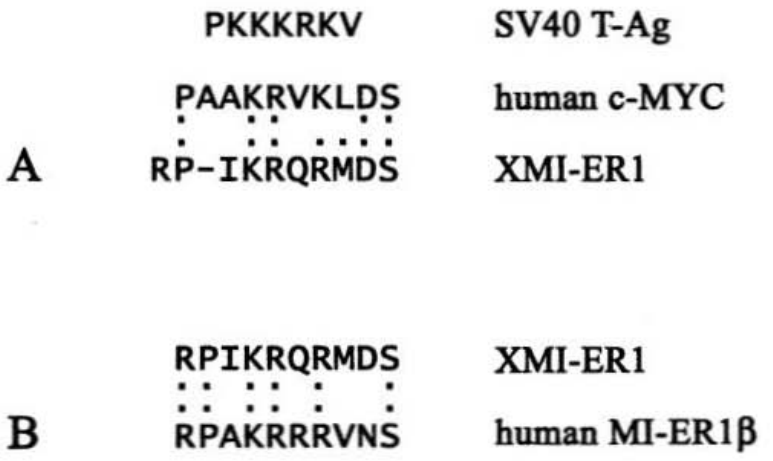
NLS1 and 2 are similar to NLS of the classic monopartite type (Figure 3.2 A). NLS3 consists of two regions of basic amino acids forming a potential bipartite NLS. NLS4 is similar to the human c-MYC NLS, in which only 3 basic amino acids in the core and a proline and acidic residues in the flanking region are present (see figure 3.2 A).

Analysis of the sequences of the putative NLS found in XMI-ER1 shows that NLS1 (^{138}RPRRCK^{143}) is flanked by bulky tyrosine and phenylalanine residues, predicted to interfere with the recognition of the NLS by importins (Boulikas, 1997), and contains a proline within the basic cluster, which can change the conformation of this basic cluster and thus interfere with the interaction with the importins. The NLS2 core cluster consists of five positively charged residues in a cluster of 11 amino acids (^{261}RRLRFNVKAAR^{271}), which contains bulky residues, however no helix breakers like glycine or prolines are present in the flanking region making it less likely to be a functional NLS. NLS3 is similar to the bipartite type of NLS (^{320}KKSERYDFFAQQTRFGKKK^{338}), but the spacer is longer than the usual 10 amino acid residues (14 amino acid residues) and contains tyrosine and phenylalanine residues, which are bulky. NLS4 (^{463}RPIKRQRMD^{471}) is most similar to the human c-MYC NLS in which the cluster of positively charged residues is flanked by a proline and an acidic residue (Makkerh *et al.*, 1996), making it the most likely candidate for a functional NLS, based on theoretical predictions.

The putative NLS4 in XMI-ER1 is not conserved between *Xenopus* and human (Figure 3.2 B). hMI-ER1β contains a classic type I NLS: ^{468}PAKRRR^{473} , and hMI-ER1α does not contain an NLS in its C-terminus. The putative NLS1 however (^{138}RPRRCK^{143}), is found in both hMI-ER1α and β as well as in XMI-ER1. Transfection of EGFP-hMI-ER1α and β into NIH 3T3 cells showed that while hMI-ER1β is found exclusively in the nuclei, hMI-ER1α is predominantly localized in the cytoplasm (data not shown).

Figure 3.2 Amino acid sequence comparison of XMI-ER1 NLS-4 to the NLS of the c-MYC protein, and the NLS in the human orthologue of XMI-ER1, hMI-ER1β.

Identical amino acids and conservative changes are indicated by double and single dots, respectively. A, The SV40 T-Ag NLS, which was the first NLS to be identified, is shown. B, Amino acid sequence comparison of the XMI-ER1 NLS and the hMI-ER1β NLS. While NLS4 of XMI-ER1 is similar to the human c-MYC NLS, the NLS of hMI-ER1β is a classic monopartite, similar to the SV40 T-Ag NLS.


PKKKRKV
SV40 T-Ag
PAAKRVKLDS
human c-MYC
A
RP-IKRQRMDS
XMI-ER1
RPIKRQRMDS
XMI-ER1
B
RPAKRRRVNS
human MI-ER1β

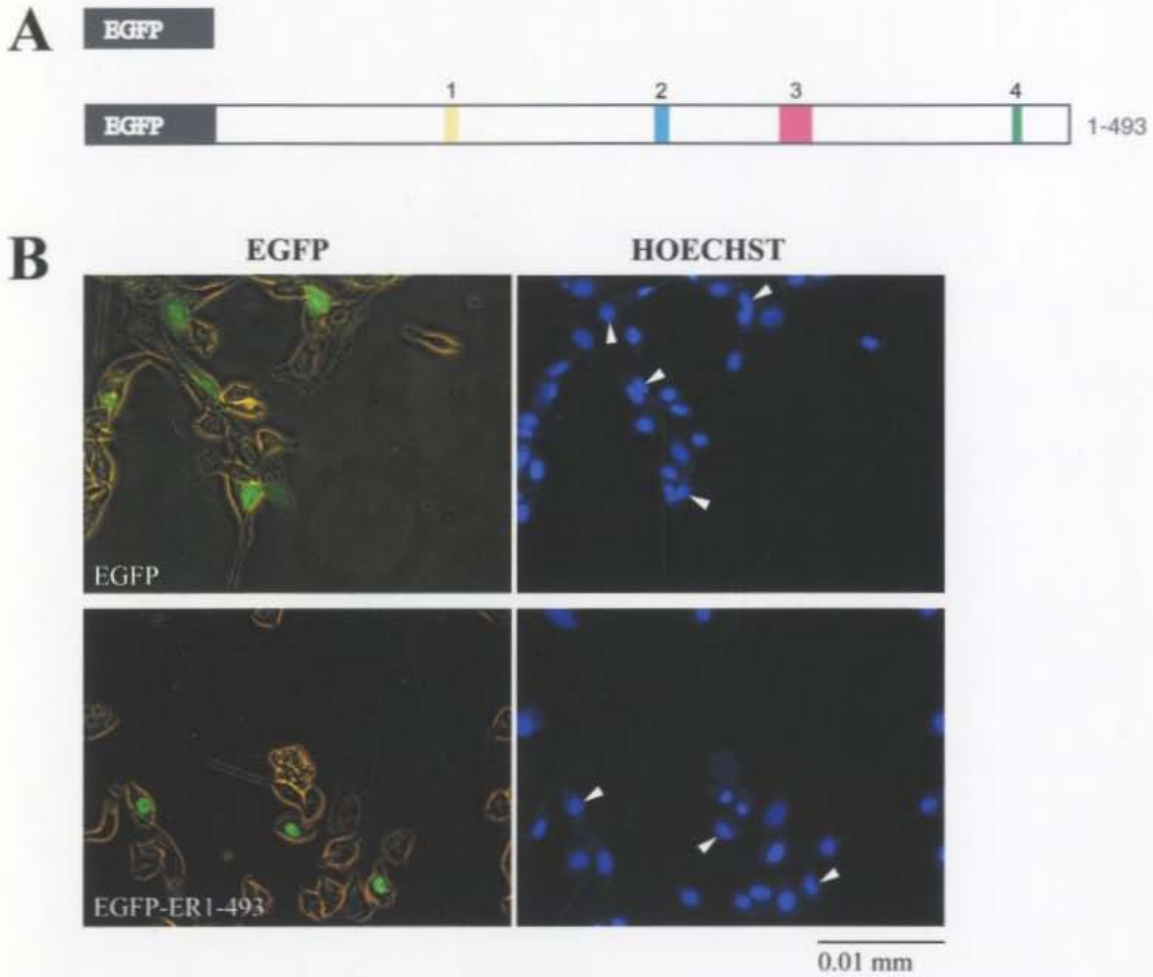
3.3 Identification of a functional NLS was performed using different fusion proteins

To reveal whether these regions of basic clusters represent functional NLSs, fusion constructs were generated in EGFP-C_2, creating fusion proteins in which EGFP is fused to the N-terminus of different regions of XMI-ER1. EGFP was chosen because of the ease of detection of fluorescent EGFP under U.V. light (495 nm). Fusion of EGFP to full-length XMI-ER1 resulted in exclusive nuclear localization, showing that EGFP did not interfere with the ability of XMI-ER1 to translocate the protein to the nucleus of NIH 3T3 cells (see figure 3.3 B). A disadvantage of using EGFP is that the protein is 29 kDa and that some EGFP fusions used were below the 40-60 kDa range, potentially allowing passive diffusion through the NPC. Cells transfected with EGFP-XMI-ER1 fusions were subjected to Western Blot analysis using an anti-GFP antibody to test the total level of transient protein expression (Figure 3.3C). This was done to test whether high protein levels interfere with the nuclear translocation. Full-length XMI-ER1, when fused to EGFP, (EGFP-ER1-493, lane 2) is not recognized well by the anti-GFP antibody, but western blot analysis using an anti-XMI-ER1 antibody shows expression levels similar to other EGFP-ER1 fusion proteins in the same assay (data not shown). Exclusive nuclear localization was observed even when the transient protein was extremely abundant as detected by Western Blot analysis, indicating that the level of transient protein expression did not interfere with the nuclear translocation, saturating the import machinery.

Figure 3.3 Fusion of EGFP to XMI-ER1 does not interfere with the nuclear translocation of XMI-ER1.

A, A construct was generated in which enhanced green fluorescent protein, EGFP (gray box), was fused to XMI-ER1 (white box). The colored bars indicate the location of the four putative NLS (1, 2, 3 and 4) in MI-ER1. B, Subcellular localization of EGFP and EGFP-ER1 was determined by fluorescence microscopy 48 hours after transfection (left panel). Nuclei were stained with Hoechst 33342 (right panel). Arrows indicate nuclei corresponding to EGFP expressing cells. C, cells were harvested 48 hours post-transfection and subjected to Western Blot analysis using an anti-GFP antibody to show the level of transiently expressed protein. Molecular weight markers are shown on the left. Arrows indicate the band corresponding to the protein of interest, EGFP-ER1, which runs around 120 kDa, and EGFP, which runs around 36 kDa.

Scale bar = 0.01 mm.


A
EGFP
1
2
3
4
EGFP
1-493
B
EGFP
HOECHST
EGFP
EGFP-ER1-493
0.01 mm
C
EGFP
EGFP-ER1-493
120
78
47
36

Different regions of XMI-ER1 were also fused to β-galactosidase, a 120 kDa protein, which is easily detectable by incubation of the cells with X-gal. This substrate for β-gal results in blue staining of β-gal expressing cells, within 15-30 minutes. An advantage of using β-gal is that diffusion into the nucleus is not possible, because of its size (fig 3.4 panel A). As well, when translocated into the nucleus, β-gal fusions are not able to leak out of the nucleus and unless a nuclear export signal (NES) is present, it will be localized in the nucleus during interphase.

Fusion of full-length XMI-ER1 to the C-terminus of β-gal resulted in nuclear localization of the protein, showing that β-gal does not interfere with the ability of XMI-ER1 to translocate to the nucleus (Figure 3.4 Panel C). When not fused to an NLS or NLS-containing protein, β-gal is excluded from the nucleus, leaving the nuclei unstained while the cytoplasm is stained blue (Figure 3.4 panel A). For all β-gal transfected cells, photographs were taken with bright field and phase contrast filters. Under bright field, only the cells expressing β-gal are visible, while when viewed under phase contrast all cells in the field are visible. This gives an indication of the transfection efficiency (number of β-gal expressing cells divided by the number of total cells) and the exact subcellular localization of the β-gal fusion proteins. Transfection efficiencies for cells transfected with β-gal-ER1 fusion constructs ranged from 50-80%.

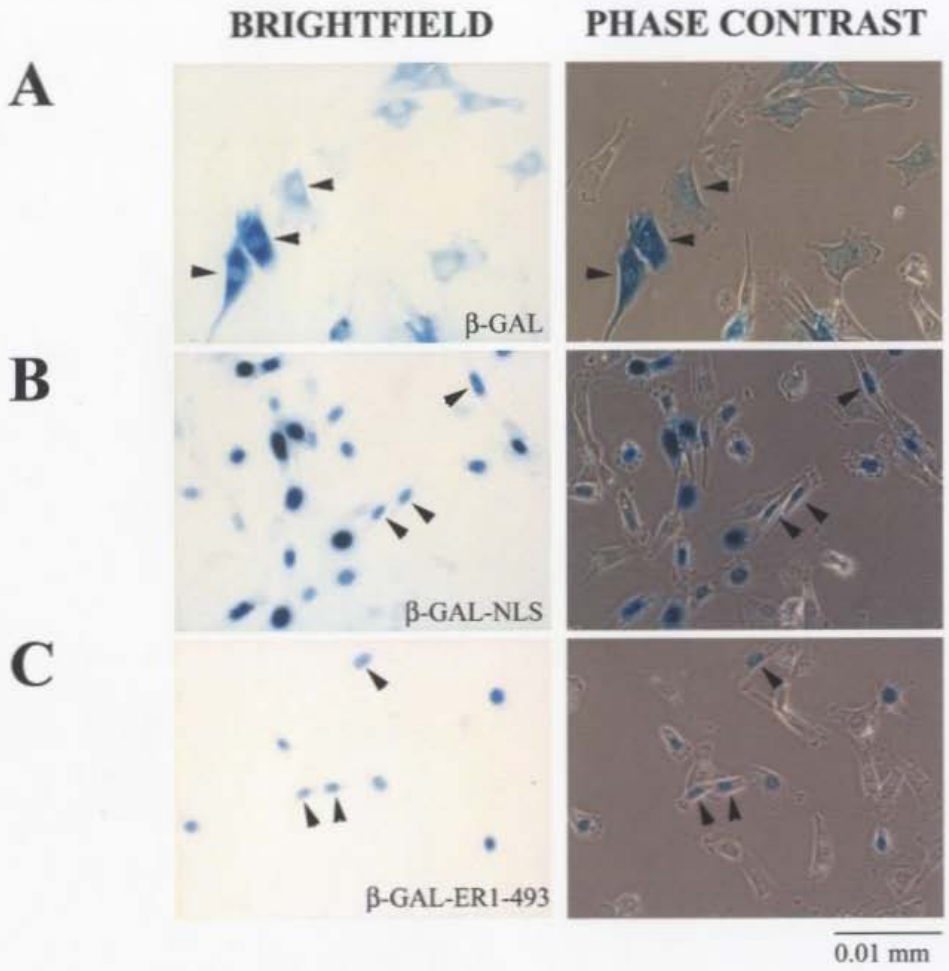
Figure 3.4 Addition of β-galactosidase to the N-terminus of MI-ER1 does not interfere with the nuclear localization.

NIH 3T3 cells were transfected with different β-GAL constructs. The subcellular localization of the β-gal fusion proteins was determined by microscopy 24 hours post-transfection. Photographs were taken of the same field using brightfield (left) and phase contrast (right) filters.

A, β-GAL localizes to the cytoplasm of transfected NIH 3T3 cells. B, When β-GAL fused to the SV40 T-Ag NLS (β-GAL-NLS) is transfected into NIH 3T3 cells, nuclear staining is observed. C, A β-GAL fusion containing full-length MI-ER1 was generated and transfected into NIH 3T3 cells. Subcellular localization of the β-GAL-ER1 fusion protein was compared to β-GAL and β-GAL-NLS.

Arrows indicate cells that show subcellular localization of the fusion protein that is most representative of the subcellular localization of the majority of cells expressing the indicated protein.

Scale bar = 0.01 mm.


BRIGHTFIELD
PHASE CONTRAST
A
B
C
β-GAL
β-GAL-NLS
β-GAL-ER1-493
0.01 mm

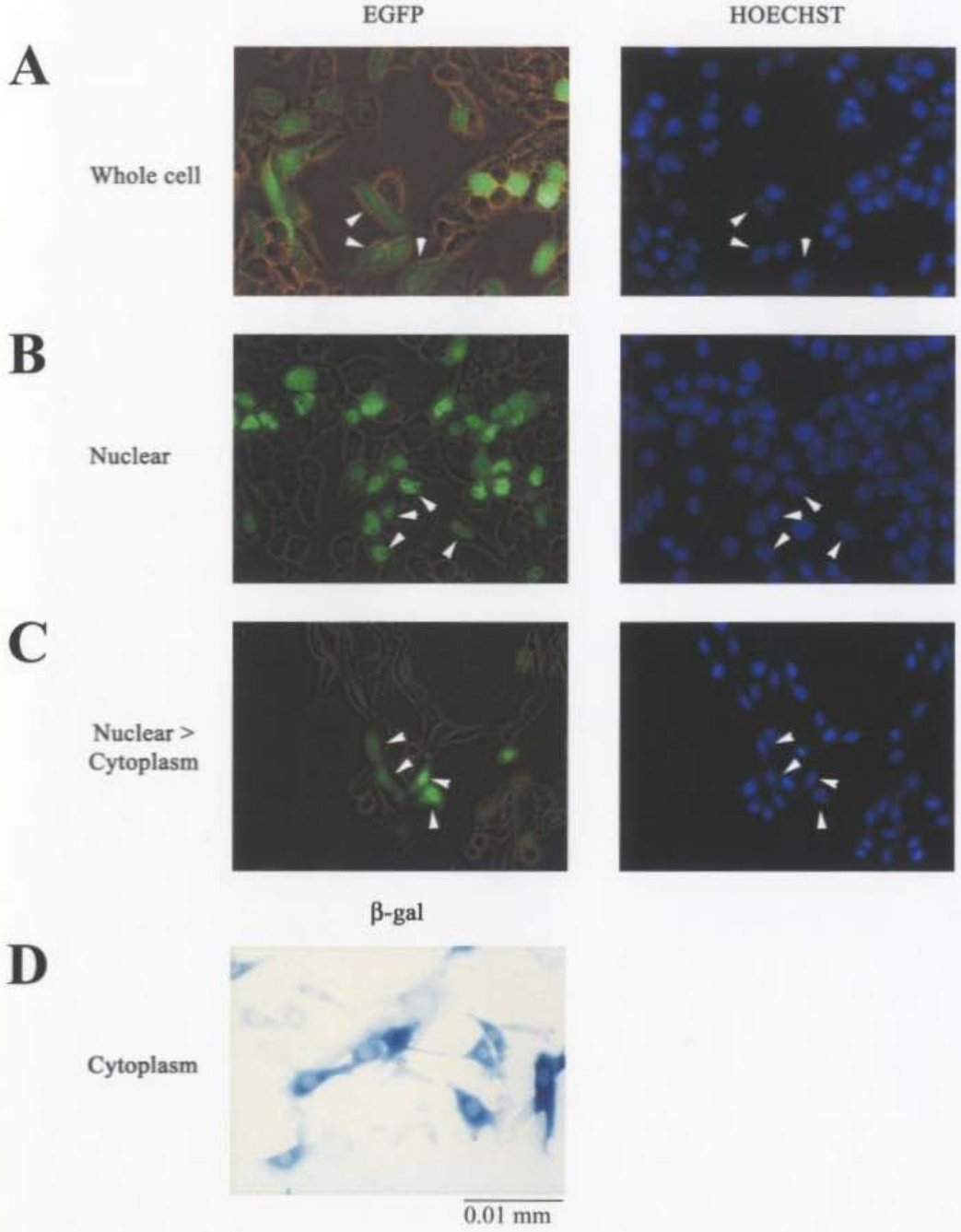
The use of two different fusion proteins is an excellent way for identifying functional NLS, because the size and structure of the fusion proteins can differ, which can result in different conformations, which in turn may influence the ability of the NLS to be recognized by the import machinery. Therefore, by using both EGFP and β-gal as fusions, the effects of these variabilities on the outcome of the experiments were minimized.

The subcellular localization of proteins in the expressing cells was defined as follows: nuclear localization indicates cells that show expression of fusion-protein exclusively in the nucleus of the cell, as is the case when XMI-ER1 full length is fused to EGFP (Figure 3.5 panel B). Location of all nuclei within the field was determined by staining of the DNA with Hoechst 33342, which resulted in a blue stain when viewed by fluorescence microscopy using a DAPI filter (406 nm; Figure 3.5). When staining was observed in the whole cell, with staining in the nucleus more intense than staining in the cytoplasm, the term "Nuclear greater than (>) Cytoplasm" was used (Figure 3.5, panel C). When no definite nuclear staining was observed, but rather staining of the nucleus and the cytoplasm in equal intensity, the term "Whole cell" staining, or "Nucleus equal to (=) Cytoplasm" was used (Figure 3.5, panel A), as is the case for EGFP when not fused to an NLS. An additional category is added when β-gal fusions are used, because the size of the protein restricts diffusion through the NPC. Therefore, β-gal, when not fused to an NLS, is characterized by an exclusion of the protein from the nucleus; indicated in figure 3.5, panel D.

Figure 3.5 Definition of subcellular localization of fusion proteins in NIH 3T3 cells.

The subcellular localization of proteins can be divided into four categories: whole cell staining (A), Nuclear staining (B), Nuclear greater than cytoplasmic staining (C) and cytoplasmic staining (D). Arrows indicate cells displaying the indicated subcellular localization.

Scale bar = 0.01 mm.


EGFP
HOECHST
A
Whole cell
B
Nuclear
C
Nuclear >
Cytoplasm
β-gal
D
Cytoplasm
0.01 mm

3.4 Identification of a functional NLS in the C-terminus of XMI-ER1

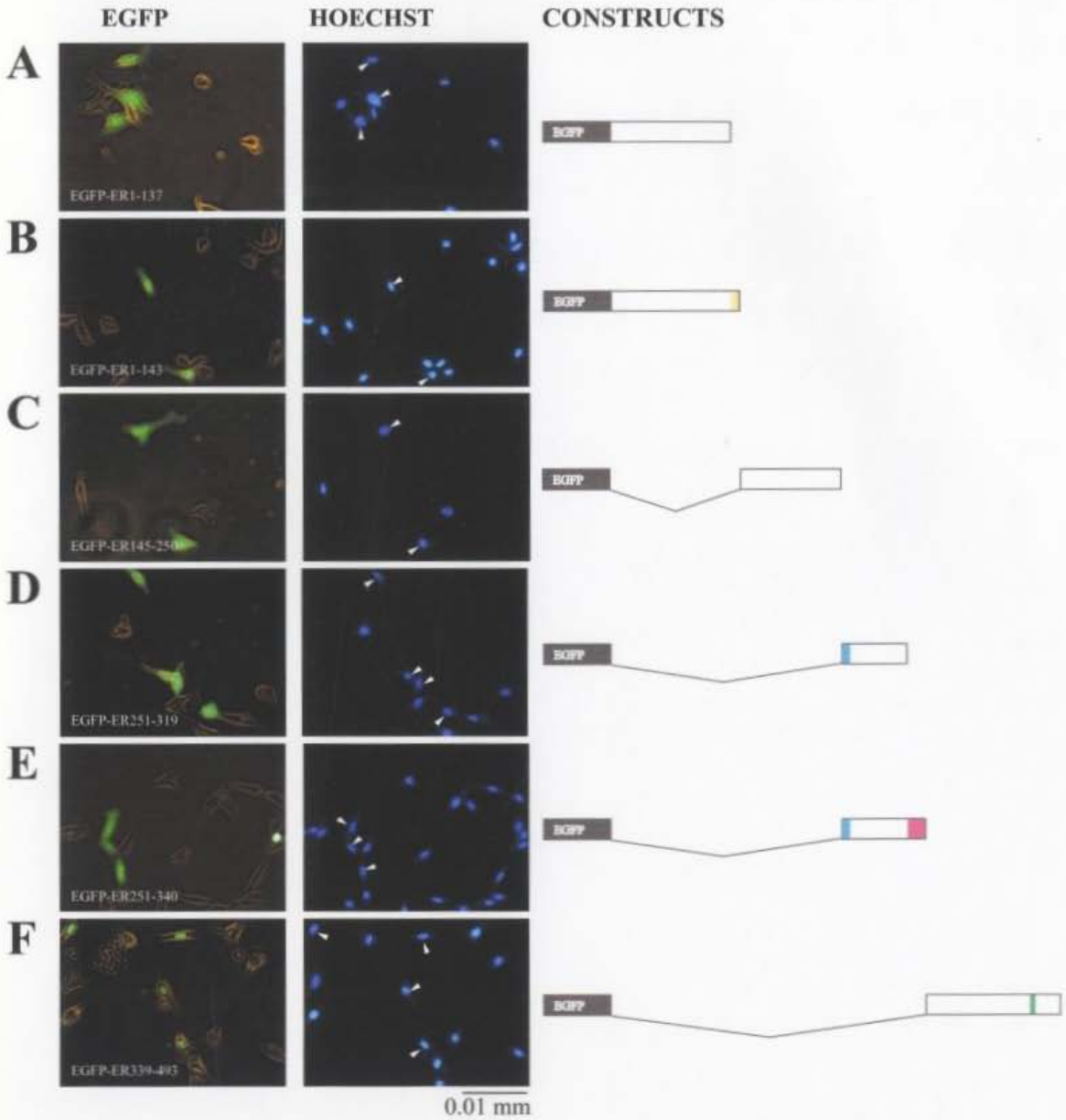
To identify whether the putative NLS are functional for translocating XMI-ER1 to the nucleus, XMI-ER1 was divided into six regions and these regions were fused to the C-terminus of EGFP. The different regions were chosen based on the position of the putative NLS in full length XMI-ER1 and were named as follows: (the numbers indicate the amino acid numbers of the protein, as shown in figure 1.4.1 and 2.1) ER1-137; ER1-143 (includes NLS1); ER145-250; ER251-319 (includes NLS2); ER251-340 (includes NLS2 and NLS3) and ER339-493 (includes NLS4) (Figure 3.6). The *egfp*-C_2 plasmids including the regions of *Xmi-er1* corresponding to these amino acids were transfected into NIH 3T3 cells. Forty-eight hours post-transfection the cells were fixed onto chamberslides and subcellular localization of the fusion proteins was observed by fluorescent microscopy. The subcellular localization of the fusion proteins was compared to EGFP and EGFP-ER1-493, as controls (full length, figure 3.3).

Cells transfected with EGFP-ER1-137 showed whole cell staining (55%) and some nuclear>cytoplasmic (45%) staining (3.6A), while EGFP-ER1-143 showed nuclear > cytoplasmic staining (88%, 3.6B). EGFP-ER145-250, EGFP-ER251-319 and EGFP-ER251-340 showed whole cell staining (3.6C, D and E).

Figure 3.6 The nuclear targeting signal of MI-ER1 is located in the C-terminus

EGFP-ER1 fusions spanning the entire XMI-ER1 protein were constructed and transfected into NIH 3T3 cells. Subcellular localization was determined by fluorescence microscopy 48 hours after transfection. A schematic illustration for each of the fusion constructs is shown on the right. The colored bars are representative for each of the putative NLS as shown in figure 3.1. Photographs illustrating the subcellular localization of each EGFP-ER1 fusion protein and Hoechst 33342 staining of the same field to visualize nuclei are shown for each construct. A: EGFP-ER1-137. B: EGFP-ER1-143. C: EGFP-ER145-250. D: EGFP-ER251-319. E: EGFP-ER251-340. F: EGFP-ER339-493. Arrows indicate nuclei of cells that express EGFP as shown in the panel on the left.

Scale bar = 0.01 mm.


EGFP
HOECHST
CONSTRUCTS
A
B
C
D
E
F
EGFP-ER1-137
EGFP-ER1-143
EGFP-ER145-250
EGFP-ER251-319
EGFP-ER251-340
EGFP-ER339-493
EGFP
0.01 mm

Only the C-terminus of ER1 (EGFP-ER339-493), containing NLS4, was localized exclusively to the nucleus, indicating that the NLS4 might function as an NLS (3.6F). Although none of the other constructs show exclusive nuclear staining, a high proportion (88%) of the cells transfected with EGFP-ER1-143 showed nuclear greater than cytoplasmic staining, indicating that NLS1 might function as a weak NLS (*vide infra*) (Table 3.1).

Table 3.1. Subcellular localization of EGFP-XMI-ER1 fusion proteins after transfection into NIH 3T3 cells

Construct	% N	% N>C	% N=C
EGFP	0	6	94
EGFP-ER1-493	98	2	0
EGFP-ER1-137	0	45	55
EGFP-ER1-143	0	88	12
EGFP-ER145-250	0	29	71
EGFP-ER251-319	0	6	94
EGFP-ER251-340	0	3	97
EGFP-ER339-493	96	4	0

To test whether the level of transiently expressed protein in the cells could saturate the nuclear translocation, cells were lysed in sample buffer 48 h post-transfection and subjected to Western Blot analysis using a monoclonal anti-GFP antibody (Figure 3.7). EGFP-ER1-493, EGFP-ER1-137 and EGFP-ER1-143 do not run at the expected size (table 3.1.1), but run larger. This is probably due to the high incidence of acidic residues (Paterno *et al.*, 1997). Presence of many acidic residues can alter the net charge of the protein, resulting in separation not

only based on the size of the protein, but also on the net charge of the protein (Bell and Bell, 1988). This results in slower migration of the proteins, so the protein appears larger than the expected size (table 3.2).

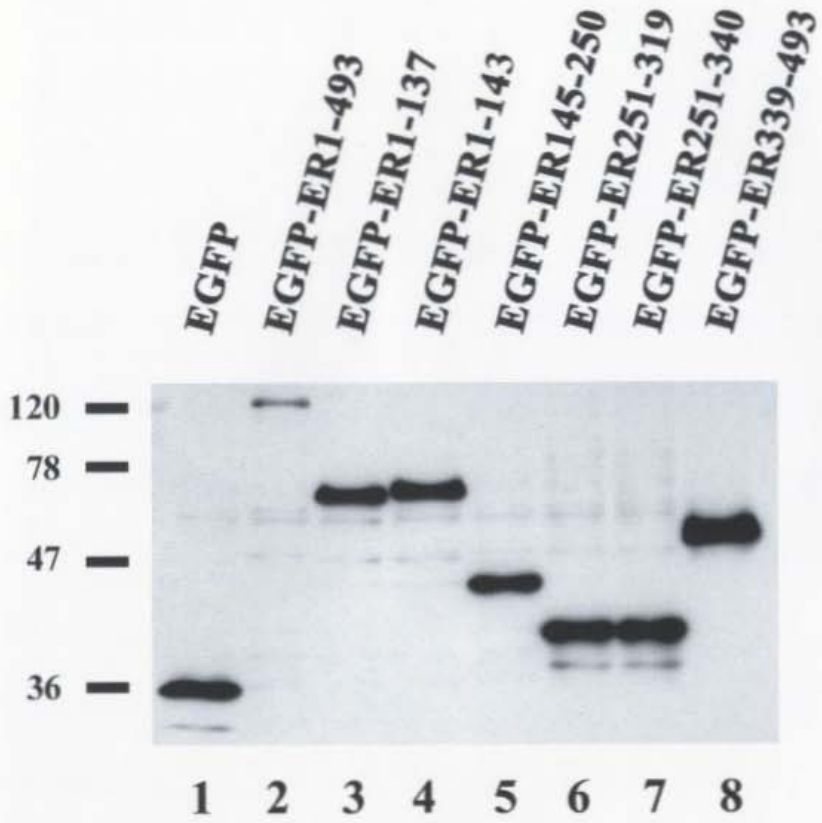
The six fusion-constructs demonstrated similar levels of expression in NIH 3T3 cells (Figure 3.7 lanes 3-8), indicating that cytoplasmic staining is due to the inability of the NLS to direct the fusion protein to the nucleus and not due to overloading of the import machinery.

Table 3.2 Predicted molecular weight of the EGFP-fusion proteins and the probability of diffusion of the EGFP-ER1 fusion proteins through the nuclear pores as described for other proteins by (Ribbeck and Gorlich, 2001).

Construct	Predicted size of EGFP-fusion protein	Probability of diffusion
EGFP	29 kDa	Yes
EGFP-ER1-493	83 kDa	No
EGFP-ER1-137	44 kDa	Yes, but slow
EGFP-ER1-143	45 kDa	Yes, but slow
EGFP-ER145-250	41 kDa	Yes
EGFP-ER251-319	37 kDa	Yes
EGFP-ER251-340	39 kDa	Yes
EGFP-ER339-493	46 kDa	Yes, but slow

Figure 3.7 Western Blot analysis of cells expressing different EGFP-ER1 fusion proteins.

Cells transfected with EGFP, EGFP-ER1, EGFP-ER1-137, EGFP-ER1-143, EGFP-ER145-250, EGFP-ER251-319, EGFP-ER251-340 and EGFP-ER339-493 were harvested and subjected to Western Blot analysis using an anti-GFP antibody. The positions of the molecular weight markers are indicated on the left. All constructs demonstrated similar protein expression levels as EGFP.


EGFP
EGFP-ER1-493
EGFP-ER1-137
EGFP-ER1-143
EGFP-ER145-250
EGFP-ER251-319
EGFP-ER251-340
EGFP-ER339-493
120
78
47
36
1 2 3 4 5 6 7 8

β-gal fusion constructs (in cβgal-CS2$^+$) were generated by addition of the regions of XMI-ER1 that showed nuclear or nuclear greater than cytoplasmic staining: ER1-493, ER1-137, ER1-143, and ER339-493. This was done to investigate whether the nuclear greater than cytoplasmic staining was a result of diffusion of the EGFP-ER1 fusion protein or whether there was active nuclear transport. As a positive control for nuclear translocation β-galactosidase was fused to the NLS of SV40 T-Ag (nβ-gal-CS2$^+$), while the β-galactosidase without an NLS was used as a control for cytoplasmic staining (cβ-gal-CS2$^+$) (figure 3.4). After transfection into NIH 3T3 cells the subcellular localization was determined. The stained cells were viewed by bright field microscopy. For each construct the percentage of cells displaying nuclear, nuclear greater than cytoplasmic nuclear, whole cell, and cytoplasmic staining were calculated (Table 3.3/ Figure 3.8).

Table 3.3 Subcellular localization of β-gal fusion constructs in NIH 3T3 cells.

Construct	% N	% N>C	% N=C	% C
β-gal	0	5	0	95
β-gal-NLS (SV40)	92	8	0	0
β-gal-ER1-493	94	6	0	0
β-gal-ER1-137	6	54	20	20
β-gal-ER1-143	27	66	3	4
β-gal-ER339-493	95	5	0	0

Transfections into NIH 3T3 cells showed nuclear localization of β-gal-NLS, ER1-493 and ER339-493 as was shown for EGFP-ER1-493 and EGFP-ER339-493 cells. β-Gal-ER1-143 showed nuclear staining in 27% of the cells, indicating that NLS1 may function as a weak NLS. No nuclear localization was observed for the negative control β-gal (see figure 3.4). Cells transfected with β-gal-ER1-137, which does not contain the NLS1, displayed little nuclear staining (6%) with 54% of the cells showing nuclear greater than cytoplasmic (N>C) staining. This suggests that ER1-137 is translocated by active transport that is not due to NLS1. As mentioned before, β-gal is too large to be able to diffuse through the NPC; therefore these results suggest that nuclear transport of ER1-137 is due to active import rather than passive diffusion of this region.

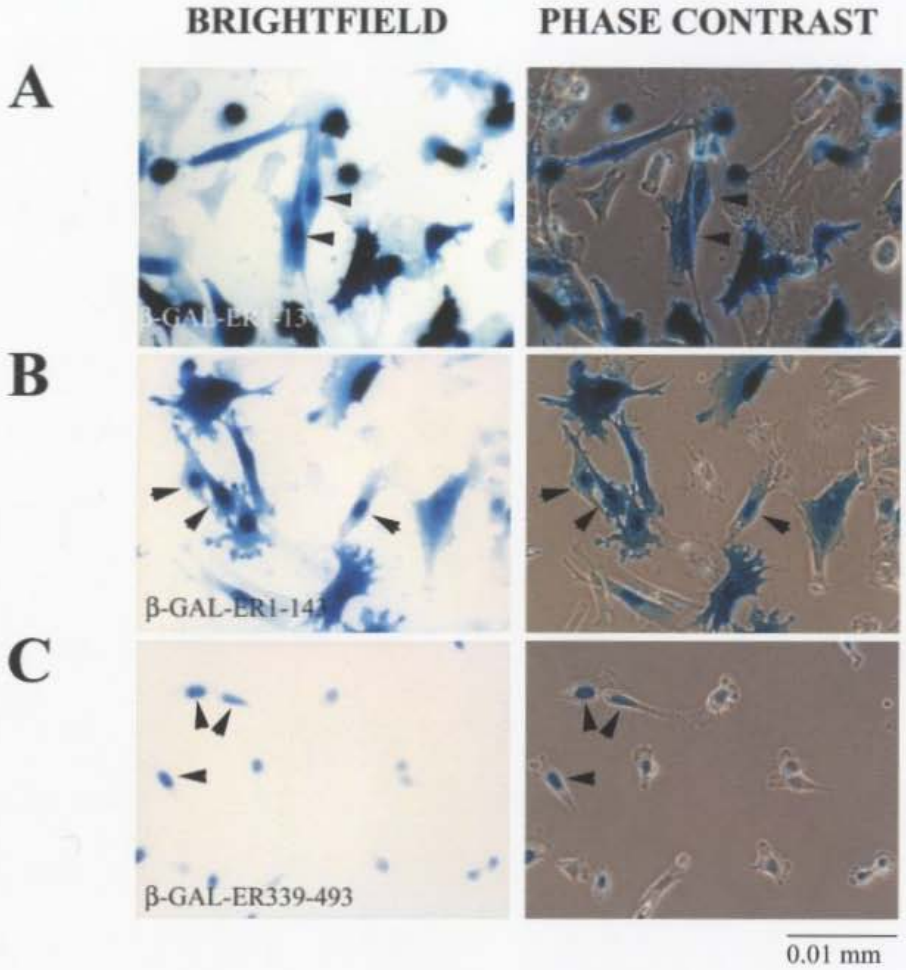
Transfection of regions of XMI-ER1 either fused to EGFP or β-gal into NIH 3T3 cells have shown that a strong NLS is located in the C-terminus of XMI-ER1, while a weak NLS is located in the N-terminal region of XMI-ER1. Also an alternative additional nuclear targeting signal may be located between residues 1-137. This may be a very weak NLS as well as a motif that can bind a molecule that is targeted to the nucleus, resulting in piggy backing of ER1-137.

Figure 3.8 NIH 3T3 cells transfected with β-gal-ER339-493 display exclusive nuclear staining, whereas cells transfected with β-gal-ER1-137 or β-gal-ER1-143 do not.

NIH 3T3 cells were transfected with β-gal-ER1-137 (A), β-gal-ER1-143 (B), or β-gal-ER339-493 (C). Subcellular localization of the fusion proteins as determined by microscopy was compared to β-gal and β-gal-ER1-493 (as shown in figure 3.4).

Arrows indicate cells that show subcellular localization of the fusion protein that is most representative of the subcellular localization of the majority of cells expressing the indicated protein.

Differences in cell size are probably caused by nutritional differences caused by cell density. Cells that lack nutrition spread out and appear to be larger than well-fed cells.


BRIGHTFIELD
PHASE CONTRAST
A
B
C
β-GAL-ER339-493
0.01 mm

3.5 NLS4 is necessary and sufficient for the nuclear transport of XMI-ER1

3.5.1 NLS4 is able to direct EGFP and β-GAL to the nuclei of NIH 3T3 cells.

The signal for nuclear localization is located in the C-terminus of XMI-ER1 between amino acids 339 and 493. To identify whether the putative NLS4, which is located between residues 463 and 469 is necessary for the nuclear translocation of XMI-ER1, two constructs were generated: EGFP-ER1-461, which is truncated just before the NLS4 and EGFP-ER1-473, which includes the basic core of NLS4 (see figure 3.9 A and B) Subcellular localization of these constructs, after transfection into NIH 3T3 cells, was compared to the subcellular localization of EGFP and EGFP-ER1-493.

As is shown in table 3.4 and figure 3.9, the EGFP-ER1-473 was almost exclusively nuclear (90%, Figure 3.9 B), while EGFP-ER1-461 was not (1%, Figure 3.9 A). This indicates that the residues between 461 and 473, where the NLS4 is located, are important for the nuclear import of XMI-ER1. Figure 3.10 shows the level of transient protein expression as analyzed by Western Blot using an anti-GFP antibody. EGFP-ER1-461 (panel A, lane 3) and EGFP-ER1-473 (panel A, lane 4) are expressed at a level similar to EGFP-ER1-493 (panel A, lane 2).

Table 3.4 Subcellular localization of EGFP-ER1-461 an EGFP-ER1-473 after transfection into NIH 3T3 cells.

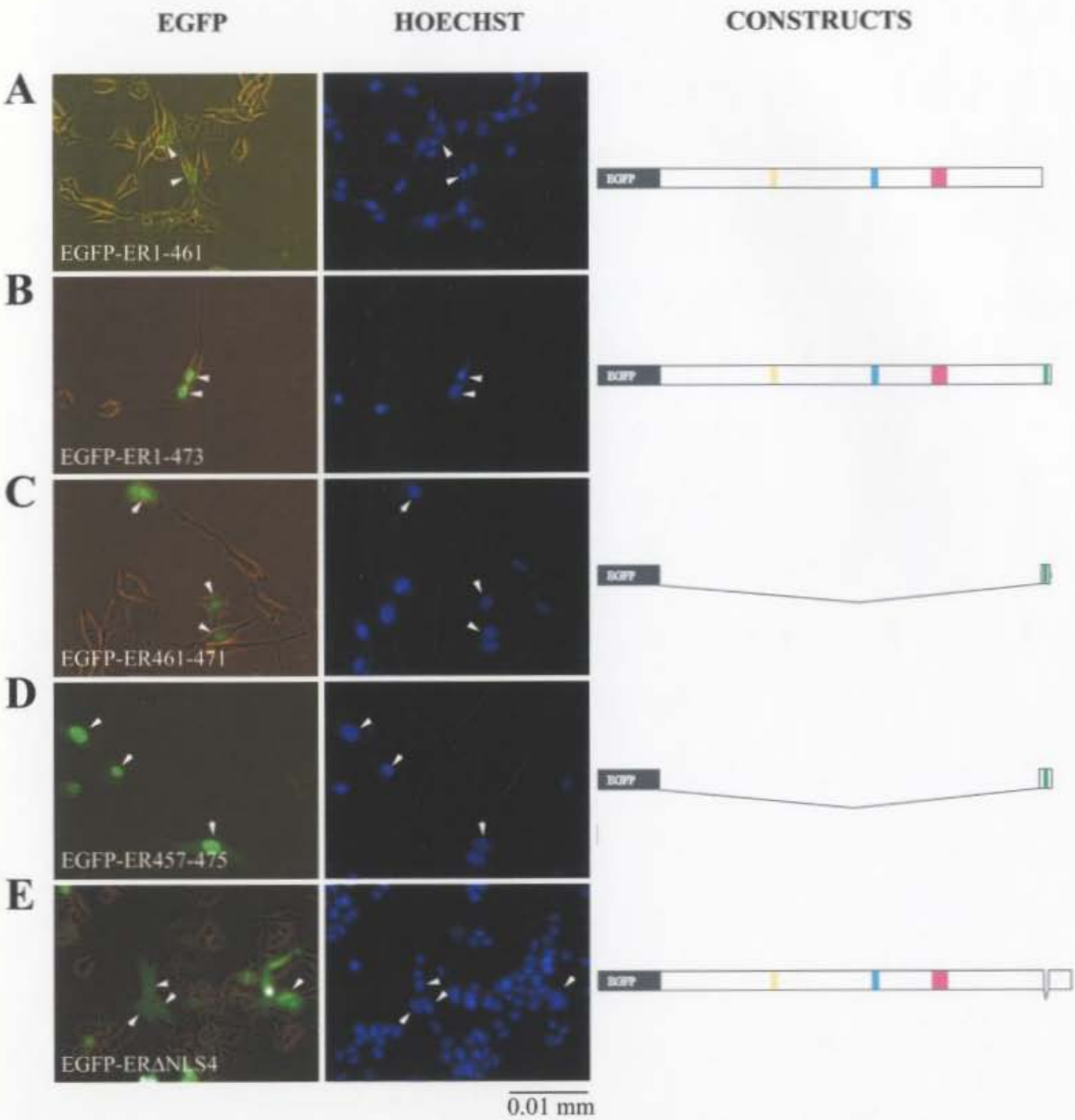
Construct	% N	% N>C	% N= C
EGFP-ER1-473	88	9	3
EGFP-ER1-461	0	21	79

Figure 3.9 NLS4 is necessary and sufficient for the nuclear translocation of MI-ER1 in NIH 3T3 cells.

EGFP-ER1 fusions used to investigate the function of the putative NLS4 were generated and transfected into NIH 3T3 cells. Subcellular localization was determined by fluorescence microscopy 48 hours after transfection. EGFP and EGFP-ER1 were transfected in the same experiment as controls for whole cell and nuclear staining, respectively (not shown). A schematic illustrating each of the fusion constructs is shown on the right. Photographs illustrating the subcellular localization of each of the fusion proteins and Hoechst 33342 staining of the same field are shown for each construct. A: EGFP-ER1-461. B: EGFP-ER1-473. C: EGFP-ER461-471. D: EGFP-ER457-475. E: EGFP-ΔNLS4.

Arrows indicate cells that show subcellular localization of the fusion protein that is most representative of the subcellular localization of the majority of cells expressing the indicated protein.

Scale bar = 0.01mm.

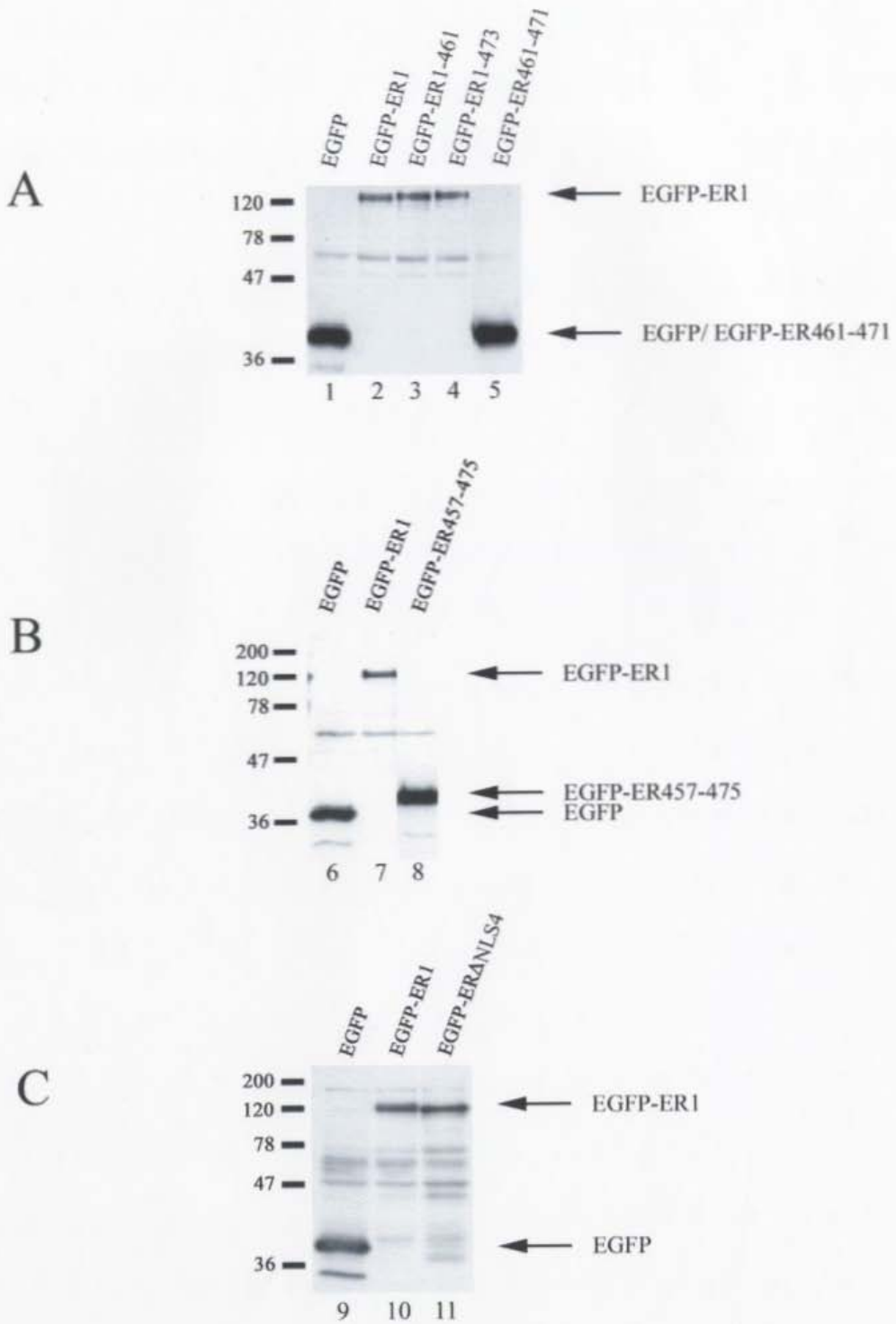

EGFP
HOECHST
CONSTRUCTS
A
EGFP-ER1-461
B
EGFP-ER1-473
C
EGFP-ER461-471
D
EGFP-ER457-475
E
EGFP-ERΔNLS4
0.01 mm

To identify the precise region of the NLS necessary for the nuclear translocation of XMI-ER1, two EGFP fusion constructs were generated: EGFP-ER457-475, which contains NLS4 including 6 amino acid residues flanking on both sides, and EGFP-ER461-471, which contains the basic core region (Figure 3.9). EGFP-ER457-475 (^{457}PDDTNE**RPIKRQR**MDSPGK^{475}) was almost as efficient as the full-length sequence at directing exclusive nuclear targeting (74%N vs. 98% N for the full-length), while EGFP-ER461-471 was less efficient with 38% nuclear staining (vs. 98% for the full-length), with the remaining cells showing nuclear greater than cytoplasmic staining (62%) (Figure 3.9 C and D/ Table 3.5). When fused to β-galactosidase amino acids 457-475 show 93% nuclear staining in transiently transfected NIH 3T3 cells (Table 3.5/ Figure 3.11, B). The greater number of cells showing nuclear staining after transfection with β-Gal-ER457-475 as compared to the EGFP-ER457-475 fusion could be due to a conformational difference, or due to the ability of the EGFP-ER457-475 to diffuse through the NPC from the nucleus to the cytoplasm, resulting in some cytoplasmic staining. A second β-gal-ER1 fusion containing amino acid residues 461-471 was 37% nuclear, with 61% nuclear greater than cytoplasmic staining (Figure 3.11 A). This indicates that amino acids flanking the NLS4 basic core are necessary for the efficient transport of XMI-ER1 to the nucleus. Figure 3.10 shows the level of protein expression of EGFP-ER461-471 (panel A, lane 5) and EGFP-ER457-475 (panel B, lane 8) as revealed by Western Blot analysis.

Figure 3.10 Western Blot analysis of EGFP-ER1 fusions used to identify the minimal NLS4 region.

Cells transfected with the indicated EGFP-ER1 fusion constructs were harvested 48 hours after transfection and subjected to Western Blot analysis using an anti-GFP antibody. The positions of the molecular weight markers are indicated on the left.

Protein expression levels of different EGFP-ER1 fusions were compared to the expression of EGFP and EGFP-ER1-493. Panels A, B and C indicate different Western blots. A, Expression of EGFP-ER1-461 (lane 3), EGFP-ER1-473 (lane 4) and EGFP-ER461-471 (lane 5), compared to EGFP (lane 1) and EGFP-ER1-493 (lane 2). B, Expression of EGFP-ER457-475 (lane 8) compared to EGFP (lane 6) and EGFP-ER1-493 (lane 7). C, Expression of EGFP-ERΔNLS4 (lane 11) compared to EGFP (lane 9) and EGFP-ER1-493 (lane 10).


A
EGFP
EGFP-ER1
EGFP-ER1-461
EGFP-ER1-473
EGFP-ER461-471
120
78
47
36
EGFP-ER1
EGFP/ EGFP-ER461-471
1 2 3 4 5
B
EGFP
EGFP-ER1
EGFP-ER457-475
200
120
78
47
36
EGFP-ER1
EGFP-ER457-475
EGFP
6 7 8
C
EGFP
EGFP-ER1
EGFP-ERΔNLS4
200
120
78
47
36
EGFP-ER1
EGFP
9 10 11

3.5.2 Deletion of the NLS4 results in cytoplasmic localization of XMI-ER1

Analysis of core NLS has shown that mutation of 1 basic residue can abolish nuclear transport (Goldfarb *et al.*, 1986; Kalderon *et al.*, 1984a). In the SV40 large T-Ag NLS substitution of the second lysine of PK^{128}**K**KRKV with threonine abolishes nuclear translocation of the protein (Goldfarb *et al.*, 1986; Kalderon *et al.*, 1984a), while substitution of either of two basic residues in the human c-MYC NLS PAA^{323}**KR**VKLDS for alanine resulted in decreased affinity for importin-α, and subsequent cytoplasmic localization (Hodel *et al.*, 2001). Thus substitution or deletion of the first and / or second basic residue in the core NLS results in abolished nuclear localization. Therefore, to verify that the NLS4 is a *bona fide* NLS necessary for the nuclear import of XMI-ER1, we deleted four amino acids of NLS4, ^{464}PIKR, from the full length EGFP-ER1-493, designated ΔNLS4. Transfection of EGFP-ΔNLS4 into NIH 3T3 cells showed nuclear staining in only 4% of the cells, while whole cell staining was observed in 75% of the cells, showing that the nuclear import of XMI-ER1 is abolished after partial deletion of the core NLS (Figure 3.9, E. Table 3.5). The level of protein expression as revealed by Western Blot analysis indicated that the cytoplasmic staining is not due to the abundance of transient protein expression in these cells and thus saturation of the import machinery (Figure 3.10, panel C, lane 3). Deletion of ^{464}PIKR in the β-gal-ER1-493 (β-gal-ΔNLS4) also resulted in a decrease in nuclear localization (11% N) with 78% of the cells showing

cytoplasmic staining (Figure 3.11, C. Table 3.5). This indicated that the NLS4 is necessary for efficient nuclear translocation of XMI-ER1.

Table 3.5 Subcellular localization of EGFP-ER461-471, EGFP-ER457-475, EGFP-ERΔPIKR, β-gal-ER461-471, β-gal-ER457-475 and β-gal-ΔNLS4 after transfection into NIH 3T3 cells.

Construct	% N	% N>C	% N=C	%C
EGFP-ER461-471	38	62	0	
EGFP-ER457-475	74	22	4	
EGFP-ERΔNLS4	4	21	75	
β-gal-ER461-471	37	61	2	0
β-gal-ER457-475	93	7	0	0
β-gal-ERΔNLS4	11	11	0	78

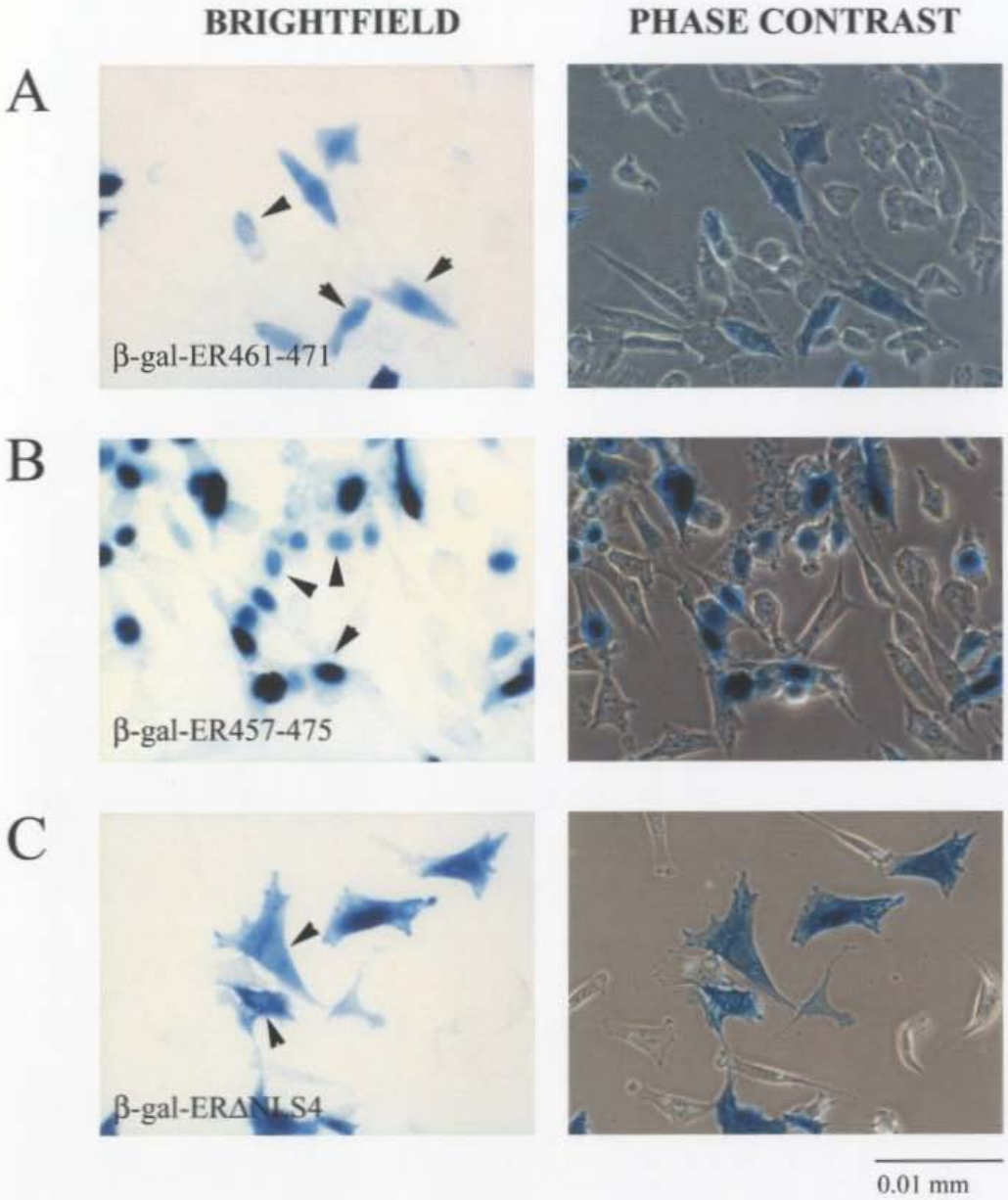
Together these results indicate that an intact NLS4 is necessary for the efficient nuclear localization of XMI-ER1 and that addition of the complete NLS4 region, the core basic cluster and the flanking region (6 residues N-terminal and 4 residues C-terminal of the basic cluster), to a cytoplasmic protein is sufficient for nuclear import. Since these are the requirements for a functional NLS, NLS4 (ER457-475) can be called an NLS.

Since deletion of the NLS4 in XMI-ER1 does not completely abolish nuclear translocation (4% N for EGFP-ΔNLS4 and 11% N for β-GAL-ΔNLS4) and ER1-143 showed a large amount of nuclear greater than cytoplasmic staining (80%), we investigated the possibility of the presence of an additional NLS in XMI-ER1.

Figure 3.11 The NLS4 basic cluster and flanking regions are necessary for the nuclear translocation of MI-ER1 in NIH 3T3 cells.

NIH 3T3 cells were transfected with β-gal-ER461-471 (A), β-gal-ER457-475 (B), or β-gal-ERΔNLS4 (C). Subcellular localization of the fusion proteins as determined by microscopy was compared to β-gal and β-gal-ER1-493 (as shown in figure 3.4).

Arrows indicate cells that show subcellular localization of the fusion protein that is most representative of the subcellular localization of the majority of cells expressing the indicated protein.


BRIGHTFIELD
PHASE CONTRAST
A
β-gal-ER461-471
B
β-gal-ER457-475
C
β-gal-ERΔNLS4
0.01 mm

3.6 Additional weak NLSs in XMI-ER1

Several of the ER1 fusion constructs, which did not contain NLS4, appeared to have a greater level of fusion protein in the nucleus than in the cytoplasm, not accountable by simple diffusion (Figures 3.6, 3.8, 3.9, 3.11). This suggested that there might be additional, albeit weak, NLSs in XMI-ER1. In an attempt to identify these regions, additional fusion constructs were generated. These included EGFP-ER136-143 (NLS1), EGFP-ERΔ138-141 (ΔNLS1), EGFP-ER259-273 (NLS2), EGFP-ER318-340 (NLS3) and EGFP-ERΔ320-338 (ΔNLS3).

After transfection into NIH 3T3 cells, only EGFP-ER136-143 demonstrated a significant number of cells (76%) in which nuclear fluorescence was greater than cytoplasmic (Figure 3.12). Also, after transfection of EGFP-ER1-143, which includes the NLS1, 88% of transfected cells demonstrated nuclear greater than cytoplasmic staining (Figure 3.6). Moreover, when the NLS4 core region is partially deleted, nuclear staining (4% for EGFP-ΔNLS4, 11% for β-gal-ΔNLS4) was observed after transfection into NIH 3T3 cells. Thus, in absence of NLS4, the predicted NLS1 appears to function as a weak nuclear targeting signal for XMI-ER1. However, deletion of NLS1 (ERΔ138-141) does not affect the nuclear localization of the EGFP-ER1 fusion (Figure 3.14), indicating that NLS1 is not necessary for import of XMI-ER1 into nuclei of NIH 3T3 cells. The percentages of nuclear, nuclear greater than cytoplasmic and nuclear equal to cytoplasmic staining for each EGFP-ER1 fusion protein were calculated, shown in table 3.6. For each construct at least 500 transfected cells were counted.

Table 3.6 Percentage nuclear staining in NIH 3T3 cells transfected with EGFP-NLS1, 2, 3 and EGFP-ΔNLS1 and 3

Construct	% N	% N>C	% N=C
EGFP-NLS1 (ER136-143)	0	76	24
EGFP-NLS2 (ER259-273)	0	29	71
EGFP-NLS3 (ER318-340)	0	14	86
EGFP-ERΔNLS1 (ERΔ138-141)	100	0	0
EGFP-ERΔNLS3 (ERΔ320-338)	100	0	0

NLS2 (ER259-273) and 3 (ER318-340) were not able to efficiently direct EGFP to the nucleus (Table 3.6). The abundance of transient protein in these transfected cells was determined by western blotting (Figure 3.13). Levels of protein expression were similar for all EGFP-NLS constructs.

These transfection experiments also showed that the putative NLS2 and 3 are not inhibited by the context in which they appear in the XMI-ER1 sequence. Directly adding the NLS region to EGFP does not change the observation that these putative NLSs are unable to direct the fusion protein to the nucleus, indicating that they are not functional NLSs.

Figure 3.12 NLS1 functions as a weak NLS when fused to EGFP, while NLS2 and NLS3 do not.

EGFP-NLS fusions containing the core NLS of putative NLS1, 2 or 3 were transfected into NIH 3T3 cells. Subcellular localization, compared to EGFP and EGFP-ER1-493, was determined 48 hours post-transfection. Schematic illustrating each EGFP-NLS construct is shown on the right. Photographs illustrating the subcellular localization of each of the fusion proteins and Hoechst 33342 staining of the same field are shown for each construct. A: EGFP-NLS1 (ER137-143). B: EGFP-NLS2 (ER259-273). C: EGFP-NLS3 (ER317-341).

Scale bar = 0.01 mm.

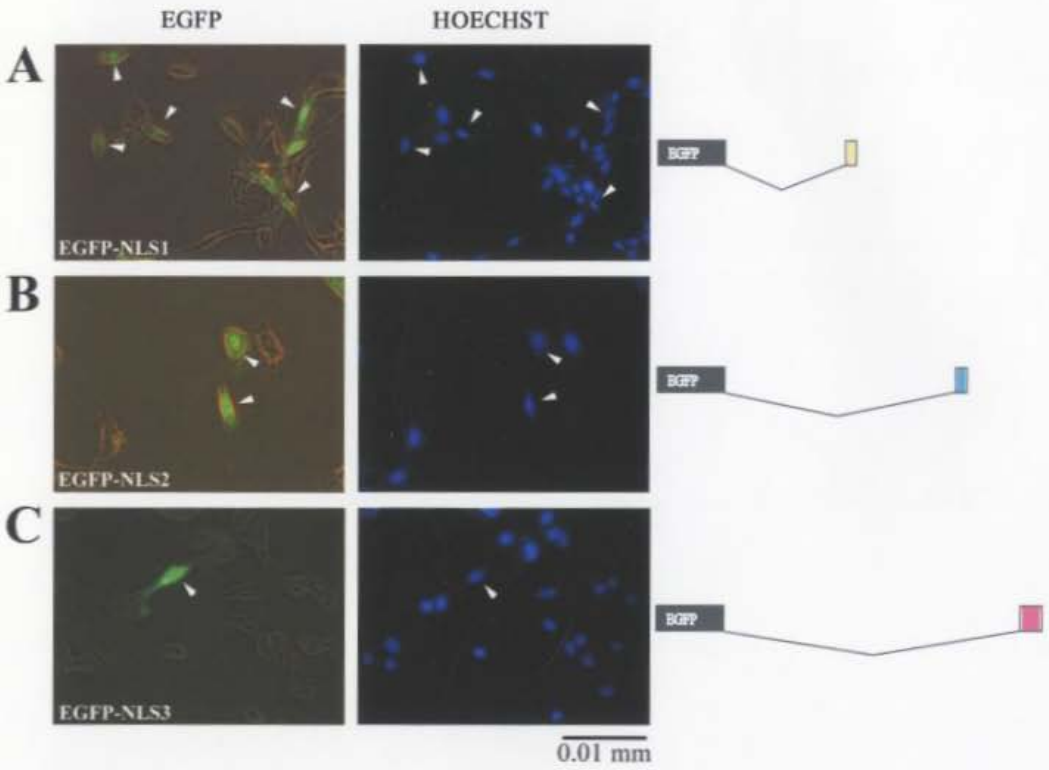
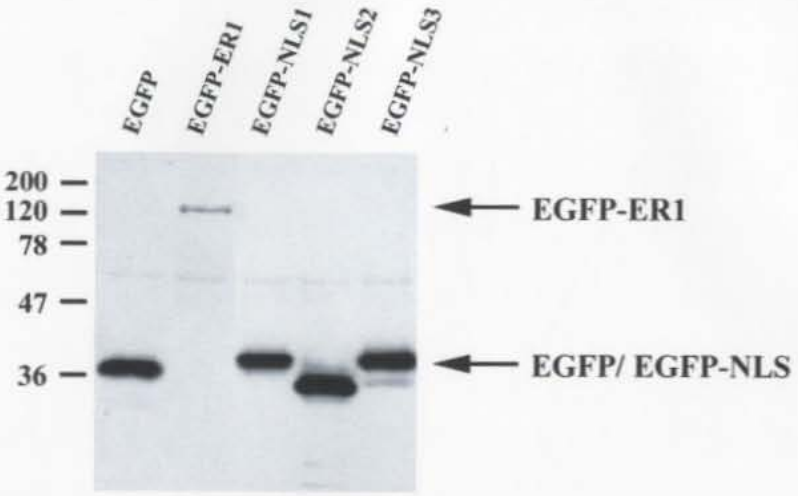

EGFP
HOECHST
A
EGFP-NLS1
EGFP
B
EGFP-NLS2
EGFP
C
EGFP-NLS3
EGFP
0.01 mm

Figure 3.13 Western Blot analysis showing the level of protein expression of EGFP-NLS1, 2 and 3

Cells transfected with the indicated EGFP fusion construct were harvested 48 hours after transfection and subjected to Western Blot analysis using an anti-GFP antibody. The positions of the molecular markers are shown on the left. All EGFP-NLS fusions demonstrated similar levels of expression in NIH 3T3 cells. Arrows indicate EGFP-ER1 (top) and EGFP and EGFP-NLS (bottom).


EGFP
EGFP-ER1
EGFP-NLS1
EGFP-NLS2
EGFP-NLS3
200
120
78
47
36
EGFP-ER1
EGFP/ EGFP-NLS

3.7 Deletion of NLS1 and/or NLS3 does not interfere with the nuclear translocation of XMI-ER1

It has been described that nuclear translocation of proteins can be dependent on multiple NLS. To investigate whether the efficiency of nuclear transport is dependent solely on NLS4 or on a co-operation of multiple putative NLS, the following constructs were generated: EGFP-ERΔ138-141, EGFP-ERΔ320-338 and EGFP-ERΔ138-141 $+^{320}$KK = AA + Δ336-338 (referred to as ΔNLS1, ΔNLS3 and ΔNLS1+ΔNLS3 respectively, Figure 3.14A).

Transfection of these constructs into NIH 3T3 cells showed exclusive nuclear localization for all these constructs (see figure 3.14B). This indicates that these NLS are not essential for the nuclear localization of XMI-ER1 and that the function of NLS4 is not dependent on a co-operation with NLS1 and/or NLS3. The level of protein expression was tested by immunoprecipitation using an anti-XMI-ER1 antibody, after which the immunoprecipitated proteins were subjected to Western Blotting with an anti-GFP antibody (Figure 3.15). Protein expression was observed for all immunoprecipitated proteins; EGFP expression levels could not be determined by this because it is not precipitated with anti-XMI-ER1. In all lanes a very abundant protein is observed around the 47-kDa markers. This staining is probably due to presence of the IgG heavy chain, which is observed when polyclonal antibodies are used for immunoprecipitation.

Figure 3.14 Nuclear translocation of MI-ER1 is not dependent on co-operation of NLS4 with NLS1, NLS3 or NLS1 and 3

A, Schematic illustrating EGFP-ER1 fusions containing deletions of the putative NLS1, NLS3 and NLS1+3, compared to full-length MI-ER1 (ER1-493). The asterisk indicates substitution of KK to AA, rather than deletion of this region. B, Subcellular localization of these constructs after transfection into NIH 3T3 cells 48 hours after transfection as shown by fluorescence microscopy using filters for EGFP (a,c,e) and corresponding Hoechst staining of the same field (b,d,f). Arrows indicate cells that show subcellular localization of the fusion protein that is most representative of the subcellular localization of the majority of cells expressing the indicated protein and the corresponding nuclei as identified by Hoechst staining.

Scale bar = 0.01 mm.

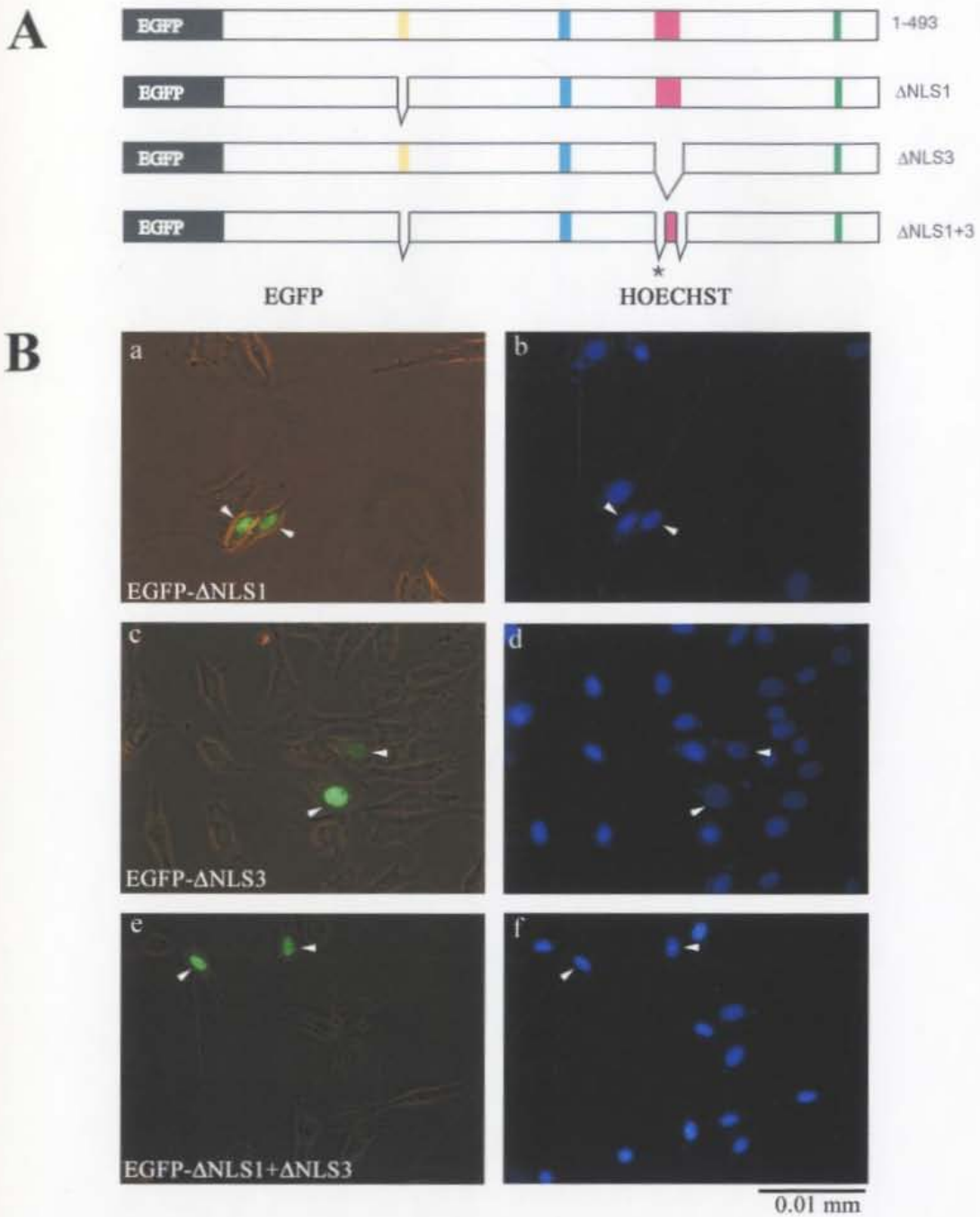
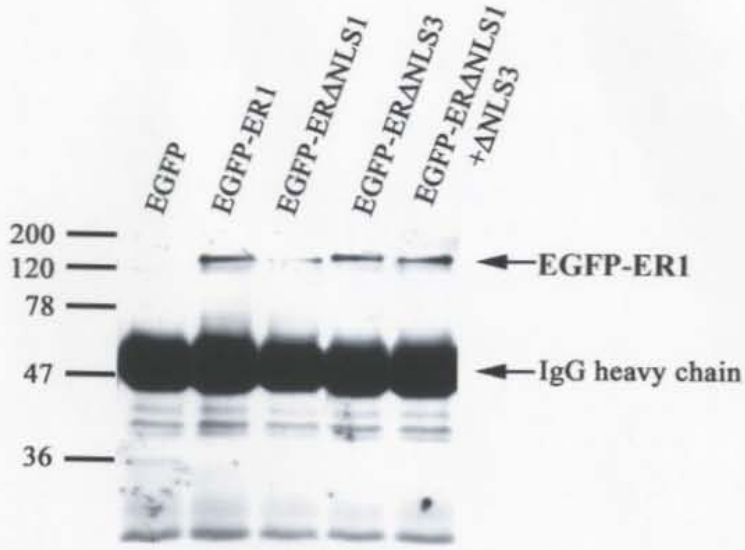

A
EGFP
1-493
ΔNLS1
ΔNLS3
ΔNLS1+3
EGFP
HOECHST
B
a
b
EGFP-ΔNLS1
c
d
EGFP-ΔNLS3
e
f
EGFP-ΔNLS1+ΔNLS3
0.01 mm

Figure 3.15 Western Blot analysis of EGFP-ER1 fusion proteins after immuno-precipitation using anti-MI-ER1 show similar levels of expression in NIH 3T3 cells

EGFP is not detected because it is not recognized by the anti-MI-ER1 antibody. In all lanes a very abundant protein is observed around the 47-kDa marker. This staining is most likely the IgG heavy chain, which can be observed when polyclonal antibodies are used for immunoprecipitation. Arrows indicate the different EGFP-ER1 proteins.


EGFP
EGFP-ER1
EGFP-ERΔNLS1
EGFP-ERΔNLS3
EGFP-ERΔNLS1 +ΔNLS3
200
120
78
47
36
EGFP-ER1
IgG heavy chain

3.8 Analysis of the putative NLS for XMI-ER1 in *Xenopus laevis* embryos

When transfected into NIH 3T3 cells, a mouse fibroblast cell line, XMI-ER1 is constitutively localized to the nucleus. However, nuclear translocation of XMI-ER1 in developing *Xenopus* embryos is stage dependent (Luchman *et al.*, 1999). No nuclear localization is observed until after stage 8 (5 hours after fertilization) and protein levels and nuclear staining diminishes at the tailbud stages (Luchman *et al.*, 1999). Regulation of nuclear import of many proteins is dependent on the developmental stage of the embryos. There are examples of proteins which possess a developmentally regulated NLS (drNLS) (Jans and Hubner, 1996). There are also proteins described that translocate during a specific phase of the cell cycle (Madeo *et al.*, 1998; Moll *et al.*, 1991). Also, different signaling pathways are activated during different stages of development. Since, XMI-ER1 was first isolated as an FGF inducible early response gene in stage 8 *Xenopus*. Activation of the FGF signaling pathway around stage 8 may be important for nuclear translocation of XMI-ER1 at this stage.

Transfection of XMI-ER1 into NIH 3T3 cells does not provide us with information about the regulation of nuclear import during *Xenopus* development. The translocation of XMI-ER1 in *Xenopus* embryos could be dependent on cell cycle dependent mechanisms that regulate the NLS, on use of an alternative drNLS, or activation of a signaling pathway like the FGF pathway. Therefore, the importance of NLS4 for nuclear translocation of XMI-ER1 in *Xenopus* needed to be determined. The observation that XMI-ER1 contains a very strong NLS

(NLS4) and an additional weak NLS (NLS1) suggests that the regulation and use of these NLS might be dependent on the context of the cell and/ or the stage of development. Therefore, it is important to identify whether the NLS4 is capable of directing the nuclear localization of XMI-ER1 in embryos, or whether alternate sequences in XMI-ER1 are responsible for nuclear translocation that is dependent on developmental events.

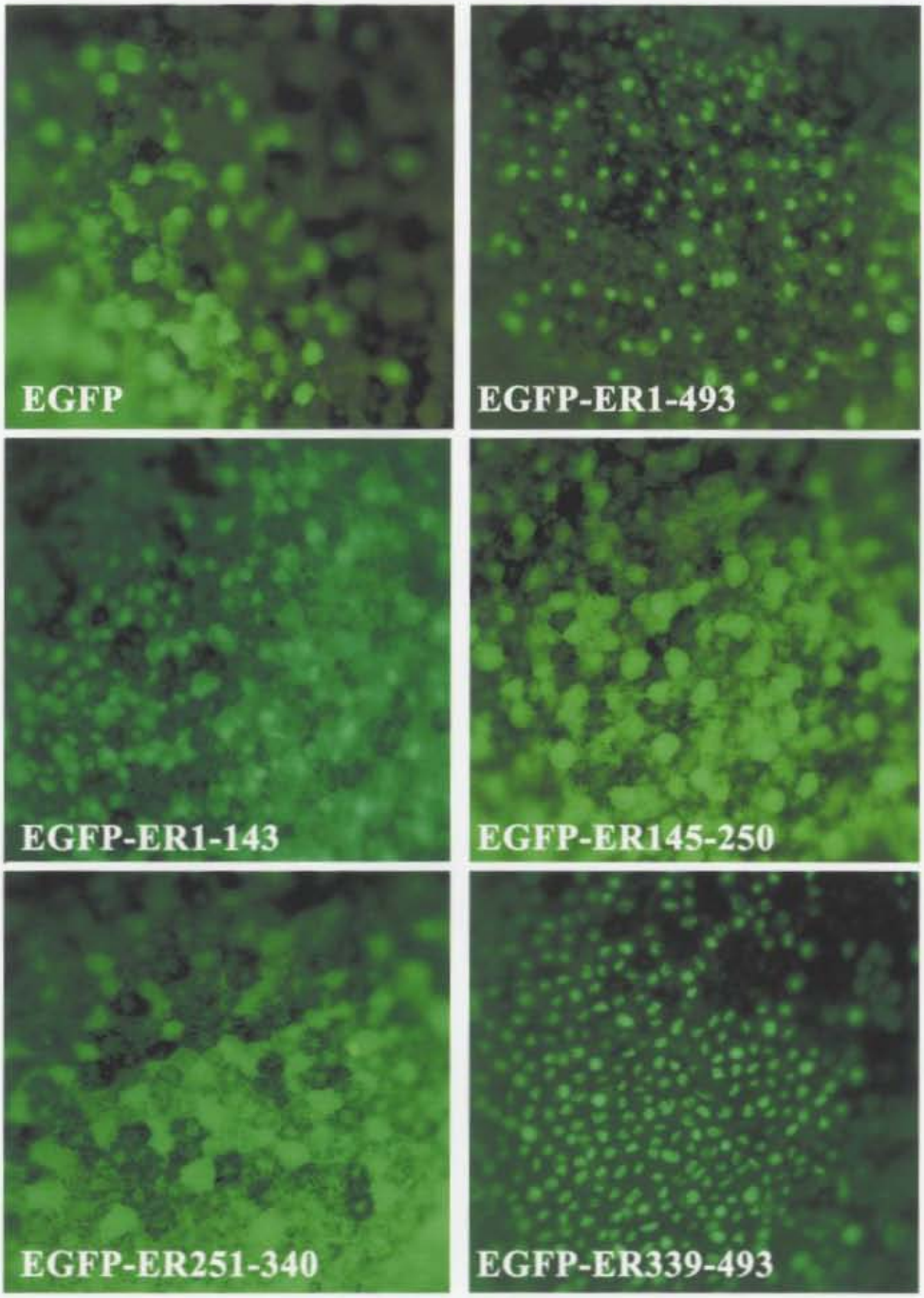
To investigate whether NLS4 is necessary for nuclear translocation of XMI-ER1 in *Xenopus* embryos, a number of different *egfp-er1* cRNAs were injected into the marginal zone of 1-cell stage wild-type embryos. The marginal zone was chosen as the injection site because it was shown that nuclear staining of endogenous XMI-ER1 is first observed in this region (Luchman *et al.*, 1999). The subcellular localization was assessed at stage 13, well past the stage, when full length XMI-ER1 was shown to be exclusively nuclear (Luchman *et al.*, 1999). The advantage of using embryos for this experiment was that *Xmi-er1* is isolated from *Xenopus*, thus expression in embryos will show how nuclear translocation of endogenous XMI-ER1 may be regulated. One disadvantage of injecting into embryos is the fact that injected RNA remains near the site of injection and thus, not all cells will contain injected cRNA. Furthermore, the cells that contain injected cRNA will contain different concentrations of cRNA and subsequently different levels of protein, depending on the distance from the injection site. Therefore, subcellular localization of the EGFP-fusion protein was determined by counting cells in the injection patch. Another disadvantage of using *Xenopus*

embryos for analyzing the subcellular localization of EGFP-ER1 fusion proteins is that *Xenopus* embryos contain yolk platelet proteins, which autofluoresce at wavelengths between 488nm and 568 nm, which includes the wavelength of the excitation energy for EGFP (495 nm), making it very difficult to investigate the subcellular localization of EGFP fusion proteins (Beumer *et al.*, 1995). The yolk protein is the embryo's main source of nutrition and is therefore very abundant, especially during earlier stages (Danilchik and Gerhart, 1987). Also, the cell nuclei are relatively small compared to the large cell (small N/C ratio) making it difficult to assess the subcellular localization of the protein. Moreover, because the injected RNA has to be translated and protein has to accumulate to reach a detectable level, it is difficult to investigate the subcellular localization before stage 8 (Beumer *et al.*, 1995). However, by assessing the nuclear localization of the EGFP-MI-ER1 fusion proteins at later stages, we were able to determine whether the NLS4 was necessary and sufficient for the nuclear translocation of MI-ER1 in *Xenopus*.

First the subcellular localization of EGFP and EGFP-ER1-493 (full-length) was assessed. The results resembled what was seen in NIH 3T3 cells, EGFP showed whole cell staining (N=C), whereas EGFP-ER1 showed exclusive nuclear staining (Figure 3.16).

Figure 3.16 EGFP-ER339-493, containing NLS4, is translocated to the nuclei of developing *Xenopus laevis* embryos at stage 13.

Xenopus laevis embryos were injected in the marginal zone at the 1-cell stage with RNAs encoding EGFP, EGFP-ER1-493, EGFP-ER1-143, EGFP-ER145-250, EGFP-ER251-340 or EGFP-ER339-493. Subcellular localization of the EGFP-ER1 fusion proteins was determined, after fixation of the embryos at stage 13, by fluorescence microscopy. Embryos injected with EGFP show whole cell staining, EGFP-ER1-493 is exclusively localized to the nucleus, EGFP-ER1-143 shows nuclear greater than cytoplasmic staining, EGFP-ER145-250 and EGFP-ER251-340 show whole cell staining. EGFP-ER339-493 is exclusively nuclear.


EGFP
EGFP-ER1-493
EGFP-ER1-143
EGFP-ER145-250
EGFP-ER251-340
EGFP-ER339-493

To investigate the whether NLS4 was the only (dominant) functional NLS in *Xenopus* embryos, cRNAs encoding the following proteins were injected: EGFP-ER1-143, EGFP-ER145-250, EGFP-ER251-340 and EGFP-ER339-493. At stage 13, the subcellular localization of the EGFP-fusion proteins was compared to EGFP and EGFP-ER1: ER1-143 showed nuclear greater than cytoplasmic staining, ER145-250 and ER251-340 showed cytoplasmic staining and ER339-493 showed exclusively nuclear staining. This corresponds to what was seen after transfection of these regions in NIH 3T3 cells, which showed that only constructs containing the NLS4, EGFP-ER1-493 and EGFP-ER339-493, are exclusively localized to the nucleus

Table 3.7 Subcellular localization of different EGFP constructs after injection into *Xenopus* embryos. Embryos were fixed at stage 13 and viewed under U.V. light (495 nm) with a 10x objective.

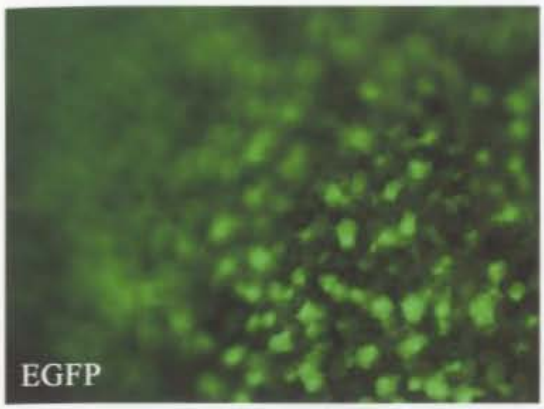
Protein	Subcellular localization
EGFP	Whole cell staining in all cells
EGFP-ER1-493	Nuclear in all cells
EGFP-ER1-144	Nuclear>Cytoplasmic in all cells
EGFP-ER145-250	Whole cell staining in all cells
EGFP-ER251-340	Whole cell staining in all cells
EGFP-ER339-493	Nuclear in all cells
EGFP-ER1-473	Nuclear in all cells
EGFP-ER1-461	Nuclear>Cytoplasmic in all cells
EGFP-ER457-475	Nuclear in all cells

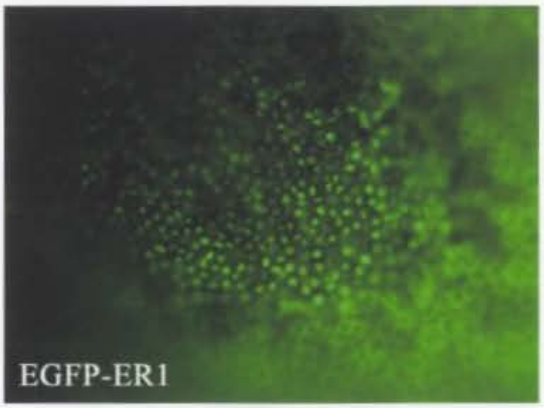
To verify that the NLS4 is necessary and sufficient for nuclear translocation of XMI-ER1 in embryos, additional cRNAs including EGFP-ER1-461 (not incl. NLS4), EGFP-ER1-473 (incl. NLS4), EGFP-ER457-475 (NLS4) and EGFP-ERΔNLS4 were injected into the marginal zone of 1-cell stage embryos. The subcellular localization was determined at stage 13 and compared to EGFP and EGFP-ER1-493. Embryos injected with EGFP-ER1-473 displayed nuclear staining, while embryos injected with EGFP-ER1-461 showed whole cell staining (Figure 3.17). Injection of EGFP-ER457-475 showed nuclear staining (Table 3.7).

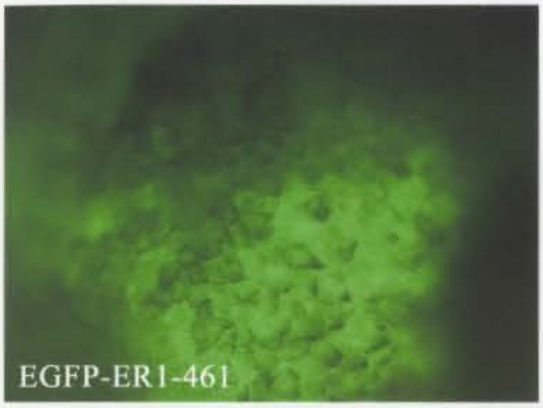
Unfortunately, no fluorescence could be detected in embryos injected with cRNA encoding EGFP-ERΔNLS4, indicating that there is no protein expression, even though embryos from the same batch injected with cRNAs encoding EGFP or EGFP-ER1-493 did express protein in these experiments. However, protein could be detected after *in vitro* translation of pT7Ts-EGFP-ERΔNLS4. Moreover, EGFP-ERΔNLS4 protein could be detected after transfection into NIH 3T3 cells by fluorescence microscopy as well as Western Blot analysis. This indicates that the lack of EGFP-ERΔNLS4 protein expression in *Xenopus* embryos is not a result of an error in the sequence. However, injection of EGFP-ER1-461, which is cytoplasmic, and EGFP-ER1-473, which is nuclear, showed that the NLS4 region was necessary for the nuclear translocation of XMI-ER1.

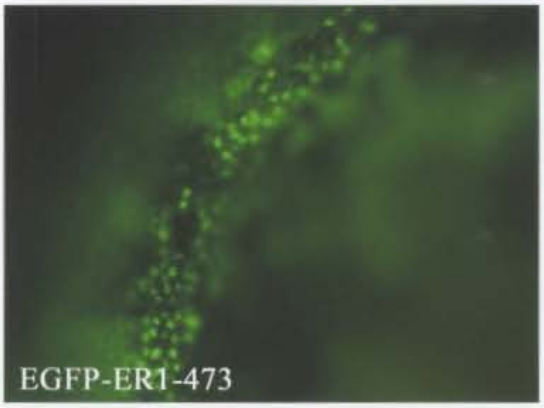
Figure 3.17 NLS4 is necessary and sufficient for the nuclear translocation of MI-ER1 in stage 13 *Xenopus* embryos.

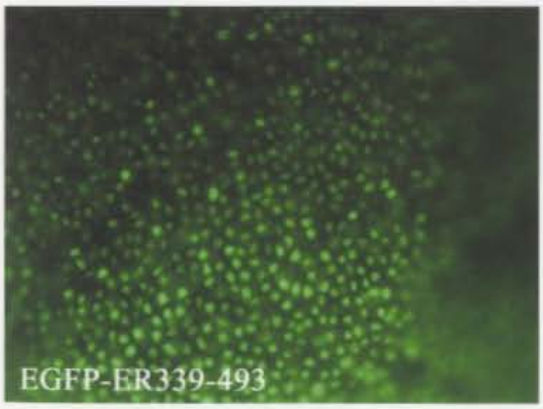
cRNAs encoding EGFP, EGFP-ER1, EGFP-ER1-461, EGFP-ER1-473, EGFP-ER339-493 and EGFP-ER457-475 were injected into the marginal zone of *Xenopus laevis* embryos. Subcellular localization of the EGFP-ER1 fusion proteins was determined by fluorescence microscopy at stage 13. All proteins containing the NLS4: EGFP-ER1, EGFP-ER1-473, EGFP-ER339-493 and EGFP-ER457-475, were localized to the nucleus. EGFP and EGFP-ER1-461 displayed whole cell staining.

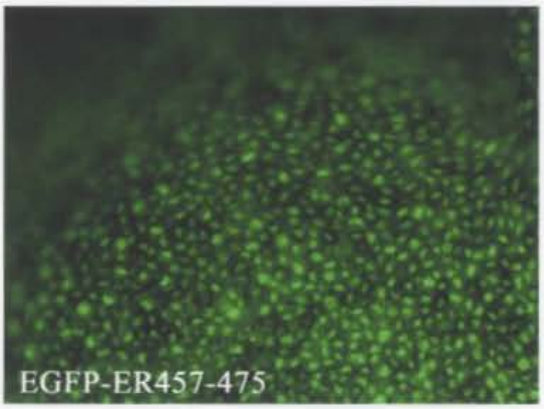

EGFP


EGFP-ER1


EGFP-ER1-461


EGFP-ER1-473


EGFP-ER339-493


EGFP-ER457-475

Thus, injection of these cRNAs into *Xenopus* embryos shows that the NLS4 is necessary and sufficient for the nuclear translocation of XMI-ER1 in *Xenopus laevis* embryos. These results are similar to what was observed after transfection of these constructs in NIH 3T3 cells, which showed NLS4 was necessary and sufficient for the nuclear localization of XMI-ER1. This indicates that the mechanisms for recognition of the XMI-ER1 NLS are present in *Xenopus* embryonic cells as well as in NIH 3T3 cells. Therefore, similar soluble transport factors are present and active in NIH 3T3 cells and *Xenopus* embryos at the stage observed. This suggests that NLS found in different animals are interchangeable, indicating that these sequences and the recognition of these sequences by the import machinery are well conserved during evolution. This is also shown for XNF7, whose bipartite NLS is functional in cells in culture as well as in *Xenopus* embryos (Li *et al.*, 1994b; Li *et al.*, 1994a).

3.9 Conclusion

In this chapter I have shown that full-length XMI-ER1 is actively transported to the nuclei of transfected NIH 3T3 cells as well as in *Xenopus* embryos. The active import of XMI-ER1 into nuclei is dependent on the presence of an intact NLS located near the C-terminus of the protein. The core of this NLS

consists of amino acids ^{463}RPIKRQRMD^{471} but it was shown that addition of amino acids flanking this region results in more efficient nuclear import.

I have also shown that an additional, albeit weak, NLS is present in the N-terminus of XMI-ER1, ^{138}RPRRCK^{143}. Although this NLS, called NLS1, is not efficient in translocating a cytoplasmic protein to the nucleus, definite nuclear greater than cytoplasmic staining of NIH 3T3 cells could be detected when the N-terminus of XMI-ER1 (ER1-143) was fused to EGFP as well as to β-galactosidase. However, ER1-137, which contains no classic NLS, translocated to the nucleus in a smaller (45%) subset of transfected cells, indicating that this region might be translocated by an alternative method, like binding to another nuclear protein (piggy backing) or even by use of an alternate, not previously described NLS.

There are different mechanisms whereby proteins can translocate into the nucleus, either by binding of the NLS directly to import receptors or by binding of the protein to another nuclear protein. Identification of a functional NLS implies a direct binding of XMI-ER1 to import receptors. However, this gives us no information about the temporal regulation of nuclear transport of XMI-ER1 in *Xenopus* embryos. Investigation of the regulation of the timing of nuclear translocation of XMI-ER1 during *Xenopus* embryonic development will be described in chapter 4.

Chapter 4

Regulation of nuclear transport of XMI-ER1 during *Xenopus laevis* development

4.1 Introduction

During oogenesis, a stockpile of structural and regulatory factors accumulates which is used during early developmental stages. As a consequence, many of these factors are present in 100 to 100,000 fold excess compared to somatic cells (Almouzni and Wolffe, 1993a). These include transcriptional activators that are inert during the rapid cell cycling that precedes the mid-blastula transition (MBT). Regulation of nuclear transport of maternally derived transcription factors may provide a mechanism by which the transcriptional machinery is constrained *in vivo* during specific stages in development.

In this chapter, it is shown that the cytoplasmic localization of XMI-ER1 before stage 8 is not a result of lack of NLS recognition and binding by the import machinery. The regulation of nuclear transport of XMI-ER1 during *Xenopus laevis* development is not regulated by phosphorylation of a site close to the NLS, as has been observed for other proteins. Furthermore, cytoplasmic localization of XMI-ER1 before stage 8 in embryos appears to be regulated by binding of XMI-ER1 to an "anchor" in the cytoplasm. The region responsible for

the cytoplasmic localization of XMI-ER1 until stage 8 is present in the N-terminal half of XMI-ER1.

4.2 Nuclear Localization of XMI-ER1 during *Xenopus* development

XMI-ER1 is localized in the cytoplasm of developing *Xenopus* embryos until the embryos reach blastula stage (stage 8), when XMI-ER1 is localized in the nuclei of cells in the marginal zone (Luchman *et al.*, 1999). Over time, more and more nuclei in the animal hemisphere become stained. By late blastula stage (9) XMI-ER1 is found in virtually all nuclei in the animal hemisphere and by early gastrula, stage 10, ubiquitous nuclear staining is observed (Luchman *et al.*, 1999).

In order to compare the temporal patterns of nuclear localization of proteins tested, the exact temporal pattern of nuclear translocation of full-length XMI-ER1 was first determined. This was done by compiling data of embryos injected with EGFP-ER1-493 or β-GAL-ER1-493. Embryos were fixed at different stages during development, followed by whole mount staining using an anti-EGFP antibody or staining with X-gal, respectively. For each stage a minimum of 40 embryos were examined (Table 4.1).

At stage 8, 7% of embryos show some nuclear staining, while cytoplasmic staining is predominant. At stage 8.5, 67% of the embryos injected with XMI-ER1 show some cells with exclusive nuclear staining, while cytoplasmic staining is

also present. At stage 9, almost all embryos show nuclear staining (97%) and at stages 10 and 13 exclusive nuclear staining is observed (100%).

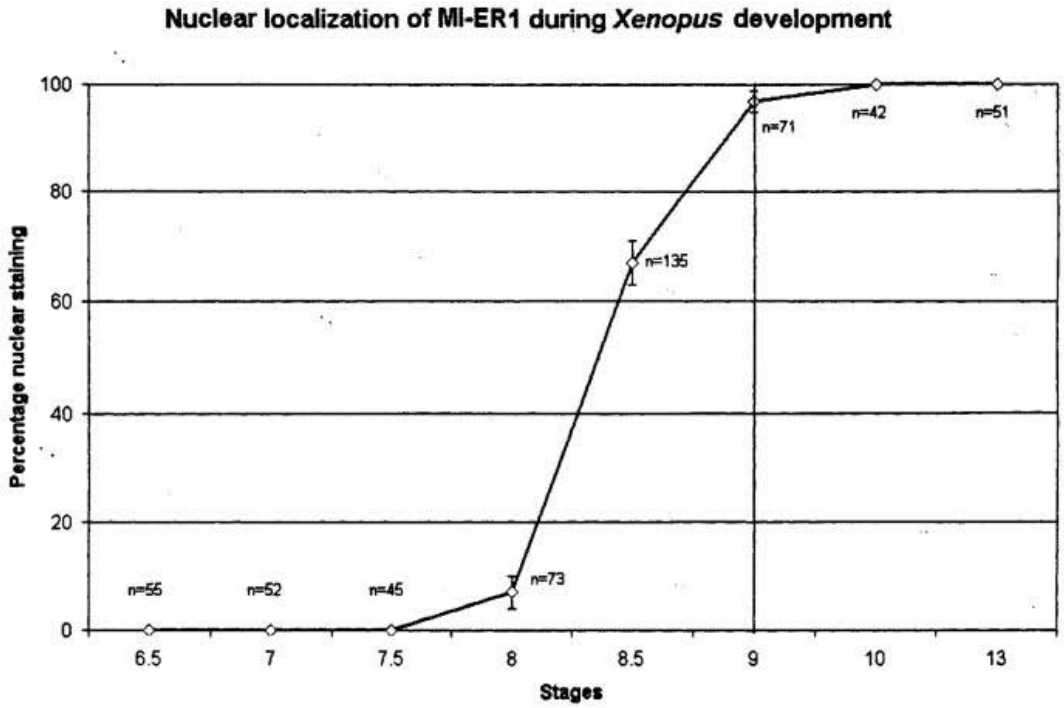
Table 4.1 Temporal pattern of nuclear translocation of XMI-ER1.

Stage	6.5	7	7.5	8	8.5	9	10	13
MI-ER1, % nuclear staining	0	0	0	7	67	97	100	100
Total number of embryos	55	52	45	73	135	71	42	51
SD of percentage	0	0	0	3	4	2	0	0

In table 4.1 and figure 4.1, the percentage of embryos showing nuclear staining at stages 6.5, 7, 7.5, 8, 8.5, 9, 10 and 13 is shown.

When XMI-ER1 is injected in the marginal zone, nuclear staining will be detected earlier than when XMI-ER1 is injected in the animal pole. To avoid misinterpretation of the timing of nuclear entry, all injections were done at the 2-cell stage and embryos were injected in one cell in the animal pole next to the first cleavage furrow. Although there may be slight variations in the development of embryos during different experiments as well as variations in determining the stage by morphology, the graph (figure 4.1) shows a sharp increase in nuclear staining from stage 8 to 9, indicating a sharp threshold. Because exclusive nuclear staining of XMI-ER1 injected cells is observed at stage 9, we define this as the threshold for XMI-ER1 nuclear entry. For our experiments, translocation of proteins before stage 9 is defined as being premature nuclear localization.

Figure 4.1 Nuclear localization of MI-ER1 during *Xenopus* embryonic development. Data of embryos injected with EGFP-ER1-493 or β-GAL-ER1-493 were compiled. The percentage of nuclear localization of ER1-493 was calculated. Nuclear translocation of MI-ER1 begins at stage 8 (7%), and 67% of embryos show nuclear staining in about 50% of the cells in the injection patch at stage 8.5, at stage 9 almost all nuclei (97%) in the injection patch show exclusive nuclear staining. We define stage 9 as the threshold for nuclear entry of MI-ER1. Translocation of proteins before stage 9 is defined as being pre-mature nuclear localization.


Nuclear localization of MI-ER1 during Xenopus development
Percentage nuclear staining
100
80
60
40
20
0
n=55
n=52
n=45
n=73
n=135
n=71
n=42
n=51
6.5
7
7.5
8
8.5
9
10
13
Stages

4.3 β-GAL-NLS4 translocates to the nucleus prematurely

It was shown that in NIH 3T3 cells and in stage 13 *Xenopus* embryos ER457-475 is localized in the nucleus (figures 3.9 and 3.17). This region includes the basic core of NLS4 and 6 amino acid residues on either side of the core NLS. This region is the smallest region tested and was shown to be able to direct EGFP and β-GAL to the nucleus in NIH 3T3 cells as well as *Xenopus* embryos. It is important therefore, to investigate whether this region behaves like full-length XMI-ER1 or whether ER457-475 is able to translocate to the nucleus of developing embryos at earlier stages when nuclear translocation of proteins like nucleoplasmin, N1 and N2 is known to occur (Dreyer, 1987). Therefore, whether or not the XMI-ER1 NLS is functional before stage 9 was investigated by generating a β-GAL-ER457-475 fusion construct and injection of cRNA encoding this protein into albino embryos.

The use of β-GAL fusions over EGFP fusions has a great advantage because of the ease of detection of β-GAL in *Xenopus* embryos. β-GAL proteins are shown to be translated to high concentrations during early stages and are easily detected by staining of the embryo with X-gal (Rupp *et al.*, 1994; Turner and Weintraub, 1994). Therefore, use of β-GAL-ER1 fusion proteins in embryos facilitates detection of subcellular localization of these proteins before stage 8. For all experiments using β-GAL-ER1 fusions, β-GAL and β-GAL-ER1-493 are used as controls; injection of cRNA encoding β-GAL and β-GAL-ER1-493 showed that β-GAL is not localized to nuclei of *Xenopus* cells during any of the

stages observed, namely stage 6.5, 7, 7.5, 8, 8.5 and 9. β-GAL-ER1-493 however, shows nuclear localization in 6% of embryos at stage 8, with increasing nuclear staining as development progresses. At stage 9, 100% of cells expressing β-GAL-ER1-493 show nuclear staining.

Injection of cRNA encoding β-GAL-ER457-475 into embryos showed premature nuclear staining (Figure 4.2, 4.3). Nuclear staining increased as the stages progressed and exclusive nuclear staining was observed in all cells expressing β-GAL-ER457-475 in about 50% of embryos at stage 8.5 (Figure 4.3). At stage 9, 93% of staining is exclusively nuclear.

These results show that the NLS4 is recognized by the import machinery and is able to cause premature translocation of the β-GAL protein. The observation that β-GAL-ER457-475 translocates prematurely suggests that regions outside ER457-475 are important for regulating the temporal translocation of XMI-ER1 to the nuclei of embryonic cells.

Regulation of nuclear transport is often dependent on specific phosphorylation patterns (Jans and Hubner, 1996), for example, XNF7 is localized in the cytoplasm in a hyperphosphorylated state. Dephosphorylation of XNF7 results in release from its cytoplasmic anchor and subsequent nuclear translocation of XNF7 (Li *et al.*, 1994b; Li *et al.*, 1994a). In the next section I investigated whether the cytoplasmic retention of XMI-ER1 is regulated by phosphorylation on a site in the vicinity of the NLS.

Figure 4.2 Injection of MI-ER457-475 (NLS) into *Xenopus* embryos results in pre-mature nuclear translocation.

Embryos were injected with cRNA encoding β-GAL-ER457-475 (D) and the subcellular localization of the fusion protein was compared to β-GAL (B), and β-GAL-ER1-493 (C). β-GAL was localized in the cytoplasm of all stages tested. β-GAL-ER1-493 showed nuclear staining beginning stage 8.5, but was not exclusively nuclear until stage 9. Control embryos were injected with DEPC-treated H_2O and showed no staining at all (A).

Scale bar = 1mm.

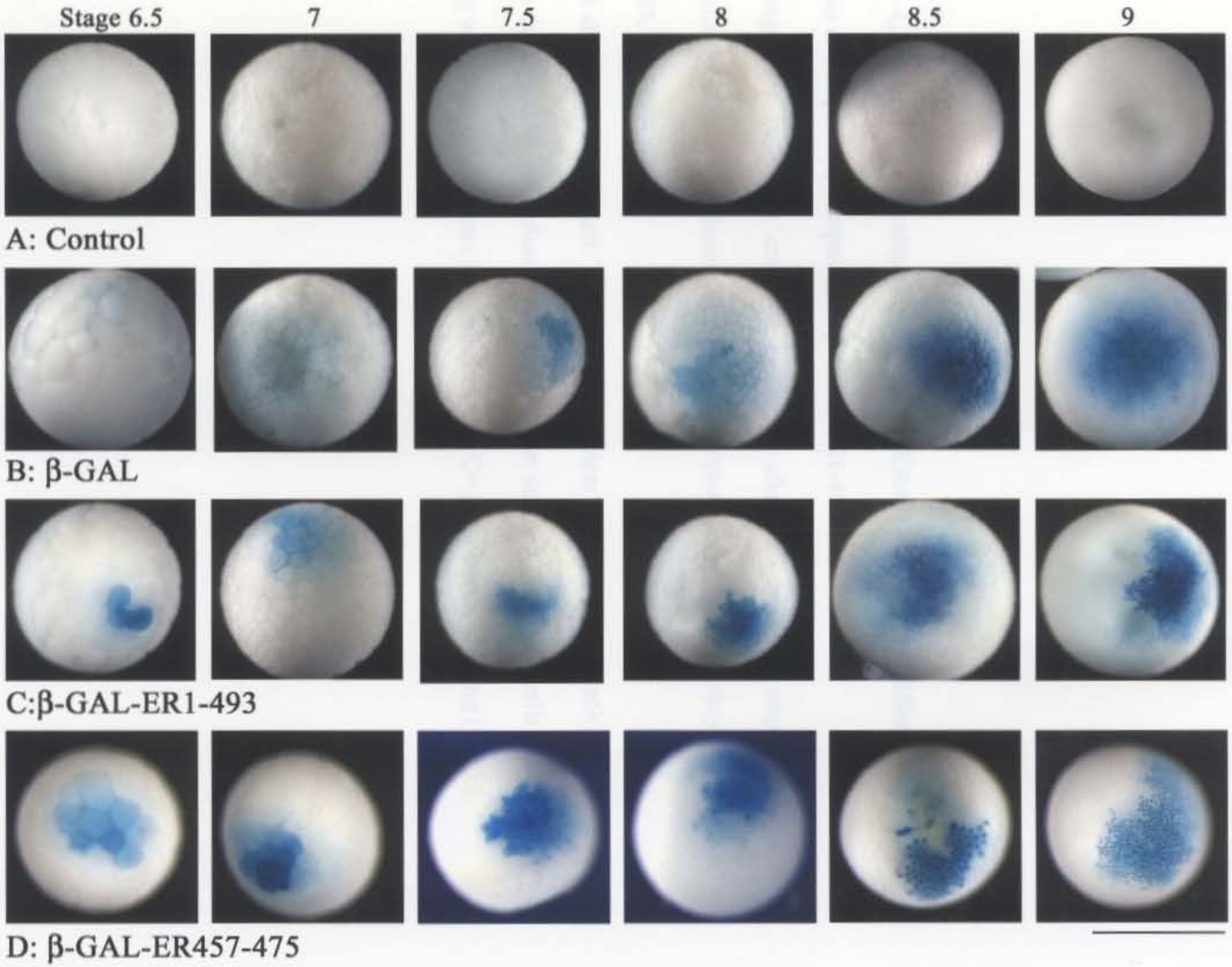
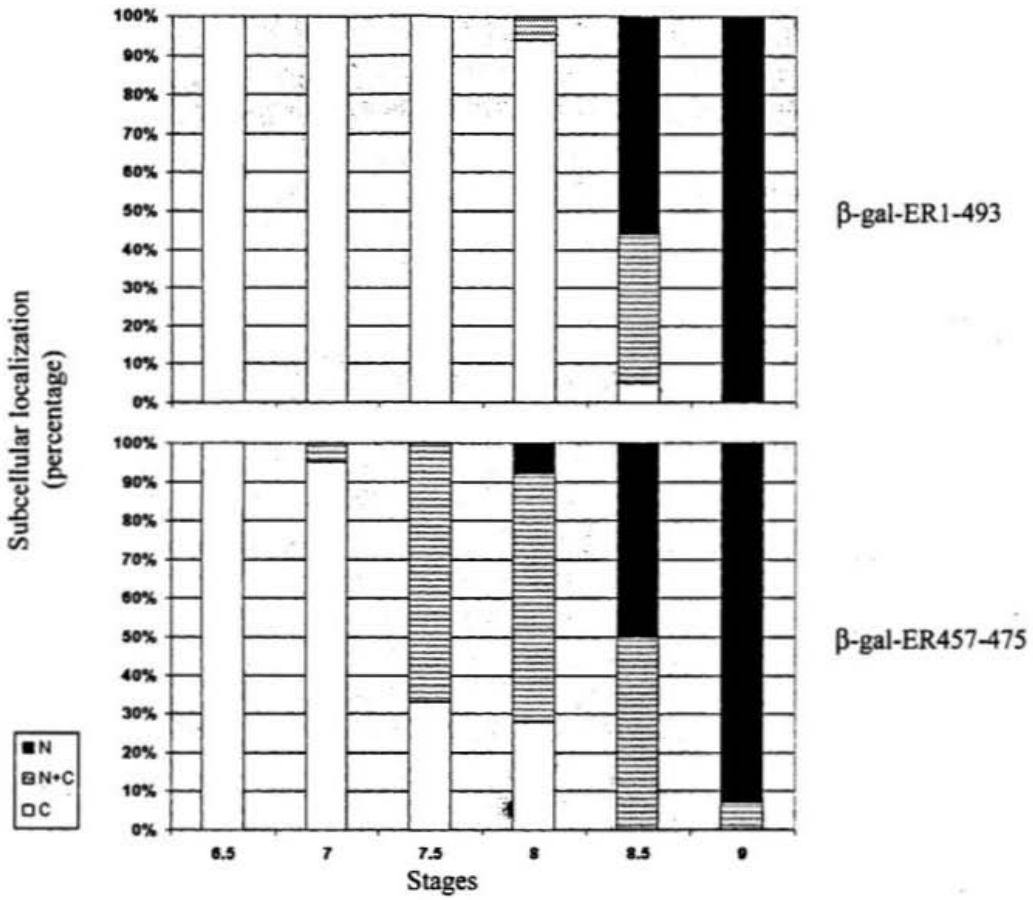

Stage 6.5
7
7.5
8
8.5
9
A: Control
B: β-GAL
C:β-GAL-ER1-493
D: β-GAL-ER457-475

Figure 4.3 Subcellular localization of β-GAL-ER457-475 during different stages of *Xenopus* development, compared to β-GAL-ER1-493.

Percentages are based on embryos examined after two injection experiments and subcellular localization of fusion protein was examined for at least 5 embryos per cRNA, per stage, per experiment.

Legend shown on bottom left: N= exclusive nuclear staining (solid black); N+C= most cells in the injection patch show exclusive nuclear staining, but some cells show nuclear staining with faint cytoplasmic staining (striped); C= cytoplasmic staining (white).


Subcellular localization
(percentage)
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
N
N+C
C
6.5
7
7.5
8
8.5
9
Stages
β-gal-ER1-493
β-gal-ER457-475

4.4 Cytoplasmic retention of XMI-ER1 is not dependent on cdc2 phosphorylation of its CcN motif

Sequences flanking NLSs, which can contain potential phosphorylation sites, can influence regulation of nuclear import (Jans and Hubner, 1996). For example, the SV40 large T-Ag NLS is flanked by two phosphorylation sites, a casein kinase II and a cdc2 kinase site. Together with the NLS this motif is called a CcN box, for Casein kinase, cdc2k and NLS.

Putative phosphorylation sites can be identified by computer analysis of the protein sequence using pbase_predict (www.cbs.dtu.dk/databases/PhosphoBase/predict/predform.html) and NetPhos (www.cbs.dtu.dk/services/NetPhos) (Blom *et al.*, 1999). Analysis of the XMI-ER1 protein sequence using these programs revealed 15 putative CKII sites, of which 4 were in the C-terminal region in the vicinity of the NLS (appendix A1), 1 putative cdc2 kinase site, 8 putative PKA sites as well as putative phosphorylation sites for PKC, CKI and others. Analysis of these putative phosphorylation sites using NetPhos gives an indication of the probability of phosphorylation. For example, the threonine in the putative CKII site KTD^{455}TNDP has a probability of phosphorylation of 0.330 and will probably not be phosphorylated, while the serine in the putative cdc2/PKA site RMD^{472}SPGK has a probability of 0.997 and therefore is more likely to be phosphorylated (appendix A2).

Figure 4.4 Amino acid sequence of XMI-ER1 indicating a putative CcN motif

A CcN motif consists of a CKII site, a cdc2 kinase site and an NLS. XMI-ER1 possesses a putative cdc2 kinase site (underlined in red) and two putative CKII sites (underlined in blue). The cdc2 kinase site, together with one of the CKII sites and the NLS (box) form a putative CcN motif.

1 MAEPSLRTAS PGGSAASDDH EFEPSADMLV HEFDDEQTLE EEEMLEGEVN

51 FTSEIEHLER ESEMPIDELL RLYGYGSTVP LPGEEDEEDM DNDCNSGCSG

101 EIKDEAIKDS SGQEDETQSS NDDPTPSFTC RDVREVIRPR RCKYFDTNHE

151 IEEESEDDED YVPSEDWKKE IMVGSMFQAE IPVGICKYRE TEKVYENDDQ

201 LLWNPEYVME ERVIDFLNEA SRRTCEERGL DAIPEGSHIK DNEQALYEHV

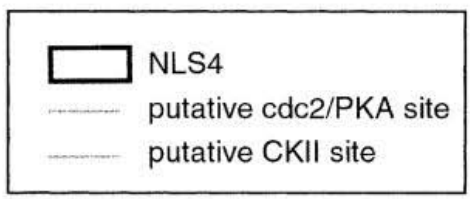
251 KCNFDTEEAL RRLRFNVKAA REELSVWTEE ECRNFEQGLK AYGKDFHLIQ

301 ANKVRTRSVG ECVAFYYMWK KSERYDFFAQ QTRFGKKKYN LHPGVTDYMD

351 RLLDESESAT SSRAPSPPPT TSNSNTSQSE KEDCTASNNT QNGVSVNGPC

401 AITAYKDEAK QGVHLNGPTI SSSDPSSNET DTNGYNRENV TDDSRFSHTS

451 GKTDTNPDDT NERPIKRQRM DSPGKESTGS SEFSQEVFSH GEV*



XMI-ER1 has two putative CKII sites in the vicinity of the NLS, as well as a putative cdc2k site (see figure 4.4). This forms a putative CcN motif and it is important to investigate whether this CcN is functional for the regulation of nuclear import of XMI-ER1 in NIH 3T3 cells as well as *Xenopus* embryos. Since phosphorylation of the CcN motif by cdc2 kinase is known to have an inhibiting effect on nuclear translocation of the SV40 large T-antigen as well as other proteins (Jans and Hubner, 1996; Jans and Jans, 1994), phosphorylation of the CcN motif by cdc2 in XMI-ER1 during early embryonic development could explain cytoplasmic localization of XMI-ER1 during early stages.

To identify whether the cdc2 kinase/ PKA site, C-terminal of the NLS, in XMI-ER1 is important in the regulation of nuclear import, different mutations were made and transfected into NIH 3T3 cells. Substitution of the serine with either alanine, asparagine or aspartic acid had no effect on the nuclear uptake when mutated in the full-length XMI-ER1, ER339-493 or ER457-475 (only S→ A and S→ D), which all showed nuclear localization of the protein. There were some differences in the numbers of nuclear cells, but no consistent pattern was found (see table 4.1 and figures 4.5, 4.6 and 4.7).

Figure 4.5 Substitution of ^{472}serine to alanine (A), asparagine (N) or aspartic acid (D) has no effect on the nuclear localization of EGFP-ER1-493 in NIH 3T3 cells.

Cells were transfected with EGFP-ER1-493 fusion constructs containing serine to alanine (A), asparagine (N) or aspartic acid (D) substitutions. A, Subcellular localization of the fusion proteins was determined by fluorescence microscopy 48 hours post-transfection. Photographs illustrating the subcellular localization of each of the fusion proteins and Hoechst 33342 staining of the same field are shown for each construct. Arrows indicate cells that show subcellular localization of the fusion protein that is most representative of the subcellular localization of the majority of cells expressing the indicated protein and the corresponding nuclei as identified by Hoechst staining. B, Protein expression levels of the EGFP-ER1-493 fusions were determined by Western Blot analysis using an anti-GFP antibody. Molecular weight markers are indicated on the left.

Scale bar = 0.01 mm.

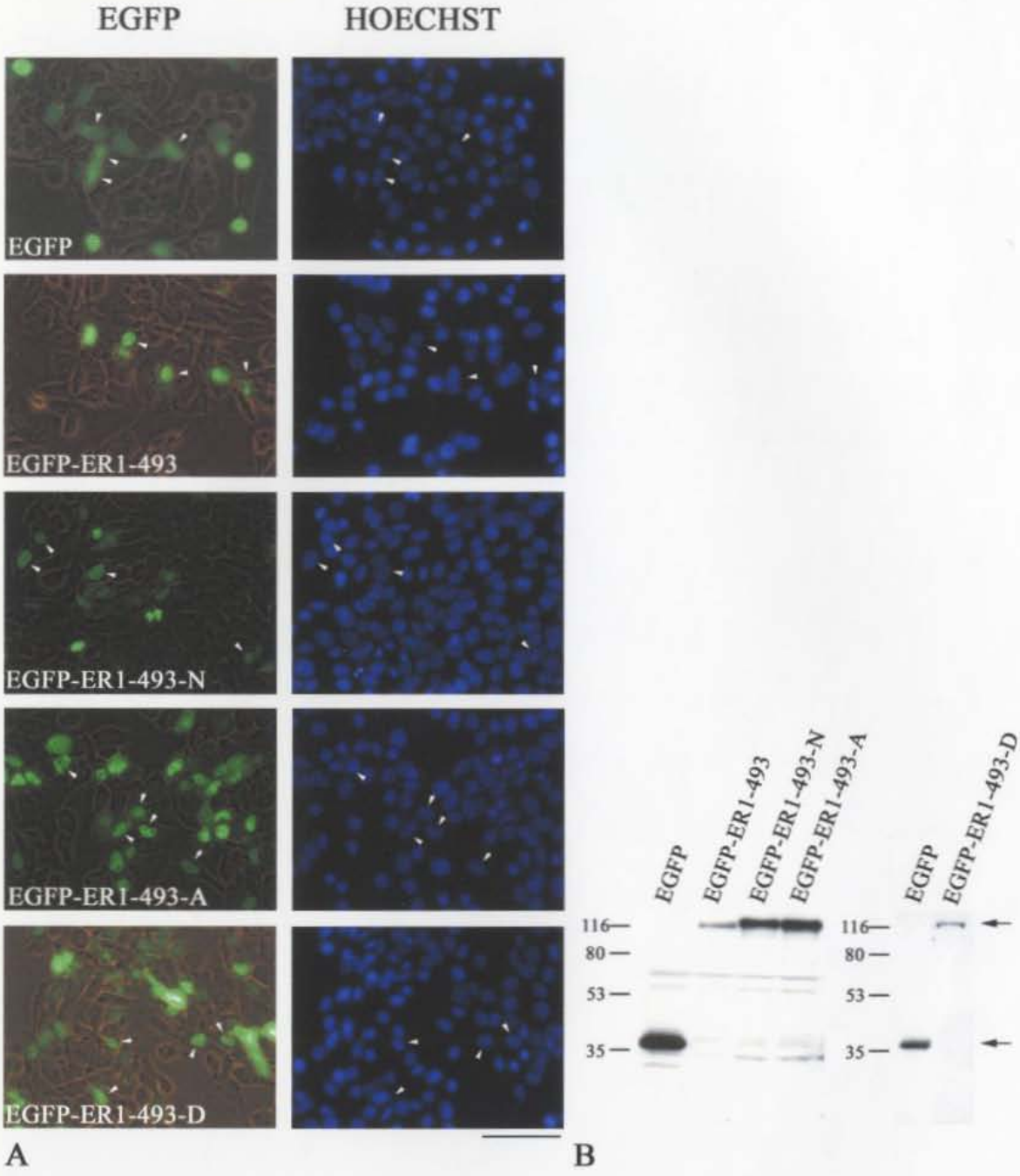

EGFP
HOECHST
EGFP
EGFP-ER1-493
EGFP-ER1-493-N
EGFP-ER1-493-A
EGFP-ER1-493-D
A
EGFP
EGFP-ER1-493
EGFP-ER1-493-N
EGFP-ER1-493-A
EGFP
EGFP-ER1-493-D
116
80
53
35
116
80
53
35
B

Figure 4.6 Substitution of ^{472}serine to alanine (A), asparagine (N) or aspartic acid (D) has no effect on the nuclear localization of EGFP-ER339-493 in NIH 3T3 cells.

Cells were transfected with EGFP-ER339-493 fusion constructs containing serine to alanine, asparagine or aspartic acid substitutions. A, Subcellular localization of the fusion proteins was determined by fluorescent microscopy 48 hours post-transfection. Photographs illustrating the subcellular localization of each of the fusion proteins and Hoechst 33342 staining of the same field are shown for each construct. Arrows indicate cells that show subcellular localization of the fusion protein that is most representative of the subcellular localization of the majority of cells expressing the indicated protein and the corresponding nuclei as identified by Hoechst staining.

B, Protein expression levels of the EGFP-ER339-493 fusions were determined by Western Blot analysis using an anti-GFP antibody (panel B). Molecular weight markers are indicated on the left.

Scale bar = 0.01 mm.

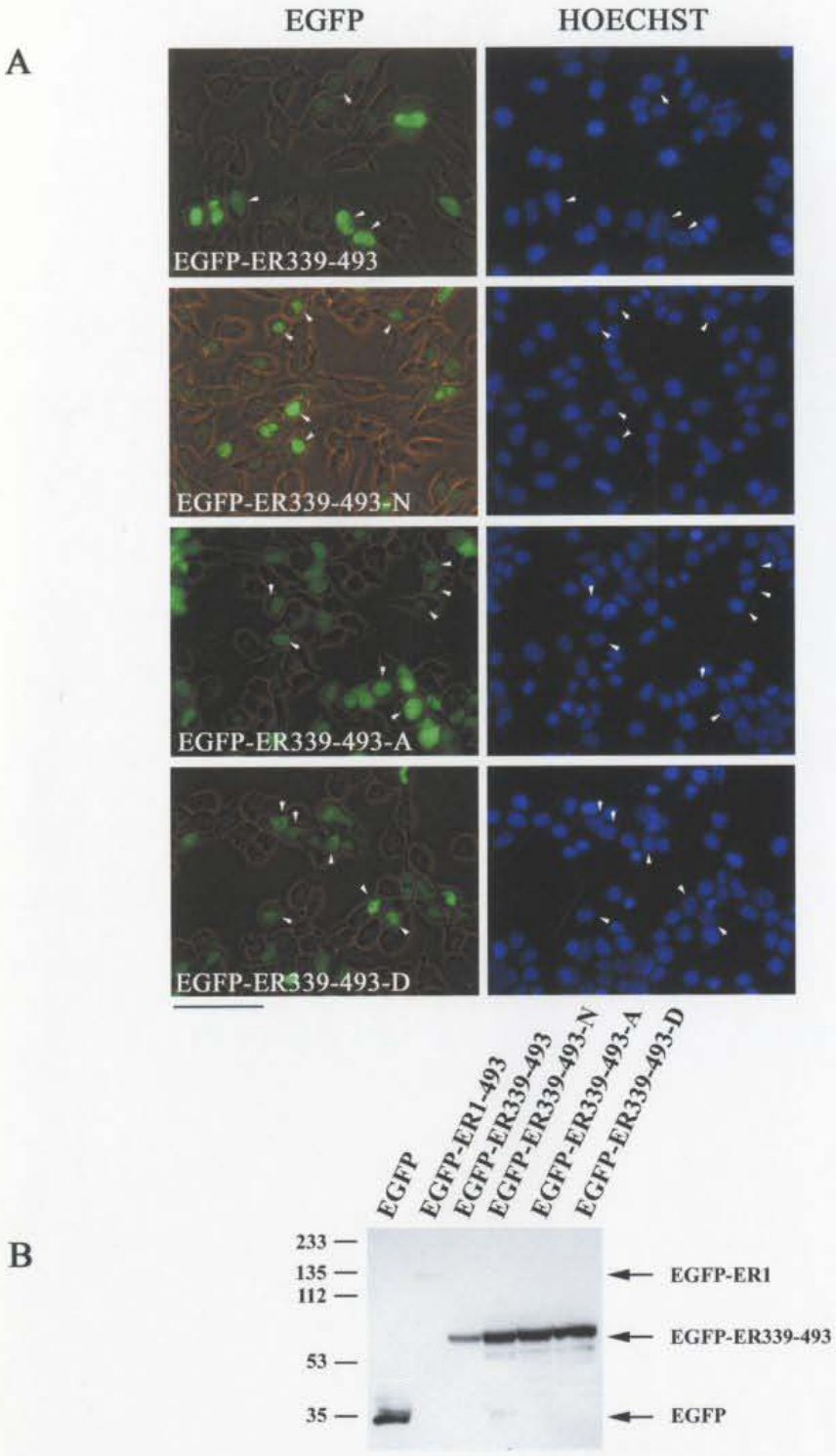

A
EGFP
HOECHST
EGFP-ER339-493
EGFP-ER339-493-N
EGFP-ER339-493-A
EGFP-ER339-493-D
B
EGFP
EGFP-ER1-493
EGFP-ER339-493
EGFP-ER339-493-N
EGFP-ER339-493-A
EGFP-ER339-493-D
233
135
112
53
35
EGFP-ER1
EGFP-ER339-493
EGFP

Figure 4.7 Substitution of ^{472}serine to alanine (A) or aspartic acid (D) has no effect on the nuclear localization of EGFP-ER457-475 in NIH 3T3 cells.

Cells were transfected with EGFP-ER457-475 fusion constructs containing serine to alanine or aspartic acid substitutions. A, Subcellular localization of the fusion proteins was determined by fluorescent microscopy 48 hours post-transfection. Photographs illustrating the subcellular localization of each of the fusion proteins and Hoechst 33342 staining of the same field are shown for each construct. Arrows indicate cells that show subcellular localization of the fusion protein that is most representative of the subcellular localization of the majority of cells expressing the indicated protein and the corresponding nuclei as identified by Hoechst staining. B, Protein expression levels of the EGFP-ER457-475 fusions were determined by Western Blot analysis using an anti-GFP antibody (panel B). Molecular weight markers are indicated on the left.

Scale bar = 0.01 mm.

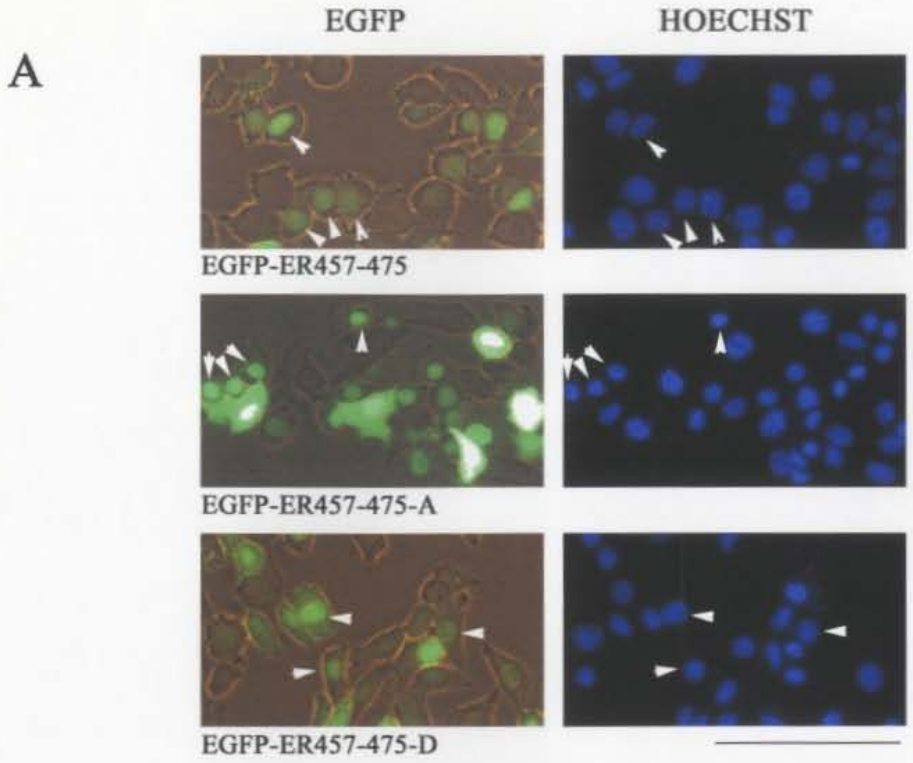

A
EGFP
HOECHST
EGFP-ER457-475
EGFP-ER457-475-A
EGFP-ER457-475-D
B
EGFP
EGFP-ER1
EGFP-ER457-475
EGFP-ER457-475-A
EGFP-ER457-475-D
233
135
112
53
35
EGFP-ER1
EGFP-ER457-475
EGFP

Table 4.2 Transfection of cdc2 kinase/ PKA phosphorylation site mutations in different ER1 constructs into NIH 3T3 cells.

Construct	% N	% N > C	% N = C
EGFP	0	5	95
EGFP-ER1-493	99	1	0
ER1-493 SPGK → APGK	89	7	4
ER1-493 SPGK → NPGK	84	13	3
ER1-493 SPGK → DPGK	99	1	0
ER339-493	97	3	0
ER339-493 SPGK → APGK	92	8	0
ER339-493 SPGK → NPGK	99	1	0
ER339-493 SPGK → DPGK	87	7	6
ER457-475	76	22	2
ER457-475 SPGK → APGK	83	15	2
ER457-475 SPGK → DPGK	71	15	14

This indicates that this putative phosphorylation site is not used for the regulation of nuclear transport of XMI-ER1 in NIH 3T3 cells, which is consistent with our previous observation that XMI-ER1 is constitutively localized to the nucleus in these cells (Paterno *et al.*, 1997). If this phosphorylation site would be important for the regulation of nuclear transport of XMI-ER1 one would expect that the nuclear localization XMI-ER1 in NIH 3T3 cells would be significantly below 100% because some of the XMI-ER1 protein would be in a phosphorylated state that prevents nuclear translocation.

The phosphorylation state of the cdc2/ PKA site in the vicinity of the NLS of XMI-ER1 in embryos might be different from the phosphorylation state of this site in an established cell line like NIH 3T3 cells. The phosphorylation state of many proteins is dependent on activation of cell surface receptors and their

signaling pathways (Karin, 1994). During embryonic development, different signaling pathways are activated at the start of zygotic transcription (Newport and Kirschner, 1982b). Activation of these pathways results in a cascade of phosphorylation of proteins downstream of the receptor, changing phosphorylation states of many proteins. In turn, these changes in phosphorylation patterns before and after the start of zygotic transcription may have a profound effect on the phosphorylation state of XMI-ER1.

In *Xenopus* embryos, a delay in the nuclear uptake of XMI-ER1 is observed until after stage 8, which is around the start of zygotic transcription, at MBT. To investigate whether this putative PKA/ cdc2 site is important for the cytoplasmic retention of XMI-ER1 seven different constructs were generated in pT7Ts, which can be used for *in vitro* RNA transcription: EGFP, EGFP-ER1-493, EGFP-ER1-493 S→A, EGFP-ER1-493 S→D, EGFP-ER339-493, EGFP-ER339-493 S→A and EGFP-ER339-493 S→D. cRNA was transcribed and injected into the animal pole of 2-cell stage albino embryos. The embryos were left to develop and were fixed at stage 7, 8-8.5, 9 and 10 to investigate whether a difference in the timing of nuclear import could be detected. The embryos were stained in a whole mount detection procedure using a GFP antibody (Clontech) and visualized by NBT/BCIP staining. At least ten embryos were examined per RNA per stage. Photographs were taken of the embryo that showed a staining pattern most representative for the subcellular localization of each EGFP-ER1 fusion

protein at the stages indicated. Expressed EGFP-ER1 fusion protein could be detected in embryos beginning at stage 7.

Injection of EGFP-ER1-493 S→ A and EGFP-ER1-493 S→ D, showed similar staining patterns compared to EGFP-ER1-493 (Figure 4.8). No nuclear localization was observed until the beginning of stage 8 and exclusive nuclear staining was observed at stage 9. This indicates that there is no difference in the timing of nuclear translocation of the protein with the alanine or aspartic acid substitution compared to the wild type counterpart. In some embryos background staining was high. For these embryos an enlargement of the injection patch is shown for better analysis of the subcellular localization of the EGFP-ER1 protein (Figure 4.8 B).

Embryos expressing EGFP-ER339-493 S→A and EGFP-ER339-493 S→ D did not show a difference in timing of nuclear translocation compared to EGFP-ER1-493 and EGFP-ER339-493 (data not shown). EGFP-ER339-493 localized to the nucleus at stage 8, the same time as the full length XMI-ER1, indicating that regulation of cytoplasmic localization may be dependent on residues between 339 and 493. The EGFP control was cytoplasmic during all stages. These results indicate that the cdc2/PKA site is not used for the regulation of timing of nuclear import during embryonic development. However, the mechanism of cytoplasmic localization of XMI-ER1 may be dependent on phosphorylation of other putative phosphorylation sites within XMI-ER1.

XMI-ER1 contains many CKII sites, which may be important for the regulation of nuclear transport. The CKII site described for the SV40 T-Ag CcN, for example, is important for the rate of nuclear uptake (Jans and Hubner, 1996; Jans and Jans, 1994). The experimental approach for investigating a role for the putative CKII sites in the temporal regulation of nuclear translocation of XMI-ER1 is microinjection of mutated constructs into cells followed by confocal microscopy. This would allow for the measurement of rates of nuclear import of proteins in live cells. This experimental set-up was not available and therefore the potential function of the CK II sites in XMI-ER1 was not investigated.

I showed that mutation of the putative cdc2/PKA site localized in the C-terminus of XMI-ER1 has no effect on nuclear import of XMI-ER1 in NIH 3T3 cells as well as in *Xenopus* embryonic cells. Furthermore, substitution of this putative cdc2/ PKA site indicated that phosphorylation on this site has no effect on the regulation of the temporal pattern of XMI-ER1 nuclear localization during *Xenopus* development.

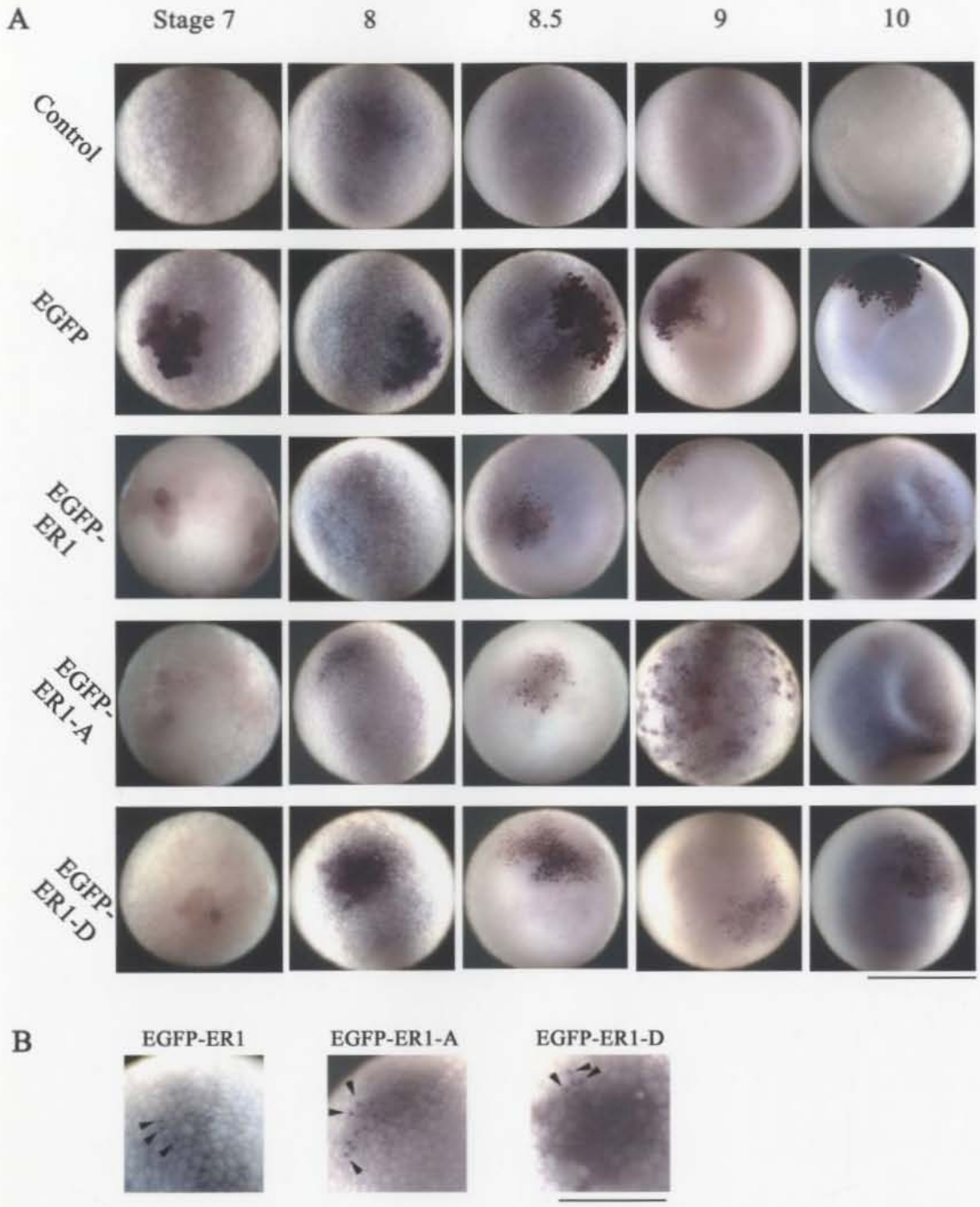
Figure 4.8 Substitution of ^{472}serine with alanine (A) or aspartic acid (D) in EGFP-ER1-493 does not influence the temporal pattern of nuclear import of MI-ER1 during *Xenopus* embryonic development.

A, cRNA encoding the indicated EGFP-ER1-493 fusion proteins was injected into the animal pole of 2-cell stage albino embryos. Embryos were fixed at different stages during development and whole mount staining was performed using an anti-GFP antibody. EGFP is localized in the cytoplasm during all stages tested. EGFP-ER1, EGFP-ER1(A) and EGFP-ER1(D) localize to the nucleus beginning stage 8 and exclusive nuclear localization is observed at stage 9. Photographs were taken of embryos that were the most representative of its group. Results are based on 2 injection experiments. For each injected cRNA at least 5 embryos were examined per stage, per experiment.

Scale bar = 1mm.

B, Enlargement of the injection patch of embryos expressing EGFP-ER1, EGFP-ER1-A and EGFP-ER1-D at stage 8. In all embryos nuclear staining was visible at this stage.

Scale bar = 0.5 mm.


A
Stage 7
8
8.5
9
10
Control
EGFP
EGFP-ER1
EGFP-ER1-A
EGFP-ER1-D
B
EGFP-ER1
EGFP-ER1-A
EGFP-ER1-D

4.5 XMI-ER1 contains a cytoplasmic retention domain, which prevents nuclear translocation before stage 8.

A protein containing a functional NLS can be localized in the cytoplasm as a result of binding to a molecule that prevents it from translocating into the nucleus. This mechanism can be divided into two categories: NLS-masking or binding of the protein to a cytoplasmic anchoring molecule. If the delay in nuclear import of XMI-ER1 in *Xenopus* embryos is due to binding to a cytoplasmic anchor, then addition of an NLS will not influence the timing of translocation. However, if the regulation is due to masking of the NLS by inter- or intramolecular binding, addition of an additional NLS will result in nuclear translocation of XMI-ER1.

To investigate whether the cytoplasmic retention of XMI-ER1 is a result of NLS-masking or binding to a cytoplasmic anchoring protein, the CS2+-nβgal vector was used, which contains the SV40 large T-Ag NLS fused to the N-terminus of the β-galactosidase cDNA, and CS2+-cβgal, which does not contain the SV40 NLS. The SV40 T-Ag NLS is shown to be a very strong NLS that is able to translocate large cytoplasmic proteins to the nuclei of cells in culture (Lanford *et al.*, 1986; Ribbeck and Gorlich, 2001). β-GAL was chosen for the fusion constructs because of the easy and very rapid detection of the β-GAL-ER1 fusion proteins, within 3 hours after fixation of the embryos. Embryos injected with nβ-gal cRNA, did not show nuclear staining before stage 9, which is simultaneously with β-GAL- ER1 injected embryos (data not shown).

Investigation as to whether cytoplasmic localization is caused by NLS-masking or cytoplasmic anchoring, is based on the ability of the additional NLS to cause premature translocation. The SV40 T-Ag NLS is not able to direct β-GAL to the nucleus prematurely. Therefore a different NLS was used to investigate whether XMI-ER1 is localized in the cytoplasm as a result of NLS-masking or as a result of binding to a cytoplasmic anchor.

Investigation of the literature on nuclear transport in *Xenopus* embryos shows that most proteins able to translocate before the start of zygotic transcription possess a bipartite NLS, like nucleoplasmin, N1 and N2 (Boulikas, 1994; Boulikas, 1997; Dreyer, 1987). This suggests that bipartite NLS can be used to translocate proteins into the nucleus prematurely, well before MBT in *Xenopus* embryos. One well-described protein, which possesses a bipartite NLS, is the zinc-finger protein, XNF7. XNF7 is retained in the cytoplasm of developing *Xenopus* embryos due to binding of its "cytoplasmic retention domain" (CRD) to an "anchor" protein in the cytoplasm. When the CRD of XNF7 was deleted, the protein was detected in the nucleus as early as stage 7 (Li *et al.*, 1994a).

Therefore, we used the XNF7 bipartite NLS fused to XMI-ER1 to test whether XMI-ER1 was retained by cytoplasmic anchoring or NLS masking. RNAs of β-GAL, β-GAL-ER1, β-GAL-XNF7NLS and β-GAL-XNF7NLS-ER1 were injected into 2-cell stage albino embryos and the subcellular localization of the corresponding proteins was examined at stage 6.5, 7, 7.5, 8, 8.5 and 9 by X-gal staining of whole embryos (Figure 4.9 and 4.10 and 4.11). The stages of the

embryos to be examined were selected to make a distinction between proteins that translocated at the same time as full length XMI-ER1, which showed nuclear staining at stage 9, and proteins that showed nuclear staining before that time, i.e. premature translocation of the protein.

At stage 7 to 7.5 the embryos injected with β-GALXNF7-NLS began showing nuclear staining, which increased as the embryo developed and exclusive nuclear staining was observed at stage 8 (Figure 4.9 and 4.10 and 4.11). In the embryos injected with β-GAL-ER1 and β-GAL-XNF7NLS-ER1, however, nuclear staining was detected beginning around stage 8.5 and staining was not exclusively nuclear until late stage 9 (see figures 4.9 and 4.10 and 4.11). The β-GAL control was always located in the cytoplasm. There was no background staining for embryos injected with DEPC-treated H_2O.

Although the XNF7NLS was able to translocate β-GAL to the nuclei prematurely, addition of the XNF7NLS to full-length MI-ER1 did not; nuclear staining was first detected at stage 8.5 and MI-ER1 was not exclusively nuclear until late during stage 9. This suggests that XMI-ER1 contains a cytoplasmic retention domain and is retained in the cytoplasm by binding of this region to an anchor in the cytoplasm during early stages of *Xenopus* embryonic development, with release of the protein and subsequent nuclear translocation during stages 8-9. Investigation of the location of the possible CRD within XMI-ER1 was performed and is described in the following sections.

Figure 4.9 MI-ER1 is retained in the cytoplasm by binding to a cytoplasmic anchor. Embryos were injected at the 2-cell stage with the RNAs indicated. At different stages during development embryos were fixed, followed by staining with X-gal to detect the subcellular localization of the β-GAL fusion proteins. The subcellular localization of β-GAL-XNF7NLS (D) and β-GAL-XNF7NLS-ER1 (E) was compared to β-GAL (B) and β-GAL-ER1 (C). Embryos injected with DEPC-treated H_2O did not show staining (A). Scale bar = 1 mm.

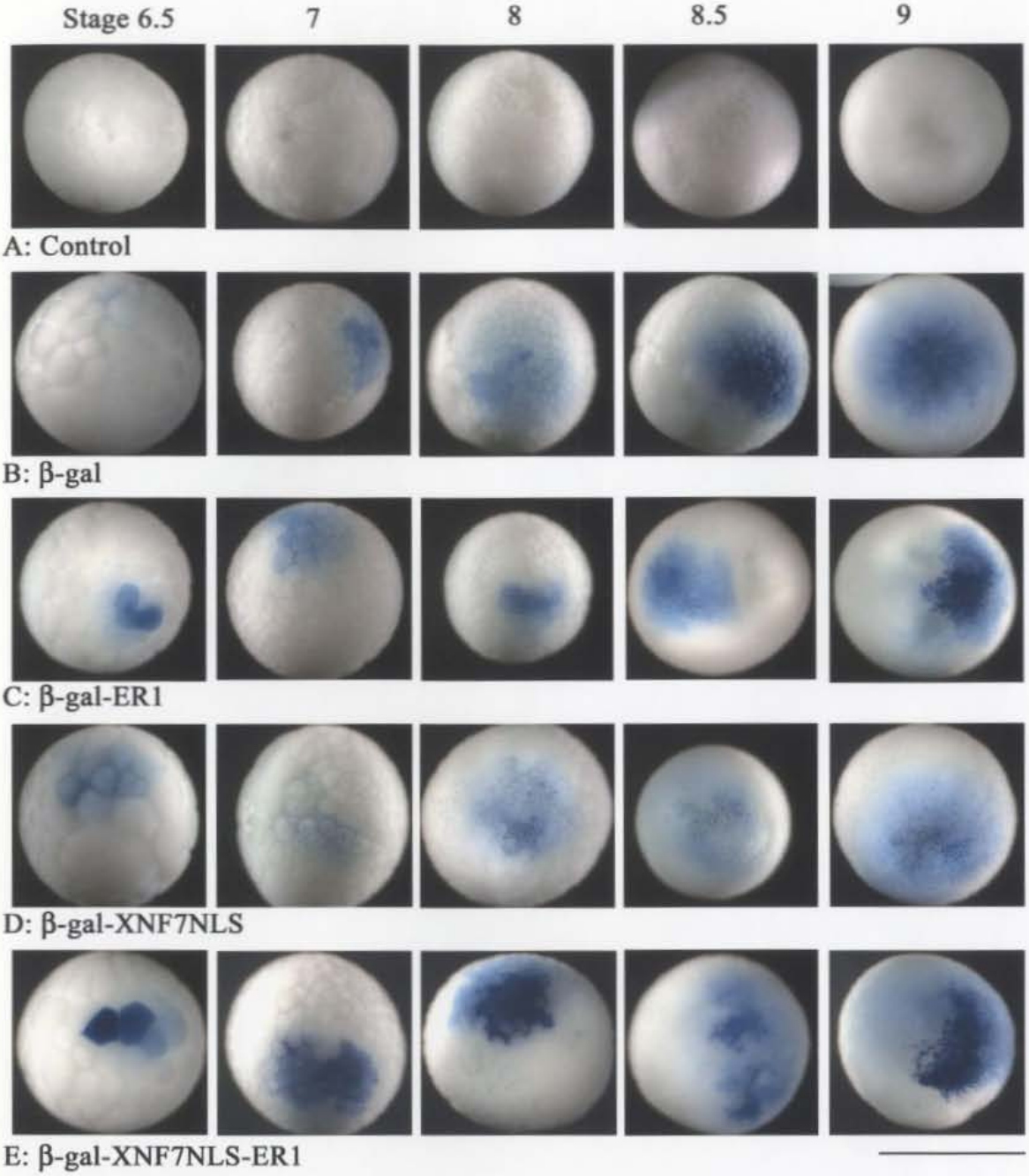

Stage 6.5
7
8
8.5
9
A: Control
B: β-gal
C: β-gal-ER1
D: β-gal-XNF7NLS
E: β-gal-XNF7NLS-ER1

Figure 4.10 Subcellular localization of β-GAL-ER1, β-GAL-XNF7NLS and β-GAL-XNF7NLS-ER1 in *Xenopus laevis* embryos.

Enlargement of the embryos shown in figure 4.9. Stages 6.5, 7, 8, 8.5 and 9 are shown. A, β-GAL-ER1. B, β-GAL-XNF7NLS. C, β-GAL-XNF7NLS-ER1.

Scale bar = 0.5 mm.

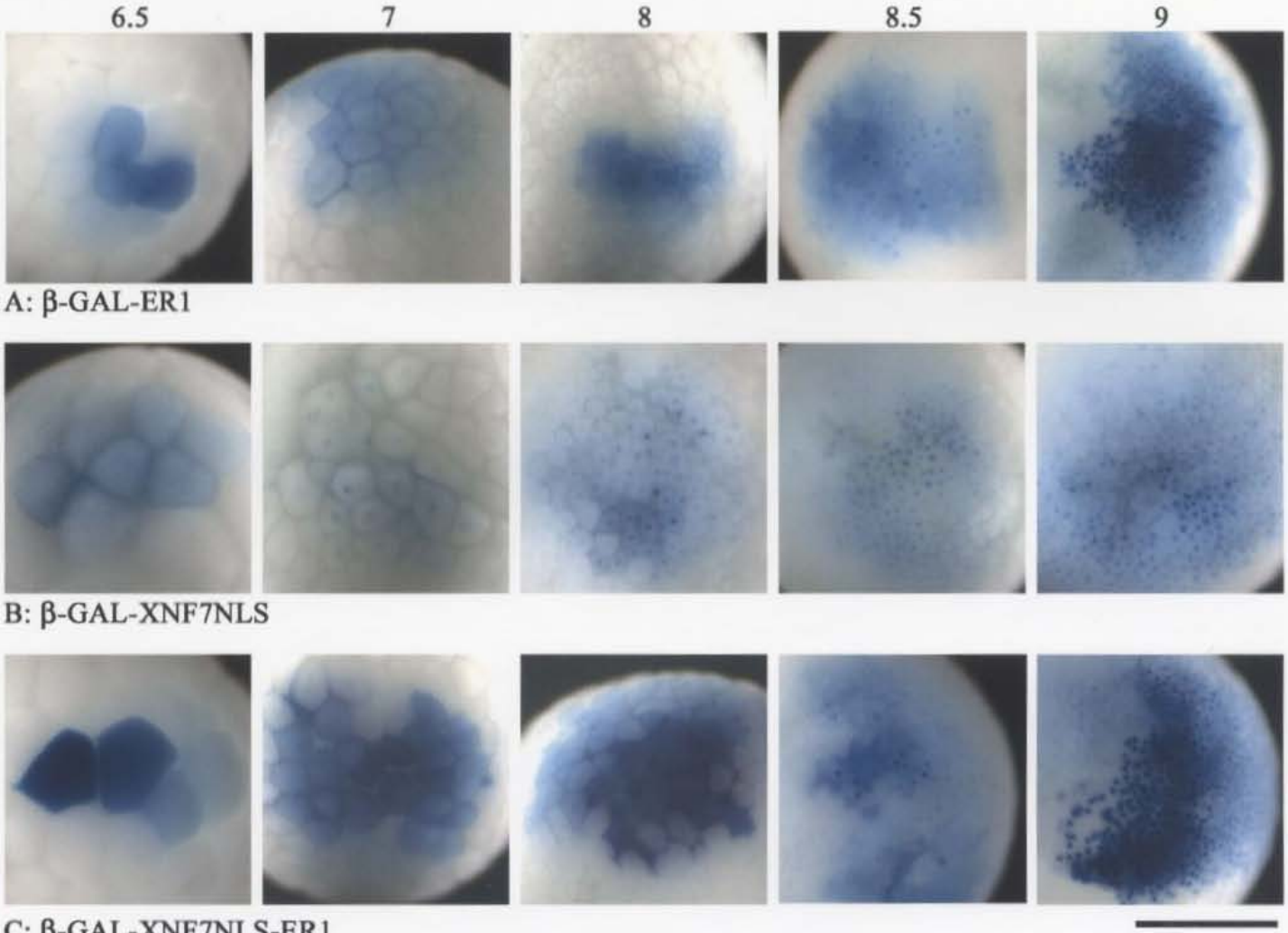
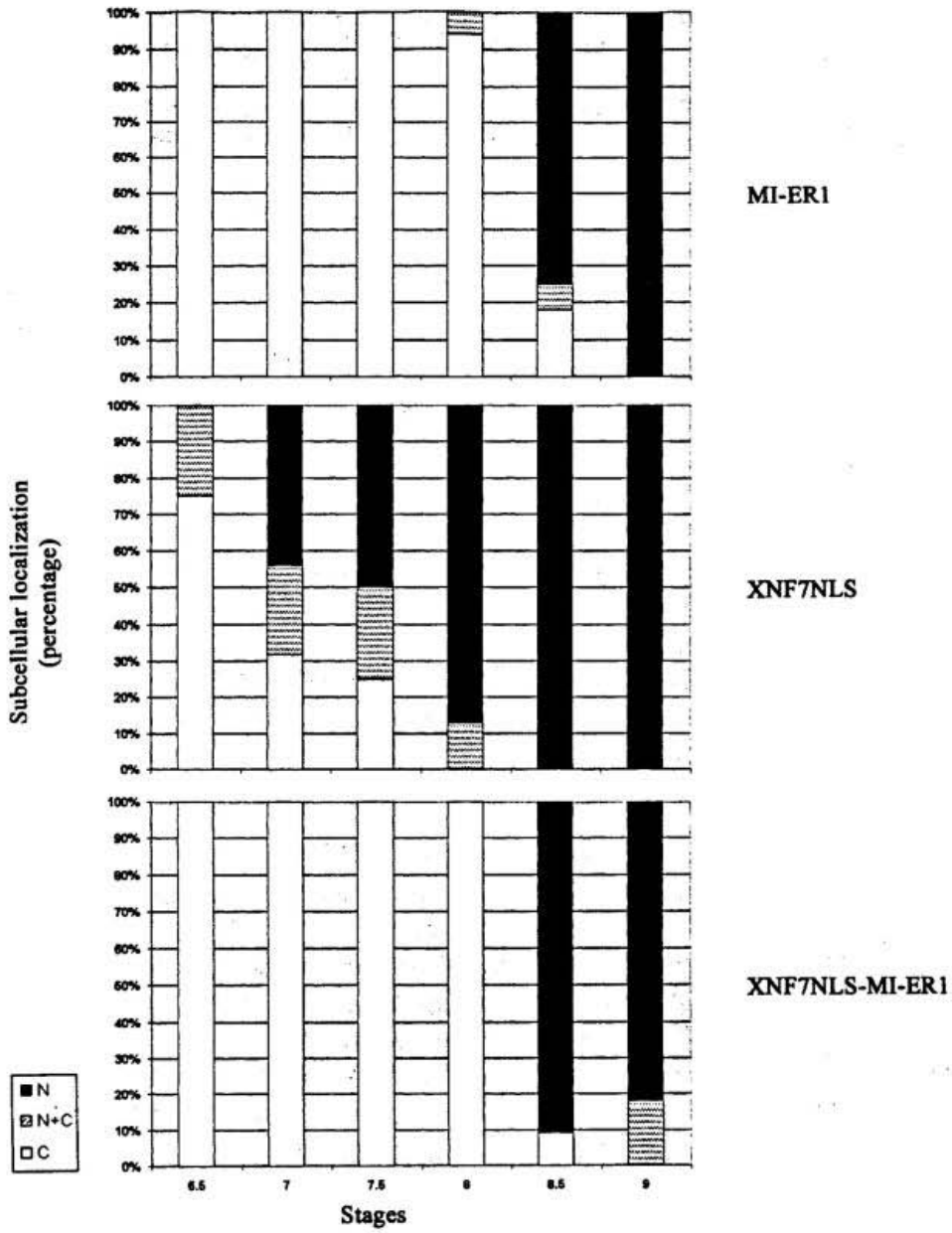

6.5
7
8
8.5
9
A: β-GAL-ER1
B: β-GAL-XNF7NLS
C: β-GAL-XNF7NLS-ER1

Figure 4.11 Subcellular localization of β-GAL- ER1, β-GAL-XNF7NLS and β-GAL-XNF7NLS-ER1 during different stages of *Xenopus* development

Percentages are based on embryos examined after two injection experiments and subcellular localization of fusion protein was examined for at least 5 embryos per cRNA, per stage, per experiment.

Legend shown on bottom left: N= exclusive nuclear staining (solid black); N+C= most cells in the injection patch show exclusive nuclear staining, but some cells show nuclear staining with faint cytoplasmic staining (striped); C= cytoplasmic staining (white).


MI-ER1
XNF7NLS
XNF7NLS-MI-ER1
Subcellular localization (percentage)
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
N
N+C
C
6.5
7
7.5
8
8.5
9
Stages

4.6 XMI-ER1 has a retention domain in its N-terminus

To investigate whether the cytoplasmic retention domain of ER-1 is localized in the N-terminal or C-terminal half of the protein, two β-GAL-XNF7NLS-ER1 fusions were generated. The first comprised amino acids 1-282, which included the acidic domains, the weak NLS (NLS1) and the ELM2 domain. The second fusion was generated by addition of amino acids 267-493, which contains the SANT domain, the proline rich region and the NLS (NLS4). cRNA was generated and injected into 2-cell stage albino embryos. At stages 6.5, 7, 7.5, 8, 8 .5 and 9, the embryos were fixed and stained to identify the subcellular localization of the fusion protein. Stages 6.5, 7, 8, 8.5 and 9 are shown in figure 4.12i and ii. As controls β-GAL-XNF7NLS and β-GAL-XNF7NLS-ER1-493 were used. As shown in figures 4.12i, ii, and 4.13 the XNF7NLS-ER267-493 translocation to the nucleus was premature, similar to that observed for the embryos injected with β-GAL-XNF7-NLS. However, XNF7NLS-ER1-282 was not nuclear until stage 8.5, showing a similar pattern to the β-GAL-XNF7-NLS-ER1-493. This indicates that the possible CRD is located in the N-terminal half of XMI-ER1.

To narrow down the region important for the cytoplasmic retention of XMI-ER1 before stage 8 in embryos, XMI-ER1 was divided into smaller regions, as described below.

Figure 4.12i MI-ER1 possesses a cytoplasmic retention domain in its N-terminal half.

Embryos were injected with cRNA encoding β-GAL-XNF7NLS, β-GAL-XNF7NLS-ER1, β-GAL-XNF7NLS-ER1-282 and β-GAL-XNF7NLS-ER267-493 at the 2-cell stage. The subcellular localization of the fusion proteins was determined by X-Gal staining at different stages during development. The subcellular localization of β-GAL-XNF7NLS-ER1-282 (C) and β-GAL-XNF7NLS-ER267-493 (D) at stages 6.5, 7 and 8 was compared to β-GAL-XNF7NLS-ER1 (A) and β-GAL-XNF7NLS (B).

Scale bar = 1mm.

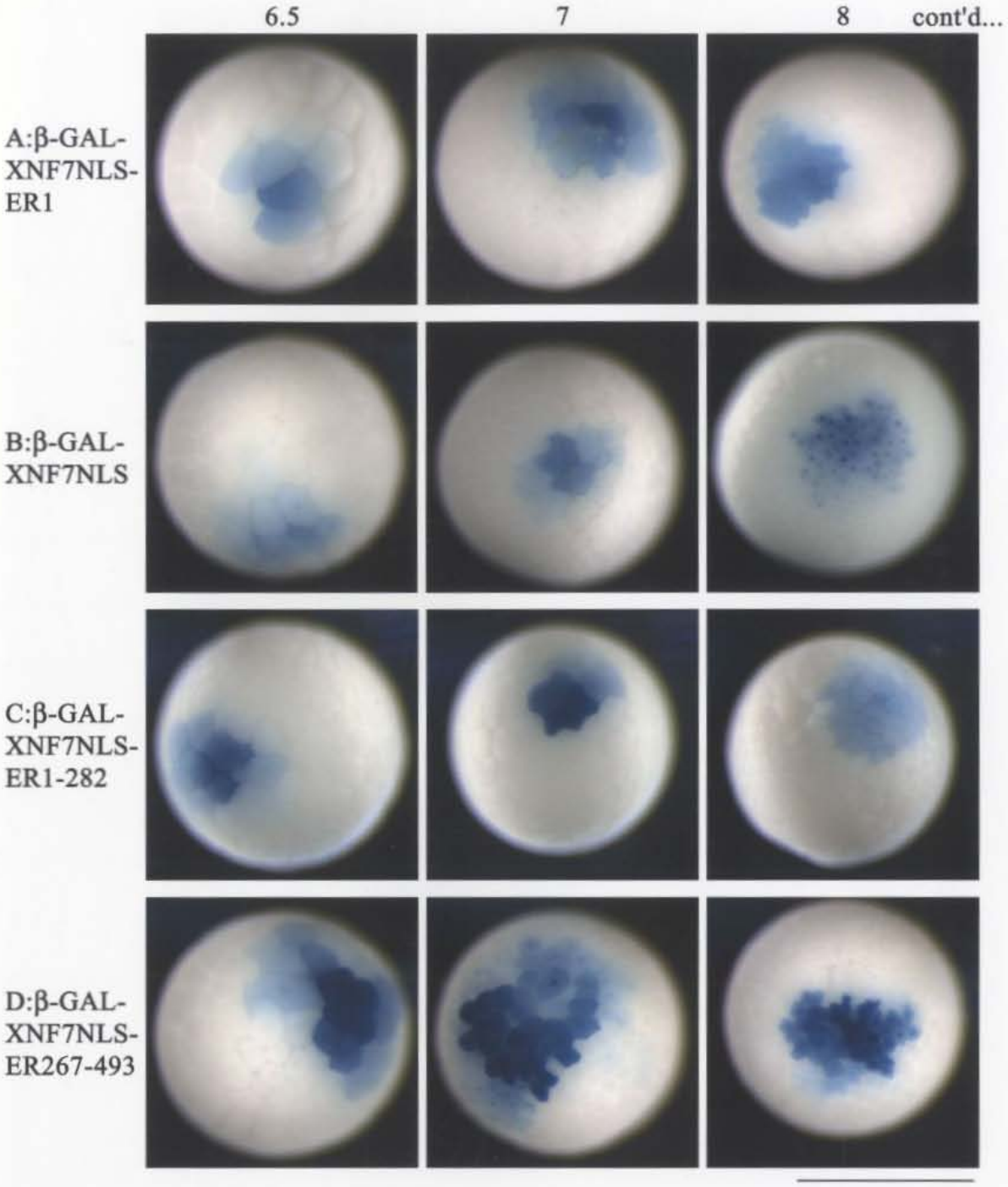

6.5
7
8
cont'd...
A:β-GAL-XNF7NLS-ER1
B:β-GAL-XNF7NLS
C:β-GAL-XNF7NLS-ER1-282
D:β-GAL-XNF7NLS-ER267-493

Figure 4.12ii MI-ER1 possesses a cytoplasmic retention domain in its N-terminus.

Embryos were injected with cRNA encoding β-GAL-XNF7NLS, β-GAL-XNF7NLS-ER1, β-GAL-XNF7NLS-ER1-282 and β-GAL-XNF7NLS-ER267-493 at the 2-cell stage. The subcellular localization of the fusion proteins was determined by X-Gal staining at different stages during development. The subcellular localization of β-GAL-XNF7NLS-ER1-282 (C) and β-GAL-XNF7NLS-ER267-493 (D) at stages 8, 8.5 and 9 was compared to β-GAL-XNF7NLS-ER1 (A) and β-GAL-XNF7NLS (B).

Scale bar = 1mm.

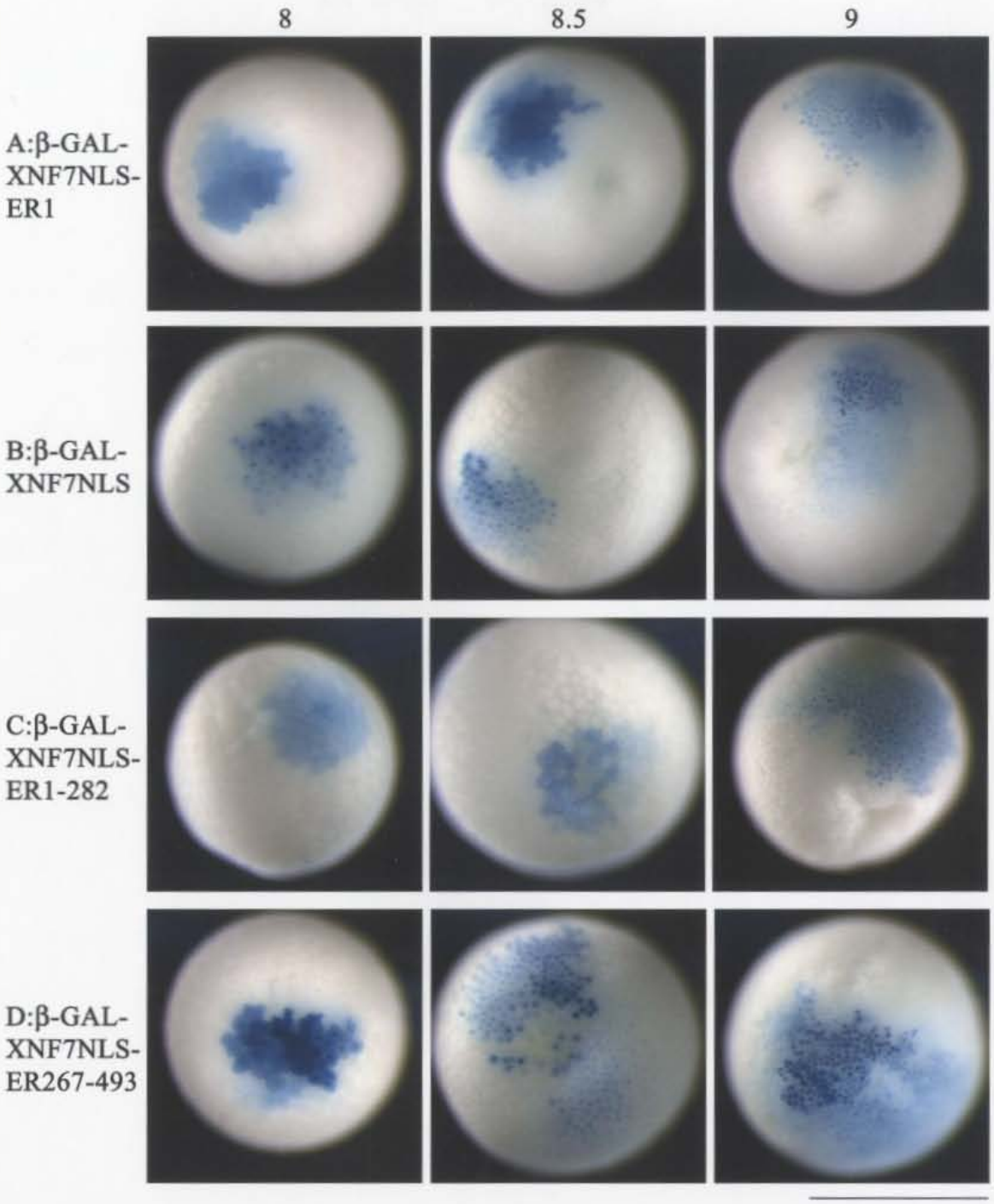
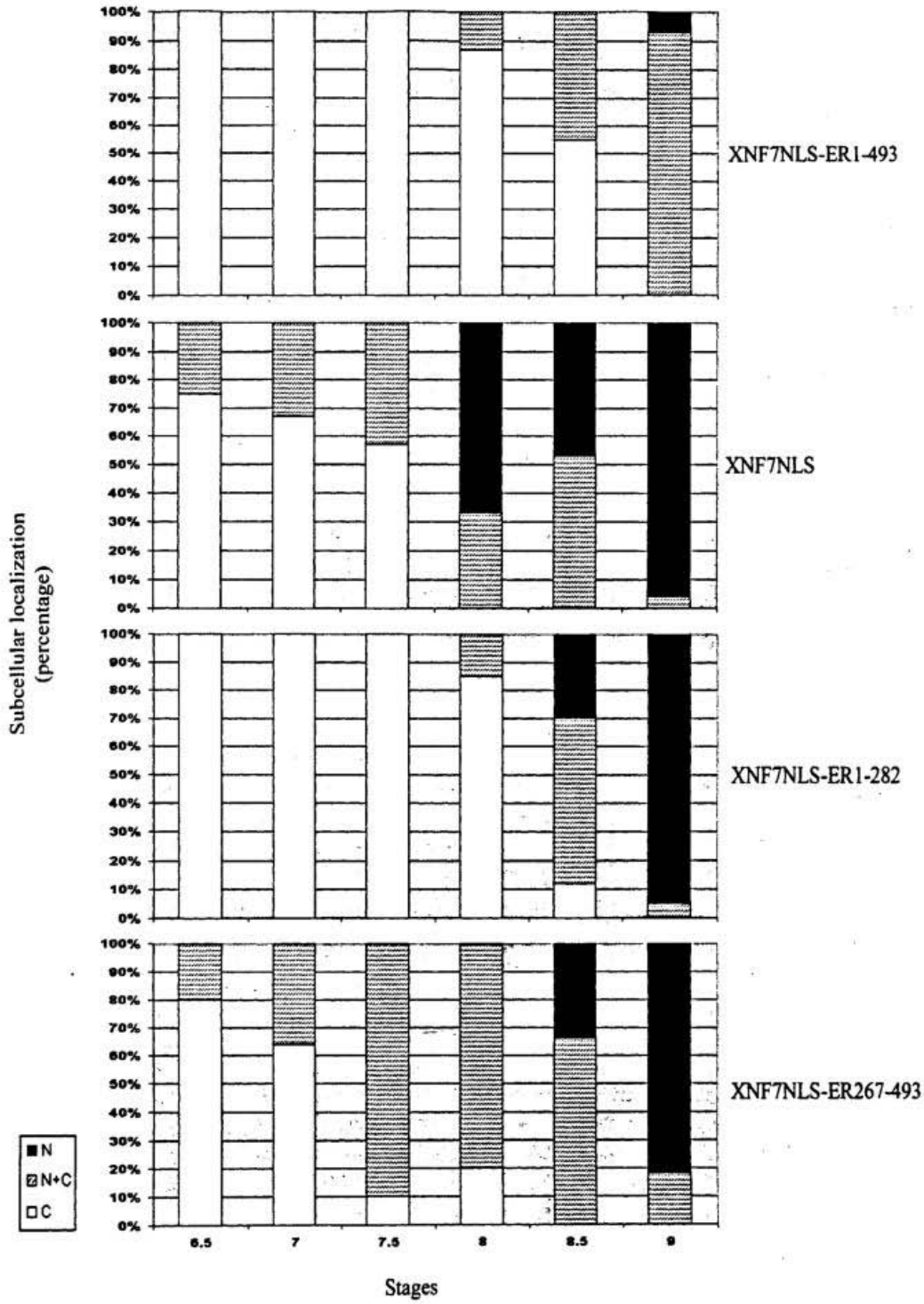

8
8.5
9
A:β-GAL-XNF7NLS-ER1
B:β-GAL-XNF7NLS
C:β-GAL-XNF7NLS-ER1-282
D:β-GAL-XNF7NLS-ER267-493

Figure 4.13 Subcellular localization of β-GAL-XNF7NLS-ER1-282 and β-GAL-XNF7NLS-ER267-493 during different stages of *Xenopus* development, compared to β-GAL-XNF7NLS-ER1 and β-gal-XNF7NLS at these stages.

Stages shown are 6.5, 7, 7.5, 8, 8.5 and 9. Percentages are based on embryos examined after two injection experiments and subcellular localization of fusion protein was examined for at least 5 embryos per cRNA, per stage, per experiment.

Legend shown on bottom left: N= exclusive nuclear staining (solid black); N+C= most cells in the injection patch show exclusive nuclear staining, but some cells show nuclear staining with faint cytoplasmic staining (striped); C= cytoplasmic staining (white).


Subcellular localization (percentage)
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
XNF7NLS-ER1-493
XNF7NLS
XNF7NLS-ER1-282
XNF7NLS-ER267-493
N
N+C
C
6.5
7
7.5
8
8.5
9
Stages

4.7 Attempts to identify the minimal region necessary for retention of XMI-ER1 during *Xenopus* development

To narrow down the region necessary for retention of XMI-ER1 during Xenopus development, cRNAs of β-GAL-XNF7NLS fusions containing different regions of ER1 were injected into 2-cell stage embryos. Since the N-terminal half of XMI-ER1 was shown to contain the cytoplasmic retention domain, this region was divided into smaller regions. To be complete, the regions used span the complete XMI-ER1 sequence. The regions used were: ER1-143, ER145-250, ER251-340 and ER339-493. Embryos were left to develop to stages 7, 7.5, 8, 8.5 or 9 and were then examined for subcellular localization of the fusion protein. The subcellular localization of β-GAL-XNF7NLS-ER1-143, β-GAL-XNF7NLS-ER145-250, β-GAL-XNF7NLS-ER251-340 and β-GAL-XNF7NLS-ER339-493 was compared to the localization of β-GAL, β-GAL-ER1-493, β-GAL-XNF7NLS, and β-GAL-XNF7NLS-ER1-493.

Nuclear staining of XNF7NLS-ER1-143 and XNF7NLS-ER251-340 was observed in embryos beginning stage 7 (figures 4.14, 4.15, 4.16 and 4.17). Nuclear staining was observed beginning stage 7.5 for XNF7NLS-ER145-250 and XNF7NLS-ER339-493 (figures 4.14, 4.15, 4.16 and 4.17). Thus, nuclear translocation of XNF7NLS-ER1-143, XNF7NLS-ER145-250, XNF7NLS-ER251-340 and XNF7NLS-ER339-493 was premature. This suggests that the possible CRD may be split in these smaller regions.

There may be many possible motifs responsible for the cytoplasmic retention of XMI-ER1 during development. However, no known CRD could be identified in the N-terminal region. All regions smaller than ER1-282 (ER1-143, ER145-250 and ER251-340) translocated to the nucleus prematurely, therefore the smallest region necessary for cytoplasmic retention of XMI-ER1 identified was ER1-282.

Figure 4.14 Addition of the XNF7NLS to ER1-143, ER145-250, ER251-340 and ER339-493 results in pre-mature nuclear translocation.

cRNA encoding β-GAL-XNF7NLS, β-GAL-XNF7NLS-ER1, β-GAL-XNF7NLS-ER1-143, β-GAL-XNF7NLS-ER145-250, β-GAL-XNF7NLS-ER251-340 and β-GAL-XNF7NLS-ER339-493 were injected into 2-cell stage *Xenopus* embryos. The embryos were fixed at different stages during development and β-GAL fusion proteins were stained with X-Gal. Embryos most representative of its group at stages 7, 7.5, 8 and 8.5 are shown. The subcellular localization of β-GAL-XNF7NLS-ER1-143 (C), β-GAL-XNF7NLS-ER145-250 (D), β-GAL-XNF7NLS-ER251-340 (E) and β-GAL-XNF7NLS-ER339-493 (F) was compared to β-GAL-XNF7NLS (A) and β-GAL-XNF7NLS-ER1 (B).

Scale bar = 1mm.

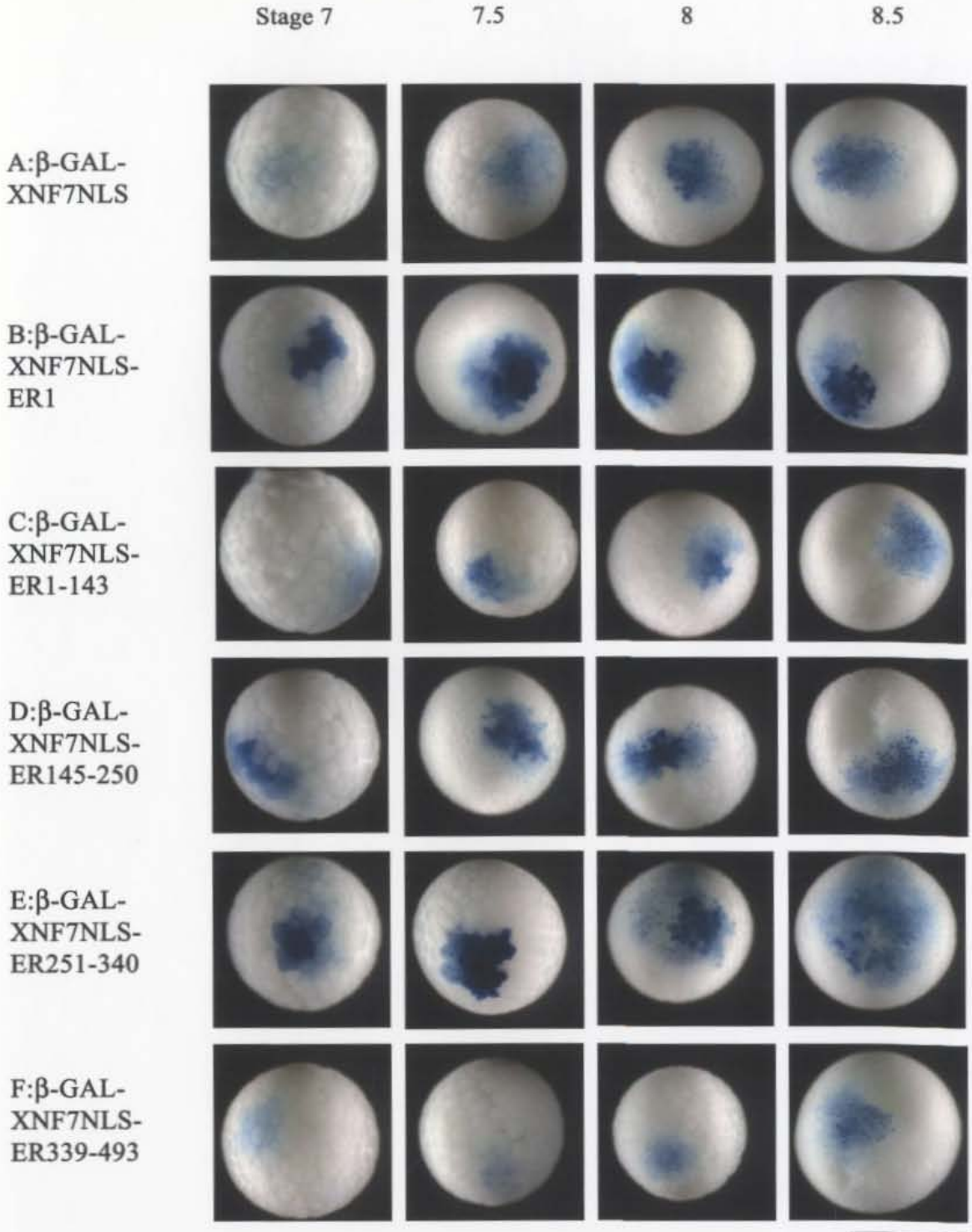

Stage 7
7.5
8
8.5
A:β-GAL-XNF7NLS
B:β-GAL-XNF7NLS-ER1
C:β-GAL-XNF7NLS-ER1-143
D:β-GAL-XNF7NLS-ER145-250
E:β-GAL-XNF7NLS-ER251-340
F:β-GAL-XNF7NLS-ER339-493

Figure 4.15 Addition of the XNF7NLS to ER1-143, ER145-250, ER251-340 or ER339-493 results in pre-mature nuclear translocation.

Enlargement of embryos shown in figure 4.14. Stages 7, 7.5, 8 and 8.5 are shown.

Scale bar = 0.5 mm.

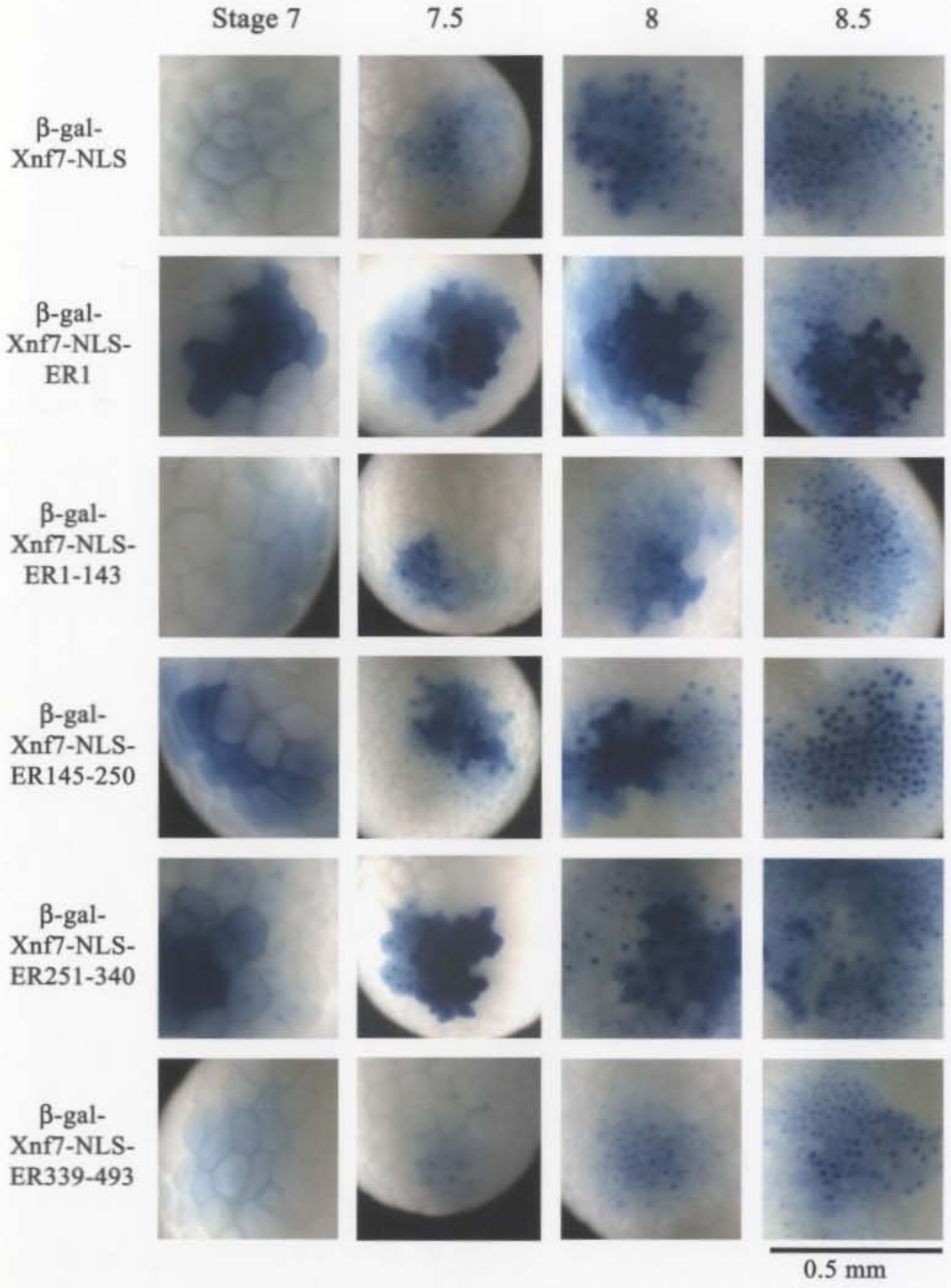

Stage 7
7.5
8
8.5
β-gal-Xnf7-NLS
β-gal-Xnf7-NLS-ER1
β-gal-Xnf7-NLS-ER1-143
β-gal-Xnf7-NLS-ER145-250
β-gal-Xnf7-NLS-ER251-340
β-gal-Xnf7-NLS-ER339-493
0.5 mm

Figure 4.16 Subcellular localization of XNF7NLS-ER1-143 and XNF7NLS-ER145-250, compared to XNF7NLS-ER1 and XNF7NLS.

Stages shown are 7, 7.5, 8, 8.5 and 9. Percentages are based on embryos examined after two injection experiments and subcellular localization of fusion protein was examined for at least 5 embryos per cRNA, per stage, per experiment.

Legend shown on bottom left: N= exclusive nuclear staining (solid black); N+C= most cells in the injection patch show exclusive nuclear staining, but some cells show nuclear staining with faint cytoplasmic staining (striped); C= cytoplasmic staining (white).

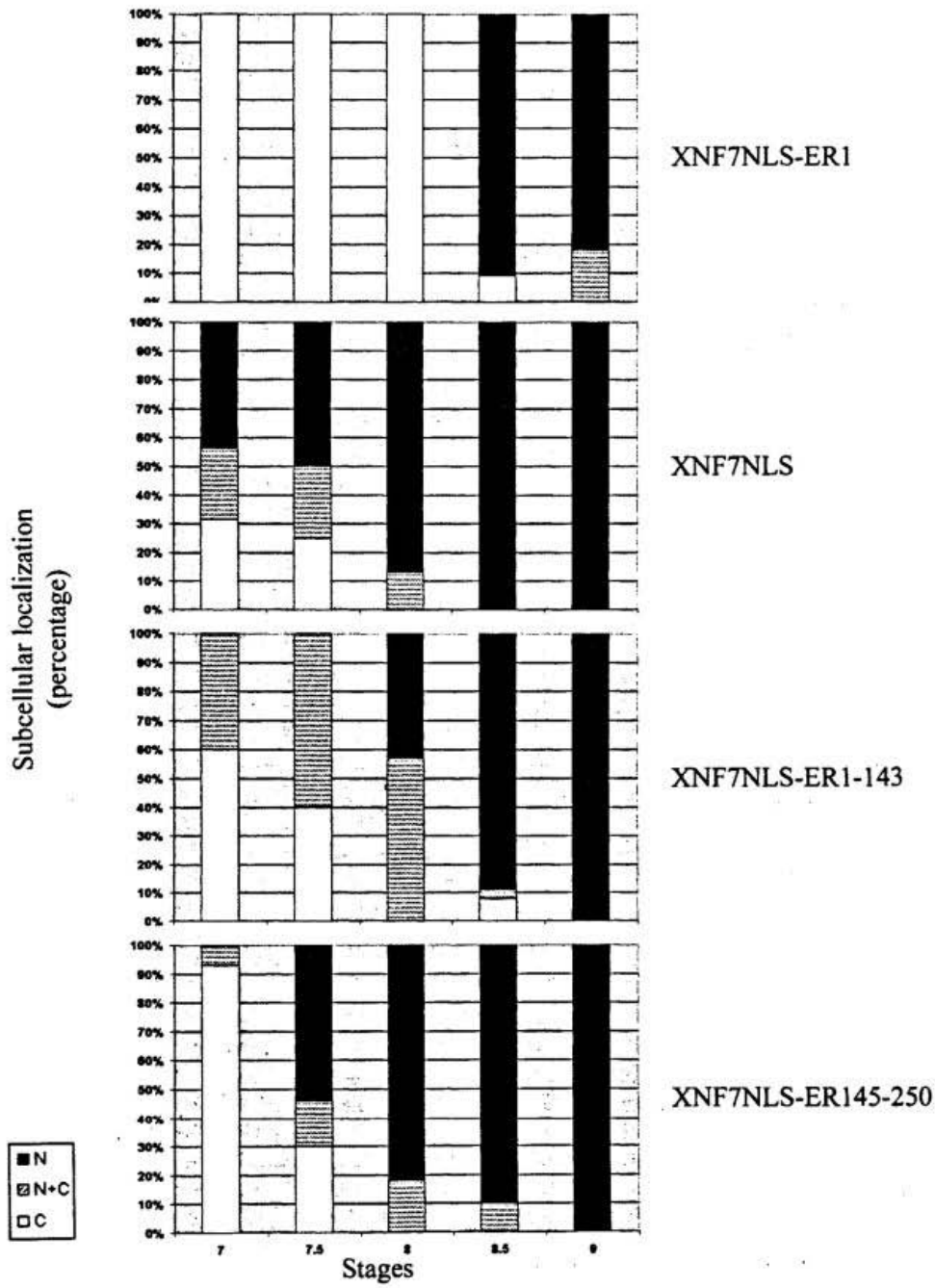

Subcellular localization
(percentage)
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
XNF7NLS-ER1
XNF7NLS
XNF7NLS-ER1-143
XNF7NLS-ER145-250
N
N+C
C
7
7.5
8
8.5
9
Stages

Figure 4.17 Subcellular localization of XNF7NLS-ER251-340 and XNF7NLS-ER339-493, compared to XNF7NLS-ER1 and XNF7NLS

Stages shown are 7, 7.5, 8, 8.5 and 9. Percentages are based on embryos examined after two injection experiments and subcellular localization of fusion protein was examined for at least 5 embryos per cRNA, per stage, per experiment.

Legend shown on bottom left: N= exclusive nuclear staining (solid black); N+C= most cells in the injection patch show exclusive nuclear staining, but some cells show nuclear staining with faint cytoplasmic staining (striped); C= cytoplasmic staining (white).

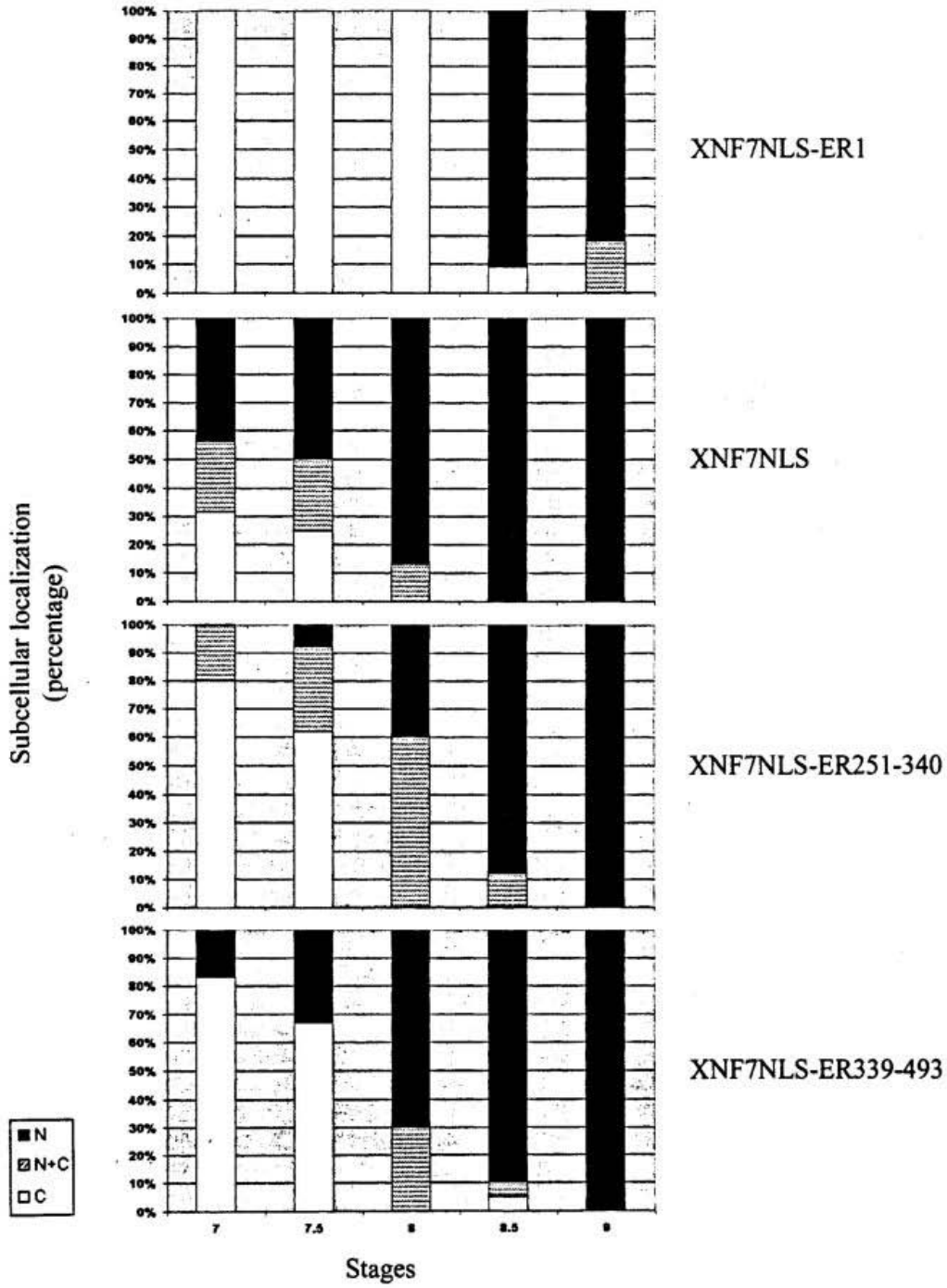
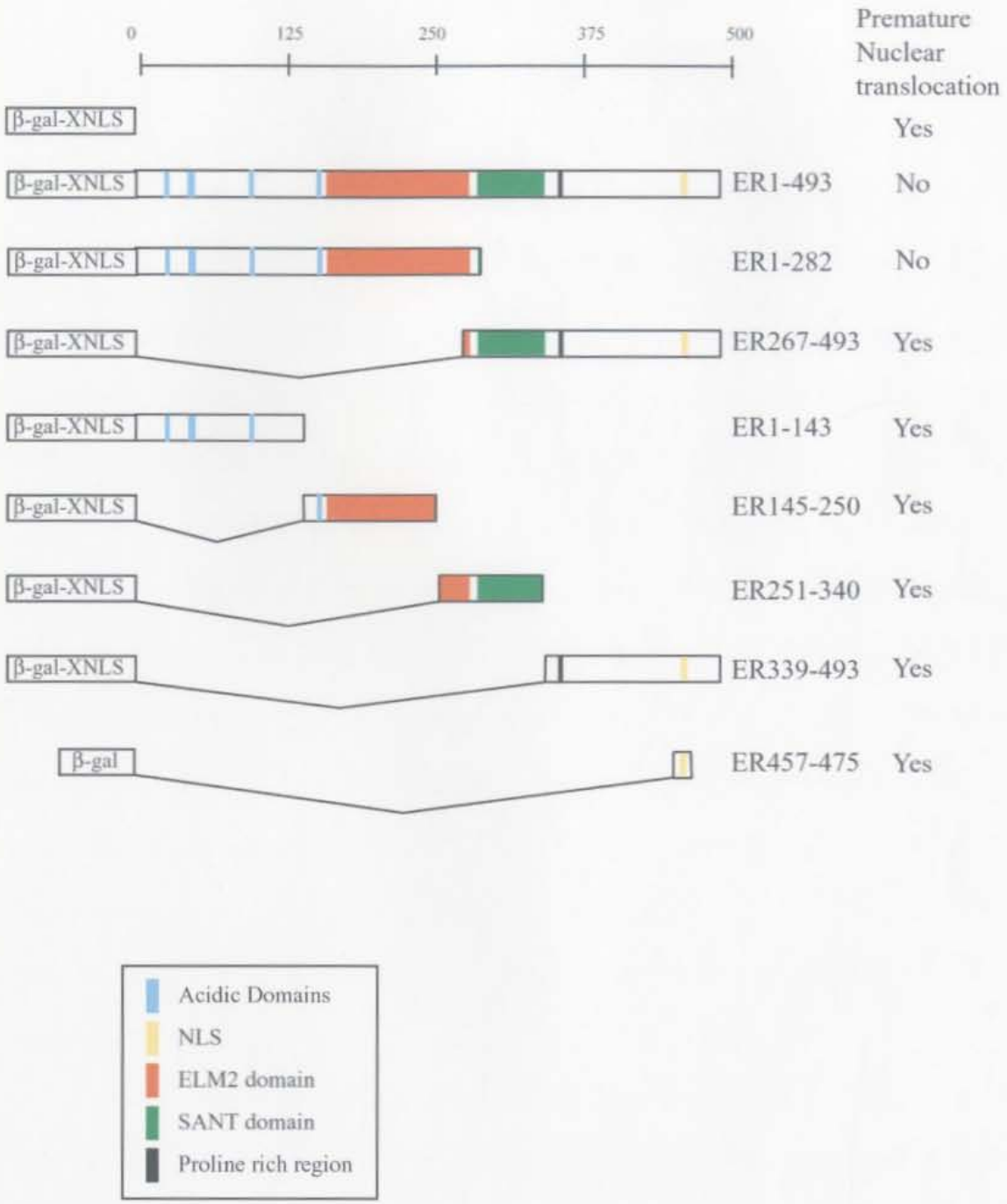

Subcellular localization
(percentage)
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
XNF7NLS-ER1
XNF7NLS
XNF7NLS-ER251-340
XNF7NLS-ER339-493
N
N+C
C
7
7.5
8
8.5
9
Stages

Figure 4.18 Schematic illustrating different β-GAL fusions used to identify the mechanism of regulation of nuclear transport of MI-ER1 during *Xenopus* embryonic development.

Colored boxes represent some of the domains present in MI-ER1, see boxed legend. Amino acid numbers of different regions of XMI-ER1 are indicated. β-GAL-XNLS= β-GAL-XNF7NLS. Whether proteins were translocated prematurely is indicated on the right.


0
125
250
375
500
Premature Nuclear translocation
β-gal-XNLS
Yes
β-gal-XNLS
ER1-493
No
β-gal-XNLS
ER1-282
No
β-gal-XNLS
ER267-493
Yes
β-gal-XNLS
ER1-143
Yes
β-gal-XNLS
ER145-250
Yes
β-gal-XNLS
ER251-340
Yes
β-gal-XNLS
ER339-493
Yes
β-gal
ER457-475
Yes
Acidic Domains
NLS
ELM2 domain
SANT domain
Proline rich region

4.8 Conclusion

The results shown in this chapter are summarized in figure 4.18. The regulation of cytoplasmic retention of XMI-ER1 in *Xenopus laevis* embryos was investigated. It was shown that the NLS (ER457-475) of XMI-ER1 was able to translocate β-GAL to the nucleus prematurely, which suggested that cytoplasmic retention of XMI-ER1 is not dependent on NLS recognition and binding by the import machinery. Investigation of a phosphorylation site in the vicinity of this NLS showed, that the subcellular localization of XMI-ER1 is not regulated by phosphorylation of this putative cdc2/ PKA site. Furthermore, I have shown that retention of XMI-ER1 before stage 8 in embryos appears to be regulated by binding of XMI-ER1 to an "anchor" molecule in the cytoplasm. This was shown by addition of the XNF7NLS to XMI-ER1. XNF7NLS-ER1 translocated to the nuclei beginning at stage 8.5, similar to what is observed for full length XMI-ER1, while the XNF7NLS by itself translocated pre-maturely. This indicates that the cytoplasmic anchoring of XMI-ER1 is dominant over the addition of a functional NLS. Although the smallest region responsible for cytoplasmic retention of XMI-ER1 has not been identified, it is localized in a region between amino acid residues 1 and 282.

In an attempt to identify the smallest region of XMI-ER1 responsible for cytoplasmic localization during early development, the N-terminal half of XMI-ER1 was divided into the following regions: ER1-143, ER145-250, and ER251-340. ER339-493, which was predicted to translocate prematurely since ER267-

493 translocated prematurely, was used as a control. These regions were fused to β-GAL-XNF7NLS and expressed in *Xenopus* embryos. Premature translocation was observed for each of these regions.

Possible motifs responsible for the cytoplasmic retention of XMI-ER1 during development include CRDs. However, no known CRD could be identified in XMI-ER1. Since all regions smaller than ER1-282 (ER1-143, ER145-250 and ER251-340) translocated to the nucleus prematurely, the smallest region necessary for cytoplasmic retention of XMI-ER1 identified was ER1-282. These results indicated that binding of a CRD in XMI-ER1 to a cytoplasmic anchoring molecule is responsible for the cytoplasmic retention of XMI-ER1 in developing *Xenopus laevis* embryos.

Chapter 5 General Discussion

XMI-ER1 is located in the cytoplasm of developing embryos until stage 8, around the start of the mid-blastula transition, when it localizes to the nucleus. MBT is characterized by a series of significant cellular and molecular changes including changes in cell cycle rate and synchrony, onset of cell movement and activation of zygotic gene transcription. The correlation between entry of XMI-ER1 into nuclei and MBT suggests a potential role for XMI-ER1 in zygotic transcription. XMI-ER1 possesses several characteristics of a regulator of gene expression. It is an immediate early gene, which localizes to the nucleus and contains acidic domains capable of trans-activating transcription (Paterno *et al.*, 1997). It also contains an ELM2 domain, which in its human homologue is shown to bind HDAC1/2 (Ding *et al.,* unpublished data), and a SANT domain, which is found to be a DNA binding motif in other proteins (Aasland *et al.*, 1996), and which possibly can bind HAT (Sterner *et al.*, 2002). Taken together, this suggests that XMI-ER1 may function in regulating the expression of at least a subset of zygotic genes at MBT.

5.1 Analysis of NLS in XMI-ER1

Nuclear translocation of proteins is dependent on the presence of a nuclear localization signal, NLS. Based on computer analysis and analysis of the basic amino acids in XMI-ER1, four putative NLS have been identified, of which one was shown to be functional. NLS4, located near the C-terminus of XMI-ER1, is able to directly localize a carrier protein to the nucleus. Moreover, deletion of the core NLS4 from the full-length XMI-ER1 sequence results in cytoplasmic localization of XMI-ER1, indicating that the nuclear translocation of XMI-ER1 is dependent on NLS4.

The sequence of the NLS4 (^{463}RPIKRQRMD^{471}) is similar to one of the NLSs found in human c-MYC (PAAKRVKLD) (Makkerh *et al.*, 1996; Post *et al.*, 2001). Both NLS have a proline in the N-terminal flanking region, followed by one (XMI-ER1) or two (c-MYC) amino acids with a non-polar side chain. Addition of PAA to a region of basic amino acids is necessary for the import of a synthetic peptide KKKK (Makkerh *et al.*, 1996). The core NLS is made up of three basic amino acids, KR.R (XMI-ER1) and KR.K (c-MYC). The C-terminal flanking region contains one nonpolar amino acid (M in XMI-ER1; L in c-MYC), followed by an acidic amino acid and an uncharged polar amino acid, DS. An acidic amino acid when preceded by the leucine residue has been shown to greatly enhance import into the nucleus for the c-MYC NLS (Makkerh *et al.*, 1996). The main difference between the two NLS is the amino acid in the core region, which contains an uncharged polar residue (Q) in the XMI-ER1 sequence (KRQR), and a nonpolar

residue (V) in the c-MYC (KRVK). Small changes between NLS have been shown to determine specific recognition between importin-α isoforms or influence the binding strength of the NLS to importin-α, thereby creating a hierarchy in transport of the protein (Kohler *et al.*, 1999; Moroianu, 1999).

Analysis of the efficiency of NLS4 to carry a cytoplasmic protein to the nucleus showed that in addition to the residues described above, flanking regions were necessary. Inclusion of these flanking regions increased the percentage of cells with nuclear staining from 38% to 74%. Interestingly, when the NLS4 including the flanking regions is fused to a larger protein, β-galactosidase, the nuclear localization increases from 74% to 93%. This is could be due to differences in the conformation of EGFP-ER457-475 vs. β-GAL-ER457-475, and/or the ability of the ER457-475 region to be recognized by the import receptors (Makkerh *et al.*, 1996). It could also be due to diffusion of EGFP-ER457-475 from the nucleus to the cytoplasm.

Analysis of the secondary structure of the EGFP-ER457-475 and β-GAL-ER457-475 by nnpredict (www.cmpharm.ucsf.edu/nnpredict), showed no difference in the presentation of the NLS. However, this is a prediction based on the conformational specification of short peptides, instead of full-length proteins and therefore may not correspond to the true conformation of the fusion protein.

Another explanation could be that EGFP-ER457-475 is able to diffuse out of the nucleus, because of its small size, whereas β-GAL-ER457-475 cannot. Diffusion is based on movement of molecules in a solution from a high to a low

concentration until equilibrium is reached, whereby the movement of molecules is random and the molecules are spread out so that the concentration is equal in the solution. In cells, molecules smaller than 40-45 kDa will diffuse freely through the NPC. The proteins will diffuse either way unless they possess nuclear export signals (NES) or nuclear retention domains (NRD) (Radtke *et al.*, 2001; Ribbeck and Gorlich, 2001). No known NRD or NES could be identified in XMI-ER1. While the protein is translated, there is a steady import of protein into the nucleus. This results in accumulation of the transiently expressed protein in the nucleus and generation of a concentration gradient, with a low concentration of the protein in the cytoplasm and a high concentration in the nucleus. When the molecular weight of the protein is below the limiting size (of 40-60 kDa), some protein may "leak" out of the nucleus by simple diffusion, resulting in cytoplasmic staining. However, active import of molecules into the nucleus is rapid, while diffusion is slow (Gorlich, 1997; Gorlich and Mattaj, 1996; Ribbeck and Gorlich, 2001), resulting in a high level of protein in the nucleus with very little protein in the cytoplasm.

The EGFP-ER457-475 has a predicted size of 32 kDa making it possible to diffuse through the NPC. Therefore, the lower percentage of cells showing nuclear staining in the EGFP-ER457-475 construct, as compared to the β-gal-ER457-475 construct, could be explained by diffusion. The β-gal-ER457-475, however, cannot diffuse out of the nucleus because its molecular weight is 120

kDa and will thus remain in the nucleus after import, while more protein is still being imported, resulting in a higher percentage of exclusive nuclear staining.

Fusion of ER461-471, containing the basic cluster of the NLS4 including 2 flanking amino acid residues, to a cytoplasmic protein resulted in nuclear staining in about 40% of NIH 3T3 cells. There was no notable difference in nuclear localization between fusion to EGFP or β-gal. This indicates that for constructs containing ER461-471 vs. ER457-475, the recognition of the basic cluster of the NLS (^{463}RPIKRQR^{469}) by import receptors is optimized when 6 flanking amino acid residues are present, compared to 2 amino acid residues in the region within ER461-471. However, the minimum flanking sequences important for the recognition of NLS4 has not been determined. Identification of the exact sequences is difficult given the importance of both primary sequence and the secondary structure in contributing to efficiency.

The sequence used for the analysis is as follows: ^{457}PDDTNE**RPIKRQR**MDSPGK^{475}, in which the basic cluster of the NLS4 is noted in bold. The NLS4 flanking regions both contain a proline (^{457}P and ^{473}P), which act as helix breakers and might be important for the presentation of the NLS to import receptors (Boulikas, 1994; Boulikas, 1997; Makkerh *et al.*, 1996). Acidic residues are present at the N-terminal flanking region, a common feature of NLS found in many nuclear proteins, indicating that these hydrophilic residues facilitate binding to the importins (Boulikas, 1994; Boulikas, 1997; Makkerh *et al.*, 1996). Introducing point mutations within the flanking regions of the ER457-475

peptide and analyzing the efficiency to translocate either β-gal or EGFP to the nucleus could help identify the importance of the flanking amino acid residues.

In addition to the NLS in the C-terminus of XMI-ER1, a weak NLS was identified in the N-terminus of XMI-ER1, NLS1 (^{138}RPRRCK^{143}). Although NLS1 was not efficient in directing a protein exclusively to the nucleus, 80% of cells transfected with EGFP-NLS1 displayed definite nuclear greater than cytoplasmic (N>C) staining, showing that it functioned as a weak NLS. Moreover, when the ER1-143, including the NLS1 region was fused to EGFP, nuclear greater than cytoplasmic staining was observed in about 90% of the transfected cells. When ER1-143 is fused to β-GAL, 30% of the cells show exclusive nuclear staining (N), indicating that nuclear accumulation is not a result of diffusion followed by binding of a nuclear retention domain to nuclear structures. However, this difference in N>C versus N staining might be the result of diffusion of EGFP-ER1-143 out of the nucleus or may be a result of export of ER1-143 from the nucleus by use of an NES.

Nuclear localization of ER1-143 is not a result of translocation by simple diffusion, because β-GAL-ER1-143 translocates to the nucleus. Furthermore, since EGFP-ER1-143 is not exclusively nuclear, it is unlikely that there is an NRD present in this region, as this would lead to accumulation of the fusion protein into the nucleus.

Experiments to determine the strength of NLS1 could include transfection of a construct containing a nuclear retention signal in addition to the EGFP-ER1-

143 fusion. If the nuclear greater than cytoplasmic staining in the cells is a result of diffusion out of the nucleus, than addition of this nuclear retention signal will result in a slow accumulation of the transient protein in the nuclei of transfected cells, similar to the results with the β-gal fusion. Also, co-transfection of a (fusion) protein containing a stronger NLS might result in a reduction in nuclear staining because the importins will bind to an NLS with a stronger affinity for the receptor and will translocate this protein before the protein containing a weak NLS (Ribbeck and Gorlich, 2001).

The observation that there is N>C staining in cells transfected with ER1-143 cannot be explained by existence of a bi-directional transport signal or a nuclear export signal (NES) (Fischer *et al.*, 1995; Schmidt-Zachmann *et al.*, 1993). I was not able to identify a consensus sequence for either a bi-directional transport signal or a NES in XMI-ER1. Moreover, the low percentage nuclear staining could only be explained if the NLS1 is the bi-directional transport signal, because EGFP-NLS1 and EGFP-ER1-143 show the same percentage of nuclear > cytoplasmic (80% N>C) staining. Moreover, this signal would have to be inhibited in the full length XMI-ER1, because transfection of EGFP-ER1-493 into NIH3T3 cells results in exclusive nuclear staining in 100% of transfected cells, which would be unlikely if an NES or bi-directional NLS were present in the sequence.

When ER1-137, which does not include the NLS1, is fused to either EGFP or β-gal, 50% of the cells show nuclear greater than cytoplasmic (N>C) staining,

and 6% of the β-gal-ER1-137 transfected cells showed exclusive nuclear (N) staining. While this percentage nuclear staining is lower than nuclear staining observed for cells transfected with EGFP-ER1-143 (no exclusive nuclear staining, 90% N>C) or β-GAL-ER1-143 (30% N), the observed level of nuclear (N) staining suggests active nuclear targeting of ER1-137 independent of the NLS1. The N-terminus of XMI-ER1 does not contain any additional regions rich in basic amino acids; actually, the N-terminus is very rich in acidic residues. An alternative NLS, consisting mainly of acidic amino acids has been shown to direct the nuclear translocation of Human Topoisomerase I (Mo *et al.*, 2000). Although XMI-ER1 does not contain a similar motif within the acidic regions in the N-terminus, the existence of an acidic NLS in this region cannot be excluded. Mutation analysis of the acidic regions might elucidate the potential of the acidic regions to direct XMI-ER1 to the nucleus. However, since transfection of ER1-137 into NIH 3T3 cells shows only a low percentage of protein to be translocated to the nucleus (6%), it seems unlikely that the nuclear localization of XMI-ER1 is the result of an alternate NLS in the N-terminus.

Active import by piggy backing with another nuclear protein might be responsible for the partial nuclear localization of ER1-137. We were not able to recognize any known protein binding sites in the N-terminus, but whether proteins bind to this region could be investigated by yeast two-hybrid analysis or analysis of labeled proteins which might co-immunoprecipitate with ER1-137.

Two additional putative NLS in XMI-ER1 were unable to translocate a fusion protein to the nucleus in NIH 3T3 cells as well as *Xenopus* embryos; NLS2, ^{279}AL**RRLRFNVKAAR**EE^{293}, contains a phenylalanine (F), which is a very bulky residue known to interfere with the binding to importin-α (Boulikas, 1994; Boulikas, 1997). Also the basic amino acids are interspersed over 11 residues resulting in a different pattern from what is normally seen in the NLS of nuclear proteins, where four basic residues are found in a hexapeptide (Boulikas, 1994; Boulikas, 1997). The sequence of NLS3 is similar to a bipartite NLS: ^{320}**KK**SERYDFFAQQTRFG**KKK**^{338}. Analysis of the functional residues of the bipartite NLS found in nucleoplasmin, **KR**PAATKKAGQA**KKKK**, has shown that the arginines and lysines shown in bold are important for the nuclear translocation of the protein (Dingwall *et al.*, 1988; Gorlich *et al.*, 1995a). Analysis of the amino acids in the spacer between these basic residues showed that the spacer can be quite variable in length and in type of residues, but in most functional NLS the spacer is 10-12 amino acids and contains no bulky amino acid residues (Jans and Hubner, 1996; Robbins *et al.*, 1991). However, Robbins *et al.* showed that the spacer length could be doubled and the protein still localized to the nucleus, although a peptide with a spacer length of 14 amino acids was not efficient for nuclear localization, probably because of the reduced ability to bind to importin-α (Robbins *et al.*, 1991). The NLS3 region in XMI-ER1 contains a spacer of 14 amino acids, indicating that the spacer length might be the reason

that it is not a functional NLS. Also, the presence of three phenylalanines, which are very bulky, might change the affinity of the NLS for the import receptors.

There are numerous proteins whose nuclear localization is dependent on the presence of multiple NLSs, which act together to direct the protein to the nucleus. One example is b-MYB (Takemoto *et al.*, 1994), whose nuclear localization is dependent on co-operation of a monopartite and a bipartite NLS in the protein. In proteins that possess multiple NLSs, the individual NLS are often weak NLS, but when used together are efficient for exclusive nuclear staining of transfected cells. The existence of a weak NLS in the N-terminus of XMI-ER1 might suggest a similar mechanism of action between NLS1 and NLS3 or NLS1 and NLS4. The possibility of NLS4 co-operating with another basic region in the XMI-ER1 sequence was investigated by deletion and/or mutation of NLS1 and NLS3. However, deletion of NLS1, NLS3 and NLS1 and 3 resulted in exclusive nuclear staining of the transfected cells. Transfection of EGFP-ER339-493, containing only the NLS4, resulted in exclusive nuclear staining, while deletion of NLS4 from full-length XMI-ER1 abolished nuclear transport, indicating that nuclear translocation of XMI-ER1 is not dependent on co-operation of NLS4 with the weak NLS1, nor on co-operation of NLS4 with the putative bipartite NLS3. Moreover, when EGFP-ΔNLS1, EGFP-ΔNLS3 and EGFP-ΔNLS1+3 were injected into *Xenopus* embryos, followed by whole mount staining using our anti–XMI-ER1 antibody, exclusive nuclear staining was observed in all embryos at stage 10 (data not shown). The putative NLS2 was excluded from these

experiments because it did not show any ability to target a cytoplasmic protein to the nucleus as well as because analysis of its sequence indicated that although there are basic amino acid residues present, it is not likely function as an NLS in NIH 3T3 cells as well as *Xenopus* embryos. These results show that the ability of NLS4 to translocate XMI-ER1 to the nucleus is not dependent on co-operation of NLS4 with any other sequences in XMI-ER1.

Investigation into the ability of four putative NLS present in XMI-ER1 to direct the protein to the nucleus showed that only NLS4 is functional. NLS4 was shown to be necessary and sufficient for the nuclear localization of XMI-ER1 in NIH 3T3 cells and in *Xenopus* embryos. An additional, weak NLS, NLS1 is located in the N-terminus of XMI-ER1. Interestingly, human MI-ER1-β isoforms (which are orthologues of XMI-ER1) contain both NLS, whereas the human MI-ER1-α isoforms only contain the NLS1 region ^{138}RPRRCK^{143}. Transfection into NIH 3T3 cells has shown that the human MI-ER1-β isoforms are constitutively localized to the nucleus, whereas the human MI-ER1-α isoforms are mainly localized in the cytoplasm (Paterno *et al.* manuscript submitted). This is consistent with what was found for XMI-ER1, where deletion of the NLS4 results in a pattern similar to that found after transfection of human MI-ER1α. This shows that the nuclear localization of XMI-ER1 is comparable to hMI-ER1-β, whereas the subcellular localization of XMI-ER1-ΔNLS4 is comparable to the subcellular localization of hMI-ER1-α.

However, the NLS found in the C-terminus of HuMI-ER1β is different from the XMI-ER1 NLS in that it is a classic monopartite NLS, ^{468}PAKRRRV^{474}. The basic amino acid residues are flanked by a proline, which acts as a helix breaker and thus the conformation results in better presentation of the NLS to the import receptors (Boulikas, 1994; Boulikas, 1997). The observation that the c-MYC-like NLS found in XMI-ER1 is not conserved between *Xenopus* and human suggests that the recognition and/or the regulation of this NLS may be different between the two proteins. There are no examples in the literature showing different NLS between orthologues of different species. Since our interest is the identification of regulation of nuclear import of XMI-ER1 during *Xenopus* development, further investigation into the differences in function of the NLS found in hMI-ER1β vs. the XMI-ER1 NLS4 has not been performed.

To investigate whether the NLS4 is necessary and sufficient for the nuclear translocation of XMI-ER1 in *Xenopus* embryos, different regions of XMI-ER1 were injected into *Xenopus* embryos. Injection of different EGFP-ER1 cRNAs into *Xenopus* embryos showed that NLS4 is necessary and sufficient for nuclear translocation of the fusion protein during stage 13 in *Xenopus* embryos. The XMI-ER1 NLS1 seems to be recognized as a weak NLS, since injection of EGFP-ER1-143 cRNA results in nuclear greater than cytoplasmic staining in stage 13 embryos, similar to what was observed in NIH 3T3 cells. This suggests that the import machinery of both mouse and *Xenopus* cells recognize the XMI-

ER1 NLS4 and are able to translocate the EGFP-XMI-ER1 fusions containing this NLS.

I showed that in NIH 3T3 cells as well as in *Xenopus* embryos NLS4 is necessary and sufficient for the nuclear translocation of XMI-ER1. Since nuclear translocation of XMI-ER1 in *Xenopus* embryos is stage dependent, additional regulatory sequences important for the regulation of nuclear transport of XMI-ER1 during early development may be present. The investigation of these possible regulatory sequences is discussed below.

5.2 Temporal regulation of nuclear import of XMI-ER1 in *Xenopus* embryos

This section will consider possible mechanisms used for the regulation of nuclear import of XMI-ER1, and using the results of experiments in this thesis, attempt to develop a possible model for the regulation of nuclear transport of XMI-ER1 during *Xenopus* embryonic development.

I showed that the NLS4, located near the C-terminus in XMI-ER1, is necessary and sufficient for the nuclear translocation of XMI-ER1 in NIH 3T3 cells as well as *Xenopus* embryos. However, since XMI-ER1 is localized in the cytoplasm of developing embryos until the mid-blastula stages, the function of NLS4 must be regulated. Therefore, I investigated the possible mechanisms by which XMI-ER1 nuclear import is regulated.

I have shown, that addition of β-GAL or EGFP to XMI-ER1 does not interfere with the temporal pattern of nuclear import in *Xenopus* embryos. Injection of β-GAL-ER1 or EGFP-ER1 into embryos showed that the fusion proteins start showing nuclear staining in the injection patch, beginning at stage 8 and that at stage 8.5 67% of the embryos show nuclear staining. During stage 9 97% of embryos show nuclear staining and at stage 10 100% of nuclei in the injection patch display nuclear staining (figure 4.1).

There are different mechanisms by which the cytoplasmic localization of XMI-ER1 could be regulated during early *Xenopus* development (also see figure 1.3.3.):

1) It was possible that NLS4, which is shown to be necessary and sufficient for the nuclear import of XMI-ER1 in NIH 3T3 cells as well as in *Xenopus* embryos, was not recognized and transported until the beginning of stage 8. This could be explained if factors necessary for import of XMI-ER1 are not present or active until the beginning of stage 8.

2) Cytoplasmic localization of XMI-ER1 could be due to binding of the protein to a cytoplasmic anchor, like RNA or a protein.

3) Cytoplasmic localization of XMI-ER1 could also be due to masking of the NLS, preventing NLS binding to the import receptors. This may be a result of inter-molecular masking by binding to another molecule or by dimer formation, or intra-molecular masking of the NLS by an XMI-ER1 precursor.

4) Regulation of nuclear import could be influenced by phosphorylation of XMI-ER1 or phosphorylation of a XMI-ER1 interacting protein.

5) Cytoplasmic localization of XMI-ER1 during *Xenopus* embryonic development could be a result of one or a combination of the mechanisms described above.

1) Cytoplasmic localization of XMI-ER1 during *Xenopus* embryonic development is not caused by lack of recognition of the NLS by import receptors.

Investigation of the literature has shown that many of the proteins translocated early during cleavage stages in *Xenopus laevis* contain a bipartite NLS (Dingwall *et al.*, 1982; Laskey *et al.*, 1993; Robbins *et al.*, 1991). Also, proteins translocating during a specific stage in development that contain a functional bipartite NLS all seem to be actively retained in the cytoplasm. Examples of this are Xnf7and Pwa33, which contain cytoplasmic retention domains (CRD), and do not translocate until MBT (El Hodiri *et al.*, 1997)) (Bellini *et al.*, 1993). CBTF, which also contains a bipartite NLS is retained in the cytoplasm until MBT by binding to maternal RNA, which is degraded at the start of zygotic transcription (Brzostowski *et al.*, 2000). XMI-ER1 does not contain a functional bipartite NLS, but the functional NLS is a rare type similar to the NLS found in human c-MYC. There are no reports describing the function of an NLS similar to the XMI-ER1 NLS in *Xenopus*. Therefore, it was difficult to predict whether this type of NLS will be recognized by the import receptors and

transported into the nucleus directly after fertilization or whether the import receptors for this type of NLS were not active/ present until stage 8.

If the NLS would not be recognized and bound by the import receptors until stage 8, then the cytoplasmic localization of XMI-ER1 until this stage is a result of lack of recognition of the NLS. However, if the NLS is bound and transported prematurely, there must be alternate mechanisms by which XMI-ER1 was retained in the cytoplasm.

I showed that β-GAL-ER457-475 was localized in the nuclei of stage 7-7.5 embryos, indicating that the NLS was functional during these stages and that cytoplasmic localization of XMI-ER1 until stage 8 is not dependent on the inability of NLS4 to be recognized and transported. Clearly, cytoplasmic retention of XMI-ER1 before stage 8 is dependent on other mechanisms of regulation of nuclear transport. Moreover, this experiment shows that nuclear import of XMI-ER1 in *Xenopus* embryos is independent of the transcription, translation or activation of new molecules. After all, if nuclear transport of XMI-ER1 were dependent on transcription or translation or activation of a molecule, premature nuclear import of ER457-475 would be unlikely. This is supported by the observation that when RNA transcription by RNA polymerases I and II is inhibited by α-amanitin, no difference in the temporal pattern of nuclear translocation was observed, indicating that nuclear import of XMI-ER1 is not dependent on protein(s) translated from zygotic messages that are newly transcribed at MBT (Luchman *et al.*, 2002).

2) Cytoplasmic localization of XMI-ER1 could be due to binding of the protein to a cytoplasmic anchor, like RNA or a protein.

And/or

3) Cytoplasmic localization could also be due to masking of the NLS, which thereby inhibits the binding of the NLS by import receptors. This may be a result of inter-molecular masking by binding to another molecule or by dimer formation, or intra-molecular masking of the NLS by a XMI-ER1 precursor.

Although cytoplasmic retention by binding to an anchor and NLS-masking are two separate mechanisms, use of these mechanisms can be differentiated by performing one experiment. The addition of another functional NLS to a protein that is localized in the cytoplasm due to NLS masking has been shown to promote its nuclear import, as shown for NF-κB/I-κB (Li *et al.*, 1994a). However, cytoplasmic sequestration by binding of the protein to a cytoplasmic anchor is dominant over the function of the additional NLS, as shown for XNF7 (Li *et al.*, 1994a). Therefore, to investigate whether XMI-ER1 is localized in the cytoplasm as a result of binding to a cytoplasmic anchoring protein or as a result of NLS-masking, the XNF7 NLS was fused to the N-terminus of XMI-ER1.

Addition of the bipartite NLS of XNF7 to full length XMI-ER1 had no effect on the timing of nuclear translocation. This indicates that XMI-ER1 is retained in the cytoplasm as a result of binding to a cytoplasmic "anchor". Interaction of proteins with cytoplasmic anchoring molecules occurs *via* a cytoplasmic retention domain, CRD. However, although this experiment indicated that XMI-ER1 is retained in the cytoplasm by binding to a cytoplasmic anchoring molecule, it does not indicate whether binding to this anchoring molecule is dependent on the NLS region or another region within XMI-ER1. Therefore, we cannot exclude that masking of the XMI-ER1 NLS occurs as an additional mechanism.

To identify the location of the CRD in XMI-ER1, the protein was divided into two regions: ER1-282 and ER267-493. Addition of the XNF7NLS to ER1-282 resulted in cytoplasmic retention, while addition of the XNF7NLS to ER267-493 did not. This indicates that the CRD is located in the N-terminal half of XMI-ER1. In an attempt to narrow down the region responsible for the cytoplasmic retention of XMI-ER1 in *Xenopus* embryos, the N-terminal half of XMI-ER1 was further subdivided: ER1-143, ER145-250 and ER251-340. Addition of the XNF7NLS to these regions resulted in premature nuclear localization of all the β-GAL-XNF7NLS-ER1 fusion proteins. This indicates that the binding of the CRD in XMI-ER1 to the cytoplasmic anchor is inhibited. This could be a result of cleavage of the CRD or change in conformation of the protein so that the CRD is no longer able to bind to the cytoplasmic anchor. It is also possible that binding of XMI-ER1 to the cytoplasmic anchor requires two regions within XMI-ER1 that

work in concert to form a binding site. Division of the XMI-ER1 into smaller regions might separate these two regions.

Investigation of the CRD could be performed using three different strategies. One strategy is based on identification of conserved regions; a second strategy is based on identification of motifs in the XMI-ER1 N-terminus. The third strategy for identifying the smallest region responsible for cytoplasmic retention on XMI-ER1 is based on progressive deletion of residues from the N-terminus and/ or the C-terminus.

To identify the function of conserved regions, hMI-ER1β may be injected into embryos to give an indication whether the CRD is conserved between human and *Xenopus*. If the temporal pattern of nuclear translocation of hMI-ER1β is similar to the temporal pattern of XMI-ER1, the CRD is present in both proteins. Comparison of conserved amino acid sequences might give an indication of the localization of the CRD. On the other hand, if the temporal regulation of HuMI-ER1β is different from XMI-ER1, non-conserved amino acid residues may be responsible for cytoplasmic retention of XMI-ER1 during embryonic development.

The second strategy is based on identification of motifs in the XMI-ER1 N-terminus, which might be responsible for cytoplasmic retention. When the XNF7NLS is added to ER1-282, the fusion protein localized to the nucleus beginning at stage 8, indicating that the CRD is located between residues 1 and 282. The motifs present in ER1-282 include the acidic regions, the NLS1 and the

ELM2 domain. The acidic domains are intact in ER 1-143 and ER145-250, although the first, second and third acidic domains are located within ER1-143 and the fourth is located in ER145-250. If the acidic domains are working together in forming a binding site for a cytoplasmic anchoring protein, dividing ER1-282 into smaller regions may disrupt this function. The NLS1 is intact in ER1-143, suggesting that the CRD does not span the NLS1. However, in dividing XMI-ER1 region into regions containing residues 145-250 and 251-340, the ELM2 domain is split. The fact that these smaller regions are not retained in the cytoplasm suggests that amino acid residues within the ELM2 domain may be responsible for binding to a cytoplasmic anchor, with release and subsequent nuclear translocation at MBT. Interestingly, the ELM2 domain is responsible for binding HDAC1 (Ding *et al,* unpublished data). Whether the ELM2 domain could be responsible for the cytoplasmic localization of XMI-ER1 before MBT requires further investigation. Investigation of the role of the ELM2 domain or sequences within this domain could be performed by deletion of the ELM2 region from the full length XMI-ER1. Injection into *Xenopus* embryos and examination of the temporal pattern will show whether this region is responsible for the cytoplasmic retention. Addition of the XNF7NLS to the XMI-ER1ΔELM2 might be necessary, since existence of NLS masking or additional regulation of the NLS of XMI-ER1 has not yet been excluded. If the fusion proteins localize prematurely, the ELM2 domain or sequences within this region are responsible for the retention in the cytoplasm.

A third strategy for identifying the smallest region responsible for cytoplasmic retention on XMI-ER1 is based on progressive deletion of residues from the N-terminus and/ or the C-terminus. Deletion of 5 residues at a time, for example, from the N-terminus and/ or the C-terminus followed by injection into *Xenopus* embryos might give an indication as to which amino acid residues are important for the cytoplasmic retention. This might also show whether the cytoplasmic retention is a result of binding of two or more regions within XMI-ER1 to the cytoplasmic anchor.

I have not yet identified the smallest region responsible for cytoplasmic retention of XMI-ER1 and can therefore not discuss how different the CRD in XMI-ER1 is from CRDs found in other proteins. However, the cytoplasmic retention domain found in XNF7 is a motif of 22 amino acid residues, conserved between XNF7 and the newt orthologue, Pwa33 (Li *et al.*, 1994a). A cytoplasmic retention domain is also described for XMyoD, and consists of two regions, one in the N-terminus which co-operates with a region in the C-terminus of the protein (Rupp *et al.*, 1994). Closer analysis of the CRD of XMI-ER1 will be necessary to be able to compare the XMI-ER1 CRD to CRDs found in other proteins. However, there are no regions of similarity between XMI-ER1 and XNF7 or XMyoD that would suggest the presence of a conserved CRD sequence.

I did not investigate the nature of the cytoplasmic anchor of XMI-ER1. To date not many cytoplasmic anchoring molecules have been identified. It is as yet unknown which molecule(s) XNF7 binds in the cytoplasm, but it is described that

the cytoplasmic anchor of XNF7 most likely is a protein (Li *et al.*, 1994a). Release of XNF7 form its cytoplasmic anchor is shown to be a result of dephosphorylation of XNF7 at MBT (Li *et al.*, 1994b; Li *et al.*, 1994a). However, the cytoplasmic anchoring molecules for the transcription factor CBTF are maternal RNA transcripts as injection of RNase into developing embryos results in premature localization of CBTF. Therefore, nuclear translocation of CBTF during MBT is believed to be a result of degradation of maternal RNA transcripts (Brzostowski *et al.*, 2000). The cytoplasmic anchor for the *Drosophila* transcription factor armadillo is cytoplasmic axin (Tolwinski and Wieschaus, 2001). Identification of the nature of the cytoplasmic anchoring molecule(s) for XMI-ER1 may be combined with investigation of the mode of release of XMI-ER1 from this molecule. For example, release of XMI-ER1 from the cytoplasmic anchoring protein could be a result of direct phosphorylation of XMI-ER1, causing a change in conformation and subsequent release from the anchor, as seen for XNF7 (Li *et al.*, 1994b; Li *et al.*, 1994a). Release from the cytoplasmic anchor could also be a result of phosphorylation of the cytoplasmic anchoring protein, as seen for NF-κB, which is released after phosphorylation of IκB (Kidd, 1992; Zabel *et al.*, 1993).

Investigation of the release of XMI-ER1 from its cytoplasmic anchor could include injection of RNAse, to test whether cytoplasmic retention is a result of binding of XMI-ER1 to maternal RNA transcripts, as in the case of CBTF. Another approach could involve *in vivo* labeling of proteins with radioactive (^{32}P)

inorganic phosphates. This would show whether there is a difference in the phosphorylation state of XMI-ER1 before and after nuclear import. Moreover, *in vitro* labeling of proteins may show whether XMI-ER1 is associated with phosphoproteins before and/ or after nuclear import (*vide infra*).

Cytoplasmic localization of XMI-ER1 during embryonic development is not likely due to NLS masking by an XMI-ER1 precursor followed by cleavage and subsequent nuclear translocation. This was shown by Luchman *et al.* who did not find an additional XMI-ER1 protein and showed that there is no difference in size of XMI-ER1 protein isolated from *Xenopus* embryos before and after nuclear import, as shown by separation of proteins by denaturing gel-electrophoresis (Luchman *et al.*, 2002).

There is evidence that the cytoplasmic localization of XMI-ER1 during *Xenopus* development is not likely to be a result of NLS-masking. NLS masking can be identified by immunohistochemistry using an antibody specifically against the NLS region. If the antibody is able to bind to the NLS, thus showing staining, it indicates there is no NLS-masking. If however, no staining is observed, the NLS could be masked. This was shown to be the case for the NF-κB p50 precursor p105: the C-terminal of p105 appears to retain the p50 subunit in the cytoplasm by intramolecular masking of its NLS. Antibodies specific to the NLS recognize p50, but not p105, implying that the NLS is inaccessible in the larger precursor (Henkel *et al.*, 1992). The mechanism of unmasking appears to be through proteolysis of the p105 C-terminal (Blank *et al.*, 1991; Jans and Hubner,

1996). Luchman *et al.* used an anti-XMI-ER1 antibody to show that endogenous XMI-ER1 was localized in the cytoplasm of *Xenopus* embryos until stage 8, and that at the beginning of stage 8 XMI-ER1 translocated to the nucleus (Luchman *et al.*, 1999). The anti-XMI-ER1 antibody was raised against amino acid residues 464-478, which spans the NLS sequence. Considering the size of the antibody, it would be unlikely that an antibody would be able to bind to the NLS4 region if this was masked by intra- or intermolecular binding. Therefore, it seems unlikely that regulation of cytoplasmic retention of XMI-ER1 is due to masking of its NLS.

Moreover, cytoplasmic localization of XMI-ER1 is not likely to be a result of activation of maternal proteins at stage 8.5 or translation of maternal transcripts at stage 8. This is shown by injection of ER457-475, which was capable of premature nuclear entry. This would be unlikely if new factors are necessary for the nuclear import of XMI-ER1.

Another possibility for regulation of cytoplasmic localization of XMI-ER1 is regulation by phosphorylation. I showed that phosphorylation of a putative cdc2/ PKA site in the vicinity of the NLS has no effect on nuclear translocation. However, although XMI-ER1 possesses multiple putative phosphorylation sites that could influence the regulation of nuclear transport, *in vivo* labeling of proteins in NIH 3T3 cells, using ^{32}P, showed that XMI-ER1 is not phosphorylated in these cells (Paterno *et al.*, unpublished results). However, since nuclear import is regulated in *Xenopus* embryos, *in vivo* labeling of XMI-ER1 with inorganic phosphate at different stages during development might prove a useful tool for

investigation of the phosphorylation state of the protein in *Xenopus* embryos, before and after nuclear import of XMI-ER1. If the protein is phosphorylated, mutations of other putative phosphorylation sites might prove useful for identifying the role of these sites in nuclear transport of XMI-ER1 during embryonic development.

4) Regulation of nuclear import could be a result of phosphorylation of XMI-ER1 or phosphorylation of an XMI-ER1 interacting protein.

Two major mechanisms of regulation of nucleocytoplasmic transport have been described: one is based on masking and unmasking the transport signal; the other is based on anchoring and release. Furthermore, both these mechanisms can be regulated by phosphorylation (reviewed in(Jans, 1995; Jans and Hubner, 1996; Vandromme *et al.*, 1996).

XMI-ER1 contains multiple putative phosphorylation sites, including several CKII sites and a cdc2 kinase site in the vicinity of the NLS, which forms a putative CcN motif. The putative cdc2/ PKA phosphorylation site, located C-terminal of the NLS in XMI-ER1, has a high probability of phosphorylation as predicted by Netphos (0.997, appendix A-2). Investigation of the putative cdc2/PKA showed that neither substitution of the serine with alanine to simulate a dephosphorylated site nor substitution of the serine for an aspartic acid to simulate a phosphorylated site had any effect on the nuclear translocation of

ER1-493 and ER339-493 in NIH 3T3 cells or *Xenopus* embryos. This indicates that phosphorylation on this site is not important for regulation of cytoplasmic retention of XMI-ER1 In *Xenopus* embryos.

XMI-ER1 contains many more potential phosphorylation sites. *In vitro* labeling of XMI-ER1 with inorganic phosphate at different stages during development might prove a useful tool for investigation of the phosphorylation state of the protein, before and after nuclear import. If the protein is phosphorylated, mutations of other putative phosphorylation sites might prove useful for identifying the role of these sites in nuclear transport of XMI-ER1 during embryonic development. However, Luchman *et al.* have examined the possibility of phosphorylation of XMI-ER1 by using phosphotyrosine, phosphoserine and phosphothreonine specific antibodies. Embryo extracts were immunoprecipitated with anti-XMI-ER1 at stages when XMI-ER1 is cytoplasmic (stage 7) or predominantly nuclear (stage 9). Western blots were stained with phosphotyrosine, phosphothreonine, phosphoserine or XMI-ER1 specific antibodies (Luchman *et al.*, 2002). It was shown that XMI-ER1 itself is not phosphorylated and that there was no difference in the phosphorylation pattern of co-migrating phosphoproteins at stage 7 vs. stage 9 when anti-phosphotyrosine or anti-phosphothreonine was used (Luchman *et al.*, 2002). However, analysis of phosphorylation using anti-phosphoserine revealed the presence of four major phosphorylated bands that are associated with XMI-ER1, indicating that XMI-ER1 itself is not phosphorylated on serine residues, but is associated with

several phosphoproteins. Phosphorylation of one of these proteins, a protein with a predicted size of 226 kDa, is very prominent at stage 7 but decreases at stage 9 when XMI-ER1 is nuclear (Luchman *et al.*, 2002). The identity of these phosphoserylproteins is unknown, therefore it could not be determined whether the decrease in band intensity was a result of changes in the amount of associated protein or due to the level of phosphorylation (Luchman *et al.*, 2002). Injection of a dominant negative FGF receptor (Amaya *et al.*, 1991) resulted in a similar phosphorylation pattern of the 226 kDa protein at stage 9 vs. stage 7, as well as a delay in nuclear translocation of XMI-ER1 (Luchman *et al.*, 2002). This indicates that activation of the FGF signaling pathway is important for the nuclear translocation of MI-ER during *Xenopus* embryonic development (Luchman *et al.*, 2002).

5) Cytoplasmic localization of XMI-ER1 during *Xenopus* embryonic development could be a result of one or a combination of the mechanisms described above.

The possible regulatory mechanisms by which XMI-ER1 can be localized in the cytoplasm during early *Xenopus* embryonic development and the translocation of the protein into the nucleus beginning at stage 8 have been discussed. Regulation of nuclear import of transcription factors is often a result of a combination of multiple mechanisms.

The results of the present study show, that XMI-ER1 is retained in the cytoplasm by binding to a cytoplasmic anchor. The minimal CRD necessary for binding to this anchor has not yet been determined, but is located between residues 1 and 282. I also found that ER457-475 translocates prematurely, indicating that this cytoplasmic localization is not due to inability of the import machinery to recognize and transport the NLS. I showed that regulation of the NLS is not a result of phosphorylation of a putative cdc2/ PKA site in the vicinity of the NLS.

Conclusion:

The results of the work described in this thesis, indicate that nuclear translocation of XMI-ER1 is directed by NLS4, which is located near the C-terminus of XMI-ER1. Inclusion of flanking regions in addition to the core NLS greatly enhances the efficiency of nuclear transport. Binding of a CRD in XMI-ER1 to a cytoplasmic anchor, is responsible for the cytoplasmic localization of XMI-ER1 in developing *Xenopus laevis* embryos. Phosphorylation of a cdc2/ PKA site in the vicinity of the NLS has no effect on the regulation of nuclear translocation of XMI-ER1 during early development of *Xenopus laevis*.

5.3 Model of regulation of nuclear import of XMI-ER1 during *Xenopus* development.

In order to prevent aberrant expression of genes normally transcribed at or soon after the start of transcription, the activity of maternal transcription factors present in the early embryo must be tightly regulated. Cytoplasmic sequestration of maternally inherited transcription factors may provide a mechanism by which the transcriptional machinery is restrained (Veenstra *et al.*, 1999). Understanding the mechanisms that might be important for the regulation of nuclear transport of transcription factors during development might provide important information about the general regulation of these transcription factors and their function in normal embryonic development.

Binding of a CRD within XMI-ER1 to a cytoplasmic anchoring molecule in combination with phosphorylation of the cytoplasmic anchor, may be responsible for the cytoplasmic retention of XMI-ER1 in developing *Xenopus laevis* embryos.

In addition, Luchman *et al.* have shown that inhibition of the FGF pathway, by injection of a dominant negative FGFR1 receptor (XFD) results in cytoplasmic retention of ER1 after stage 8, indicating that FGFR1 activation is necessary for the transport (Luchman *et al.*, 2002). Moreover, Luchman *et al.* showed that a phosphoserylprotein with a predicted molecular weight of 226 kDa co-immunoprecipates with XMI-ER1 and that phosphorylation of this phosphoserylprotein decreases at stage 9, when XMI-ER1 is nuclear. However,

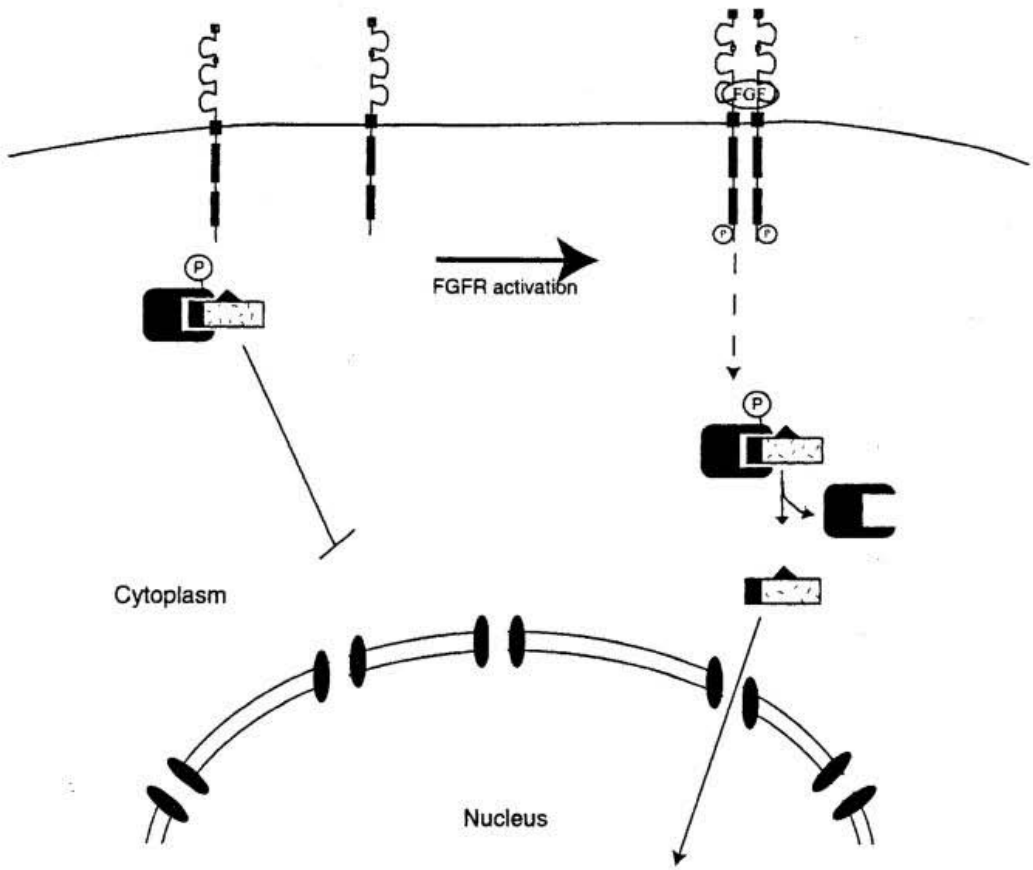
after injection of XFD, the phosphorylation of the phosphoserylprotein does not decrease, but is similar to the phosphorylation at stage 7 (Luchman *et al.*, 2002). This, together with the observation that injection of XFD results in a delay in nuclear import of XMI-ER1, suggests that activation of the FGF signaling pathway is important for the regulation of nuclear import of XMI-ER1 during *Xenopus* embryonic development (Luchman *et al.*, 2002).

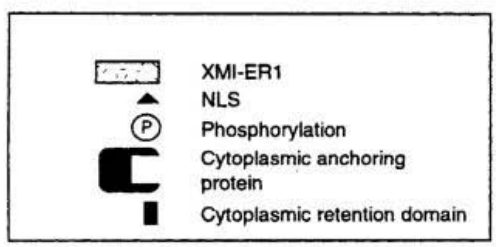
As described before, activation of the FGF signaling pathway results in a cascade of proteins being activated until the signal reaches the cell nucleus, where subsequent activation or inhibition of transcription can occur (Karin, 1994). Activation of these proteins occurs by phosphorylation or dephosphorylation. There are many possible combinations of phosphorylation on either XMI-ER1 or the cytoplasmic anchoring protein or both.

Figure 5.1 shows a model of the possible regulation of nuclear import of XMI-ER1 during *Xenopus* embryonic development: The transcription factor XMI-ER1 is an FGF induced immediate early gene. FGF signaling may play a role in regulation of nuclear transport of transcription factors like XMI-ER1. Therefore, activation of the FGF signaling pathway during mid blastula stages in *Xenopus laevis* could be necessary for the release of XMI-ER1 from its cytoplasmic anchor, followed by translocation to the nucleus.

Figure 5.1 Schematic illustrating a possible mechanism of regulation of cytoplasmic retention of XMI-ER1.

Binding of a CRD in XMI-ER1 to a cytoplasmic anchor, is responsible for the cytoplasmic localization of XMI-ER1 in developing *Xenopus laevis* embryos. Dephosphorylation of this cytoplasmic anchor after activation of the FGF signaling pathway may be responsible for the release of XMI-ER1 from its anchor and subsequent nuclear translocation.


FGF
P
FGFR activation
P
Cytoplasm
Nucleus


XMI-ER1
NLS
P
Phosphorylation
Cytoplasmic anchoring protein
Cytoplasmic retention domain

5.4 Future work

Investigation into the regulation of cytoplasmic retention of XMI-ER1 in *Xenopus* embryos showed that import of XMI-ER1 during Xenopus development is dependent on the NLS located near the C-terminus of ER1. Also, XMI-ER1 is retained in the cytoplasm by binding to a cytoplasmic anchor, with release and subsequent nuclear translocation beginning at stage 8. Dephosphorylation of the cytoplasmic anchoring protein is proposed to be the mechanism of release of XMI-ER1.

Future work would include investigation of the CRD sequence and the regulation of release of XMI-ER1from the cytoplasmic anchor. The ELM2 domain is intact in the region ER1-282, which is retained in the cytoplasm, and split in the smaller regions tested. Therefore, sequences within this region may be important for the cytoplasmic retention of XMI-ER1. Injection of cRNA encoding β-GAL-ΔELM2, followed by assessment of the subcellular localization of this protein at stages 7, 7.5, 8, 8.5 and 9, would identify whether the CRD is located within this region. If the β-GAL-ΔELM2 translocates prematurely, the CRD is within the region of the ELM2 domain. Further deletion of this region would help identify the minimal CRD. Injection of XMI-ER1 constructs, in which the CRD is deleted, into *Xenopus* embryos might give us information about the possible function of XMI-ER1 in embryonic development. Also, injection of ΔCRD and XMI-ER1 full-length into embryos might help us identify possible zygotic genes that are transcribed in response to XMI-ER1 translocation.

If the CRD is not localized within the ELM2 domain, progressive deletions of XMI-ER1, followed by injection into *Xenopus* embryos may give information on the location of the CRD. Constructs containing a functional NLS (like the XMI-ER1 NLS4 or the XNF7NLS) fused to the mutated regions of XMI-ER1 should be used to ensure that a functional NLS is present in all of the proteins tested.

The cytoplasmic anchoring protein of XMI-ER1 should be identified. Isolation of the anchor by co-immunoprecipitation with an anti-XMI-ER1 antibody, followed by proteomic analysis and protein sequencing, may provide information about the identity of the anchoring protein. Knowing the identity of the anchoring protein may provide us with information about the FGF signal transduction pathway important for the regulation of nuclear translocation of XMI-ER1 during embryonic development.

References

Aasland, R., Stewart, A.F., and Gibson, T. (1996) The SANT domain: a putative DNA-binding domain in the SWI-SNF and ADA complexes, the transcriptional co-repressor N-CoR and TFIIIB. *Trends Biochem.Sci.* **21,** 87-88.

Adam, E.J. and Adam, S.A. (1994) Identification of cytosolic factors required for nuclear location sequence-mediated binding to the nuclear envelope. *J Cell Biol* **125,** 547-555.

Alexandropoulos, K., Cheng, G., and Baltimore, D. (1995) Proline-rich sequences that bind to Src homology 3 domains with individual specificities. *Proc.Natl.Acad.Sci.U.S.A* **92,** 3110-3114.

Almouzni, G. and Wolffe, A.P. (1993a) Nuclear assembly, structure, and function: the use of Xenopus in vitro systems. *Exp.Cell Res.* **205,** 1-15.

Almouzni, G. and Wolffe, A.P. (1993b) Replication-coupled chromatin assembly is required for the repression of basal transcription in vivo. *Genes Dev.* **7,** 2033-2047.

Almouzni, G. and Wolffe, A.P. (1995) Constraints on transcriptional activator function contribute to transcriptional quiescence during early Xenopus embryogenesis. *EMBO J.* **14,** 1752-1765.

Amaya, E., Musci, T.J., and Kirschner, M.W. (1991) Expression of a dominant negative mutant of the FGF receptor disrupts mesoderm formation in Xenopus embryos. *Cell* **66,** 257-270.

Amaya, E., Stein, P.A., Musci, T.J., and Kirschner, M.W. (1993) FGF signalling in the early specification of mesoderm in Xenopus. *Development* **118,** 477-487.

Armelin, H.A. (1973) Pituitary extracts and steroid hormones in the control of 3T3 cell growth. *Proc.Natl.Acad.Sci.U.S.A* **70,** 2702-2706.

Bagley, S., Goldberg, M.W., Cronshaw, J.M., Rutherford, S., and Allen, T.D. (2000) The nuclear pore complex. *J.Cell Sci.* **113 (Pt 22),** 3885-3886.

Baird, A. and Klagsbrun, M. (1991) The fibroblast growth factor family. *Cancer Cells* **3,** 239-243.

Bell,J.E. and E.T.Bell. (1988). *Proteins and Enzymes.* Prentice-Hall, Division of Simon & Schuster, New Jersey. Englewood Cliffs.

Bellini, M., Lacroix, J.C., and Gall, J.G. (1993) A putative zinc-binding protein on lampbrush chromosome loops. *EMBO J.* **12,** 107-114.

Bellot, F., Crumley, G., Kaplow, J.M., Schlessinger, J., Jaye, M., and Dionne, C.A. (1991) Ligand-induced transphosphorylation between different FGF receptors. *EMBO J.* **10,** 2849-2854.

Beumer, T.L., Veenstra, G.J., Hage, W.J., and Destree, O.H. (1995) Whole-mount immunohistochemistry on Xenopus embryos using far-red fluorescent dyes. *Trends Genet.* **11,** 9.

Bischoff, F.R. and Gorlich, D. (1997) RanBP1 is crucial for the release of RanGTP from importin beta-related nuclear transport factors. *Febs Lett.* **419,** 249-254.

Bischoff, F.R., Krebber, H., Smirnova, E., Dong, W., and Ponstingl, H. (1995) Co-activation of RanGTPase and inhibition of GTP dissociation by Ran-GTP binding protein RanBP1. *EMBO J.* **14,** 705-715.

Bischoff, J.R., Friedman, P.N., Marshak, D.R., Prives, C., and Beach, D. (1990) Human p53 is phosphorylated by p60-cdc2 and cyclin B-cdc2. *Proc.Natl.Acad.Sci.U.S.A* **87,** 4766-4770.

Blair, W.S., Bogerd, H.P., Madore, S.J., and Cullen, B.R. (1994) Mutational analysis of the transcription activation domain of RelA: identification of a highly synergistic minimal acidic activation module. *Mol.Cell Biol.* **14,** 7226-7234.

Blank, V., Kourilsky, P., and Israel, A. (1991) Cytoplasmic retention, DNA binding and processing of the NF-kappa B p50 precursor are controlled by a small region in its C-terminus. *EMBO J.* **10,** 4159-4167.

Blom, N., Gammeltoft, S., and Brunak, S. (1999) Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *J.Mol.Biol.* **294,** 1351-1362.

Bohlen, P., Esch, F., Baird, A., and Gospodarowicz, D. (1985) Acidic fibroblast growth factor (FGF) from bovine brain: amino-terminal sequence and comparison with basic FGF. *EMBO J.* **4,** 1951-1956.

Boulikas, T. (1994) Putative nuclear localization signals (NLS) in protein transcription factors. *J.Cell Biochem.* **55,** 32-58.

Boulikas, T. (1997) Nuclear import of DNA repair proteins. *Anticancer Res.* **17,** 843-863.

Brzostowski, J., Robinson, C., Orford, R., Elgar, S., Scarlett, G., Peterkin, T., Malartre, M., Kneale, G., Wormington, M., and Guille, M. (2000) RNA-dependent cytoplasmic anchoring of a transcription factor subunit during Xenopus development. *Embo J* **19,** 3683-3693.

Bugler, B., Amalric, F., and Prats, H. (1991) Alternative initiation of translation determines cytoplasmic or nuclear localization of basic fibroblast growth factor. *Mol.Cell Biol.* **11,** 573-577.

Burgess, W.H. and Maciag, T. (1989) The heparin-binding (fibroblast) growth factor family of proteins. *Annu.Rev.Biochem.* **58,** 575-606.

Capdevila, J. and Izpisua Belmonte, J.C. (2000) Perspectives on the evolutionary origin of tetrapod limbs. *J.Exp.Zool.* **288,** 287-303.

Cleaver, O.B., Patterson, K.D., and Krieg, P.A. (1996) Overexpression of the tinman-related genes XNkx-2.5 and XNkx-2.3 in Xenopus embryos results in myocardial hyperplasia. *Development* **122,** 3549-3556.

Cohen, G.B., Ren, R., and Baltimore, D. (1995) Modular binding domains in signal transduction proteins. *Cell* **80,** 237-248.

Cordes, V.C., Reidenbach, S., and Franke, W.W. (1995) High content of a nuclear pore complex protein in cytoplasmic annulate lamellae of Xenopus oocytes. *Eur.J.Cell Biol.* **68,** 240-255.

Coutavas, E., Ren, M., Oppenheim, J.D., D'Eustachio, P., and Rush, M.G. (1993) Characterization of proteins that interact with the cell-cycle regulatory protein Ran/TC4. *Nature* **366,** 585-587.

Cress, W.D. and Triezenberg, S.J. (1991) Critical structural elements of the VP16 transcriptional activation domain. *Science* **251,** 87-90.

Cunliffe, V. and Smith, J.C. (1994) Specification of mesodermal pattern in Xenopus laevis by interactions between Brachyury, noggin and Xwnt-8. *EMBO J.* **13,** 349-359.

Danilchik, M.V. and Gerhart, J.C. (1987) Differentiation of the animal-vegetal axis in Xenopus laevis oocytes. I. Polarized intracellular translocation of platelets establishes the yolk gradient. *Dev.Biol.* **122,** 101-112.

Dent, J.A., Polson, A.G., and Klymkowsky, M.W. (1989) A whole-mount immunocytochemical analysis of the expression of the intermediate filament protein vimentin in Xenopus. *Development* **105,** 61-74.

Dingwall, C. and Laskey, R.A. (1986) Protein import into the cell nucleus. *Annu.Rev.Cell Biol* **2:367-90,** 367-390.

Dingwall, C. and Laskey, R.A. (1991) Nuclear targeting sequences--a consensus? [see comments]. *Trends.Biochem.Sci* **16,** 478-481.

Dingwall, C. and Laskey, R.A. (1998) Nuclear import: a tale of two sites. *Curr Biol* **8,** R922-R924.

Dingwall, C., Robbins, J., Dilworth, S.M., Roberts, B., and Richardson, W.D. (1988) The nucleoplasmin nuclear location sequence is larger and more complex than that of SV-40 large T antigen. *J Cell Biol* **107,** 841-849.

Dingwall, C., Sharnick, S.V., and Laskey, R.A. (1982) A polypeptide domain that specifies migration of nucleoplasmin into the nucleus. *Cell* **30,** 449-458.

Dong, Z., Xu, R.H., Kim, J., Zhan, S.N., Ma, W.Y., Colburn, N.H., and Kung, H. (1996) AP-1/jun is required for early Xenopus development and mediates mesoderm induction by fibroblast growth factor but not by activin. *J.Biol.Chem.* **271,** 9942-9946.

Donohue, P.J., Alberts, G.F., Guo, Y., and Winkles, J.A. (1995) Identification by targeted differential display of an immediate early gene encoding a putative serine/threonine kinase. *J.Biol.Chem.* **270,** 10351-10357.

Dreyer, C. (1987) Differential accumulation of oocyte nuclear proteins by embryonic nuclei of Xenopus. *Development* **101,** 829-846.

El Hodiri, H.M., Che, S., Nelman-Gonzalez, M., Kuang, J., and Etkin, L.D. (1997) Mitogen-activated protein kinase and cyclin B/Cdc2 phosphorylate Xenopus nuclear factor 7 (xnf7) in extracts from mature oocytes. Implications for regulation of xnf7 subcellular localization. *J.Biol.Chem.* **272,** 20463-20470.

Fambrough, D., McClure, K., Kazlauskas, A., and Lander, E.S. (1999) Diverse signaling pathways activated by growth factor receptors induce broadly overlapping, rather than independent, sets of genes. *Cell* **97,** 727-741.

Fanger, G.R. (1999) Regulation of the MAPK family members: role of subcellular localization and architectural organization. *Histol.Histopathol.* **14,** 887-894.

Feldherr, C.M. and Akin, D. (1990) The permeability of the nuclear envelope in dividing and nondividing cell cultures. *J.Cell Biol.* **111,** 1-8.

Feldherr, C.M. and Akin, D. (1991) Signal-mediated nuclear transport in proliferating and growth-arrested BALB/c 3T3 cells. *J.Cell Biol.* **115,** 933-939.

Feldherr, C.M. and Akin, D. (1993) Regulation of nuclear transport in proliferating and quiescent cells. *Exp.Cell Res.* **205,** 179-186.

Ferrell, J.E.J. (1998) How regulated protein translocation can produce switch-like responses. *Trends.Biochem.Sci* **23,** 461-465.

Ferrell, J.E., Jr., Wu, M., Gerhart, J.C., and Martin, G.S. (1991) Cell cycle tyrosine phosphorylation of p34cdc2 and a microtubule- associated protein kinase homolog in Xenopus oocytes and eggs. *Mol.Cell Biol.* **11,** 1965-1971.

Fiorini, M., Alimandi, M., Fiorentino, L., Sala, G., and Segatto, O. (2001) Negative regulation of receptor tyrosine kinase signals. *Febs Lett.* **490,** 132-141.

Fischer, U., Huber, J., Boelens, W.C., Mattaj, I.W., and Luhrmann, R. (1995) The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. *Cell* **82,** 475-483.

Florkiewicz, R.Z. and Sommer, A. (1989) Human basic fibroblast growth factor gene encodes four polypeptides: three initiate translation from non-AUG codons. *Proc.Natl.Acad.Sci.U.S.A* **86,** 3978-3981.

Fontes, M.R., Teh, T., and Kobe, B. (2000) Structural basis of recognition of monopartite and bipartite nuclear localization sequences by mammalian importin-alpha. *J Mol Biol* **297,** 1183-1194.

Fukuda, M., Gotoh, I., Adachi, M., Gotoh, Y., and Nishida, E. (1997a) A novel regulatory mechanism in the mitogen-activated protein (MAP) kinase cascade. Role of nuclear export signal of MAP kinase kinase. *J Biol Chem.* **272,** 32642-32648.

Fukuda, M., Gotoh, I., Gotoh, Y., and Nishida, E. (1996) Cytoplasmic localization of mitogen-activated protein kinase kinase directed by its NH2-terminal, leucine-rich short amino acid sequence, which acts as a nuclear export signal. *J Biol Chem.* **271,** 20024-20028.

Fukuda, M., Gotoh, Y., and Nishida, E. (1997b) Interaction of MAP kinase with MAP kinase kinase: its possible role in the control of nucleocytoplasmic transport of MAP kinase. *Embo J* **16,** 1901-1908.

Garcia-Bustos, J., Heitman, J., and Hall, M.N. (1991) Nuclear protein localization. *Biochim.Biophys.Acta* **1071,** 83-101.

Gimenez-Gallego, G., Rodkey, J., Bennett, C., Rios-Candelore, M., DiSalvo, J., and Thomas, K. (1985) Brain-derived acidic fibroblast growth factor: complete amino acid sequence and homologies. *Science* **230,** 1385-1388.

Godsave, S.F., Isaacs, H.V., and Slack, J.M. (1988) Mesoderm-inducing factors: a small class of molecules. *Development* **102,** 555-566.

Goldfarb, D.S., Gariepy, J., Schoolnik, G., and Kornberg, R.D. (1986) Synthetic peptides as nuclear localization signals. *Nature* **322,** 641-644.

Gorlich, D. (1997) Nuclear protein import. *Curr Opin Cell Biol* **9,** 412-419.

Gorlich, D., Kostka, S., Kraft, R., Dingwall, C., Laskey, R.A., Hartmann, E., and Prehn, S. (1995a) Two different subunits of importin cooperate to recognize nuclear localization signals and bind them to the nuclear envelope. *Curr Biol* **5,** 383-392.

Gorlich, D. and Kutay, U. (1999) Transport between the cell nucleus and the cytoplasm. *Annu.Rev.Cell Dev.Biol* **15:607-60,** 607-660.

Gorlich, D. and Mattaj, I.W. (1996) Nucleocytoplasmic transport. *Science* **271,** 1513-1518.

Gorlich, D., Pante, N., Kutay, U., Aebi, U., and Bischoff, F.R. (1996) Identification of different roles for RanGDP and RanGTP in nuclear protein import. *Embo J* **15,** 5584-5594.

Gorlich, D., Prehn, S., Laskey, R.A., and Hartmann, E. (1994) Isolation of a protein that is essential for the first step of nuclear protein import. *Cell* **79,** 767-778.

Gorlich, D., Vogel, F., Mills, A.D., Hartmann, E., and Laskey, R.A. (1995b) Distinct functions for the two importin subunits in nuclear protein import. *Nature* **377,** 246-248.

Gorovsky, M.A. (1973) Macro- and micronuclei of Tetrahymena pyriformis: a model system for studying the structure and function of eukaryotic nuclei. *J.Protozool.* **20,** 19-25.

Gospodarowicz, D. (1974) Localisation of a fibroblast growth factor and its effect alone and with hydrocortisone on 3T3 cell growth. *Nature* **249,** 123-127.

Graham, C.F. and Morgan, R.W. (1966) Changes in the cell cycle during early amphibian development. *Dev.Biol.* **14,** 439-460.

Hartley, R.S., Rempel, R.E., and Maller, J.L. (1996) In vivo regulation of the early embryonic cell cycle in Xenopus. *Dev.Biol.* **173,** 408-419.

Hartley, R.S., Sible, J.C., Lewellyn, A.L., and Maller, J.L. (1997) A role for cyclin E/Cdk2 in the timing of the midblastula transition in Xenopus embryos. *Dev.Biol.* **188,** 312-321.

Hartwell, L.H. and Weinert, T.A. (1989) Checkpoints: controls that ensure the order of cell cycle events. *Science* **246,** 629-634.

Henderson, B.R. and Eleftheriou, A. (2000) A comparison of the activity, sequence specificity, and CRM1-dependence of different nuclear export signals. *Exp.Cell Res.* **256,** 213-224.

Henkel, T., Zabel, U., van Zee, K., Muller, J.M., Fanning, E., and Baeuerle, P.A. (1992) Intramolecular masking of the nuclear location signal and dimerization domain in the precursor for the p50 NF-kappa B subunit. *Cell* **68,** 1121-1133.

Henningfeld, K.A., Friedle, H., Rastegar, S., and Knochel, W. (2002) Autoregulation of Xvent-2B; Direct Interaction and Functional Cooperation of Xvent-2 and Smad1. *J.Biol.Chem.* **277,** 2097-2103.

Hessabi, B., Schmidt, I., and Walther, R. (2000) The homeodomain of Nkx2.2 carries two cooperatively acting nuclear localization signals. *Biochem.Biophys.Res.Commun.* **270,** 695-700.

Hodel, M.R., Corbett, A.H., and Hodel, A.E. (2001) Dissection of a nuclear localization signal. *J.Biol.Chem.* **276,** 1317-1325.

Holmquist, G. (1975) Hoechst 33258 fluorescent staining of Drosophila chromosomes. *Chromosoma* **49,** 333-356.

Hope, I.A. and Struhl, K. (1986) Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. *Cell* **46,** 885-894.

Horiuchi, J., Silverman, N., Marcus, G.A., and Guarente, L. (1995) ADA3, a putative transcriptional adaptor, consists of two separable domains and interacts with ADA2 and GCN5 in a trimeric complex. *Mol.Cell Biol.* **15,** 1203-1209.

Horlein, A.J., Naar, A.M., Heinzel, T., Torchia, J., Gloss, B., Kurokawa, R., Ryan, A., Kamei, Y., Soderstrom, M., Glass, C.K., and . (1995) Ligand-independent repression by the thyroid hormone receptor mediated by a nuclear receptor co-repressor. *Nature* **377,** 397-404.

Hu, W. and Jans, D.A. (1999) Efficiency of importin alpha/beta-mediated nuclear localization sequence recognition and nuclear import. Differential role of NTF2. *J Biol Chem.* **274,** 15820-15827.

Huang, T.T. and Miyamoto, S. (2001) Postrepression activation of NF-kappaB requires the amino-terminal nuclear export signal specific to IkappaBalpha. *Mol.Cell Biol.* **21,** 4737-4747.

Humphrey, G.W., Wang, Y., Russanova, V.R., Hirai, T., Qin, J., Nakatani, Y., and Howard, B.H. (2001) Stable histone deacetylase complexes distinguished by the presence of SANT domain proteins CoREST/kiaa0071 and Mta-L1. *J.Biol.Chem.* **276,** 6817-6824.

Imagawa, M., Sakaue, R., Tanabe, A., Osada, S., and Nishihara, T. (2000) Two nuclear localization signals are required for nuclear translocation of nuclear factor 1-A. *Febs Lett.* **484,** 118-124.

Imamoto, N., Shimamoto, T., Kose, S., Takao, T., Tachibana, T., Matsubae, M., Sekimoto, T., Shimonishi, Y., and Yoneda, Y. (1995) The nuclear pore-targeting complex binds to nuclear pores after association with a karyophile. *Febs Lett.* **368,** 415-419.

Imamura, T., Engleka, K., Zhan, X., Tokita, Y., Forough, R., Roeder, D., Jackson, A., Maier, J.A., Hla, T., and Maciag, T. (1990) Recovery of mitogenic activity of a growth factor mutant with a nuclear translocation sequence. *Science* **249,** 1567-1570.

Isaacs, H.V., Pownall, M.E., and Slack, J.M. (1994) eFGF regulates Xbra expression during Xenopus gastrulation. *EMBO J.* **13,** 4469-4481.

Izaurralde, E. and Adam, S. (1998) Transport of macromolecules between the nucleus and the cytoplasm. *RNA.* **4,** 351-364.

Izaurralde, E., Kann, M., Pante, N., Sodeik, B., and Hohn, T. (1999) Viruses, microorganisms and scientists meet the nuclear pore. Leysin, VD, Switzerland, February 26-March 1, 1998. *Embo J* **18,** 289-296.

Izaurralde, E., Kutay, U., von, K.C., Mattaj, I.W., and Gorlich, D. (1997) The asymmetric distribution of the constituents of the Ran system is essential for transport into and out of the nucleus. *Embo J* **16,** 6535-6547.

Jakel, S., Albig, W., Kutay, U., Bischoff, F.R., Schwamborn, K., Doenecke, D., and Gorlich, D. (1999) The importin beta/importin 7 heterodimer is a functional nuclear import receptor for histone H1. *EMBO J.* **18,** 2411-2423.

Jans, D.A. (1995) The regulation of protein transport to the nucleus by phosphorylation. *Biochem.J.* **311 (Pt 3),** 705-716.

Jans, D.A., Ackermann, M.J., Bischoff, J.R., Beach, D.H., and Peters, R. (1991) p34cdc2-mediated phosphorylation at T124 inhibits nuclear import of SV- 40 T antigen proteins. *J.Cell Biol.* **115,** 1203-1212.

Jans, D.A. and Hubner, S. (1996) Regulation of protein transport to the nucleus: central role of phosphorylation. *Physiol Rev.* **76,** 651-685.

Jans, D.A. and Jans, P. (1994) Negative charge at the casein kinase II site flanking the nuclear localization signal of the SV40 large T-antigen is mechanistically important for enhanced nuclear import. *Oncogene* **9,** 2961-2968.

Jans, D.A., Moll, T., Nasmyth, K., and Jans, P. (1995) Cyclin-dependent kinase site-regulated signal-dependent nuclear localization of the SW15 yeast transcription factor in mammalian cells. *J.Biol.Chem.* **270,** 17064-17067.

Jans, D.A., Xiao, C.Y., and Lam, M.H. (2000) Nuclear targeting signal recognition: a key control point in nuclear transport? *Bioessays* **22,** 532-544.

Jaye, M., Howk, R., Burgess, W., Ricca, G.A., Chiu, I.M., Ravera, M.W., O'Brien, S.J., Modi, W.S., Maciag, T., and Drohan, W.N. (1986) Human endothelial cell growth factor: cloning, nucleotide sequence, and chromosome localization. *Science* **233,** 541-545.

Johnson, D.E. and Williams, L.T. (1993) Structural and functional diversity in the FGF receptor multigene family. *Adv.Cancer Res.* **60,** 1-41.

Jones, C.M. and Smith, J.C. (1999) An overview of Xenopus development. *Methods Mol.Biol.* **97,** 331-340.

Kaffman, A. and O'Shea, E.K. (1999) Regulation of nuclear localization: a key to a door. *Annu.Rev.Cell Dev.Biol.* **15,** 291-339.

Kalderon, D., Richardson, W.D., Markham, A.F., and Smith, A.E. (1984a) Sequence requirements for nuclear location of simian virus 40 large-T antigen. *Nature* **311,** 33-38.

Kalderon, D., Roberts, B.L., Richardson, W.D., and Smith, A.E. (1984b) A short amino acid sequence able to specify nuclear location. *Cell* **39,** 499-509.

Karin, M. (1994) Signal transduction from the cell surface to the nucleus through the phosphorylation of transcription factors. *Curr.Opin.Cell Biol.* **6,** 415-424.

Kassavetis, G.A., Nguyen, S.T., Kobayashi, R., Kumar, A., Geiduschek, E.P., and Pisano, M. (1995) Cloning, expression, and function of TFC5, the gene encoding the B" component of the Saccharomyces cerevisiae RNA polymerase III transcription factor TFIIIB. *Proc.Natl.Acad.Sci.U.S.A* **92,** 9786-9790.

Kataoka, N., Bachorik, J.L., and Dreyfuss, G. (1999) Transportin-SR, a nuclear import receptor for SR proteins. *J Cell Biol* **145,** 1145-1152.

Kay, B.K., Williamson, M.P., and Sudol, M. (2000) The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. *FASEB J.* **14,** 231-241.

Kidd, S. (1992) Characterization of the Drosophila cactus locus and analysis of interactions between cactus and dorsal proteins. *Cell* **71,** 623-635.

Kimelman, D., Kirschner, M., and Scherson, T. (1987) The events of the midblastula transition in Xenopus are regulated by changes in the cell cycle. *Cell* **48,** 399-407.

Kistanova, E., Dell, H., Tsantili, P., Falvey, E., Cladaras, C., and Hadzopoulou-Cladaras, M. (2001) The activation function-1 of hepatocyte nuclear factor-4 is an acidic activator that mediates interactions through bulky hydrophobic residues. *Biochem.J.* **356,** 635-642.

Kleinschmidt, J.A. and Seiter, A. (1988) Identification of domains involved in nuclear uptake and histone binding of protein N1 of Xenopus laevis. *Embo J* **7,** 1605-1614.

Klint, P. and Claesson-Welsh, L. (1999) Signal transduction by fibroblast growth factor receptors. *Front Biosci.* **4,** D165-D177.

Knauf, J.A., Pendergrass, S.H., Marrone, B.L., Strniste, G.F., MacInnes, M.A., and Park, M.S. (1996) Multiple nuclear localization signals in XPG nuclease. *Mutat.Res.* **363,** 67-75.

Koch, C.A., Anderson, D., Moran, M.F., Ellis, C., and Pawson, T. (1991) SH2 and SH3 domains: elements that control interactions of cytoplasmic signaling proteins. *Science* **252,** 668-674.

Kohler, M., Speck, C., Christiansen, M., Bischoff, F.R., Prehn, S., Haller, H., Gorlich, D., and Hartmann, E. (1999) Evidence for distinct substrate specificities of importin alpha family members in nuclear protein import. *Mol Cell Biol* **19,** 7782-7791.

Kouhara, H., Hadari, Y.R., Spivak-Kroizman, T., Schilling, J., Bar-Sagi, D., Lax, I., and Schlessinger, J. (1997) A lipid-anchored Grb2-binding protein that links FGF-receptor activation to the Ras/MAPK signaling pathway. *Cell* **89,** 693-702.

Kumagai, A., Yakowec, P.S., and Dunphy, W.G. (1998) 14-3-3 proteins act as negative regulators of the mitotic inducer Cdc25 in Xenopus egg extracts. *Mol.Biol.Cell* **9,** 345-354.

Kussel, P. and Frasch, M. (1995) Pendulin, a Drosophila protein with cell cycle-dependent nuclear localization, is required for normal cell proliferation. *J.Cell Biol.* **129,** 1491-1507.

Kutay, U., Lipowsky, G., Izaurralde, E., Bischoff, F.R., Schwarzmaier, P., Hartmann, E., and Gorlich, D. (1998) Identification of a tRNA-specific nuclear export receptor. *Mol.Cell* **1,** 359-369.

Lanford, R.E. and Butel, J.S. (1984) Construction and characterization of an SV40 mutant defective in nuclear transport of T antigen. *Cell* **37,** 801-813.

Lanford, R.E., Kanda, P., and Kennedy, R.C. (1986) Induction of nuclear transport with a synthetic peptide homologous to the SV40 T antigen transport signal. *Cell* **46,** 575-582.

Laskey, R.A., Mills, A.D., Philpott, A., Leno, G.H., Dilworth, S.M., and Dingwall, C. (1993) The role of nucleoplasmin in chromatin assembly and disassembly. *Philos.Trans.R.Soc.Lond.B.Biol Sci* **339,** 263-269.

Lemmon, M.A. and Schlessinger, J. (1994) Regulation of signal transduction and signal diversity by receptor oligomerization. *Trends Biochem.Sci.* **19,** 459-463.

Li, X., Shou, W., Kloc, M., Reddy, B.A., and Etkin, L.D. (1994a) Cytoplasmic retention of Xenopus nuclear factor 7 before the mid blastula transition uses a unique anchoring mechanism involving a retention domain and several phosphorylation sites. *J.Cell Biol.* **124,** 7-17.

Li, X., Shou, W., Kloc, M., Reddy, B.A., and Etkin, L.D. (1994b) The association of Xenopus nuclear factor 7 with subcellular structures is dependent upon phosphorylation and specific domains. *Exp.Cell Res.* **213,** 473-481.

Liang, P. and Pardee, A.B. (1992) Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science* **257,** 967-971.

Lin, J., Chen, J., Elenbaas, B., and Levine, A.J. (1994) Several hydrophobic amino acids in the p53 amino-terminal domain are required for transcriptional activation, binding to mdm-2 and the adenovirus 5 E1B 55-kD protein. *Genes Dev.* **8,** 1235-1246.

Loewinger, L. and McKeon, F. (1988) Mutations in the nuclear lamin proteins resulting in their aberrant assembly in the cytoplasm. *EMBO J.* **7,** 2301-2309.

Longshaw, V.M., Dirr, H.W., Blatch, G.L., and Lassle, M. (2000) The in vitro phosphorylation of the co-chaperone mSTI1 by cell cycle kinases substantiates a predicted casein kinase II-p34cdc2-NLS (CcN) motif. *Biol.Chem.* **381,** 1133-1138.

Lowenstein, E.J., Daly, R.J., Batzer, A.G., Li, W., Margolis, B., Lammers, R., Ullrich, A., Skolnik, E.Y., Bar-Sagi, D., and Schlessinger, J. (1992) The SH2 and SH3 domain-containing protein GRB2 links receptor tyrosine kinases to ras signaling. *Cell* **70,** 431-442.

Luchman, H.A., Paterno, G.D., and Gillespie, L.L. (2002) Nuclear localization of XMI-ER1 is regulated by FGF signaling during early development in *Xenopus laevis. manuscript submitted.*

Luchman, H.A., Paterno, G.D., Kao, K.R., and Gillespie, L.L. (1999) Differential nuclear localization of ER1 protein during embryonic development in Xenopus laevis. *Mech.Dev.* **80,** 111-114.

Ma, J. and Ptashne, M. (1987) A new class of yeast transcriptional activators. *Cell* **51,** 113-119.

MacArthur, M.W. and Thornton, J.M. (1991) Influence of proline residues on protein conformation. *J.Mol.Biol.* **218,** 397-412.

Madeo, F., Schlauer, J., Zischka, H., Mecke, D., and Frohlich, K.U. (1998) Tyrosine phosphorylation regulates cell cycle-dependent nuclear localization of Cdc48p. *Mol Biol Cell* **9,** 131-141.

Makkerh, J.S., Dingwall, C., and Laskey, R.A. (1996) Comparative mutagenesis of nuclear localization signals reveals the importance of neutral and acidic amino acids. *Curr Biol* **6,** 1025-1027.

Maller, J.L., Gross, S.D., Schwab, M.S., Finkielstein, C.V., Taieb, F.E., and Qian, Y.W. (2001) Cell cycle transitions in early Xenopus development. *Novartis.Found.Symp.* **237,** 58-73.

Marks,F. (1996) *Protein phosphorylation.* Wiley, John & Sons, incorporated.

Masui, Y. and Wang, P. (1998) Cell cycle transition in early embryonic development of Xenopus laevis. *Biol.Cell* **90,** 537-548.

Meehan, R.R. and Stancheva, I. (2001) DNA methylation and control of gene expression in vertebrate development. *Essays Biochem.* **37,** 59-70.

Melchior, F., Guan, T., Yokoyama, N., Nishimoto, T., and Gerace, L. (1995) GTP hydrolysis by Ran occurs at the nuclear pore complex in an early step of protein import. *J Cell Biol* **131,** 571-581.

Merle, P.L., Feige, J.J., and Verdetti, J. (1995) Basic fibroblast growth factor activates calcium channels in neonatal rat cardiomyocytes. *J.Biol.Chem.* **270,** 17361-17367.

Michael, W.M., Eder, P.S., and Dreyfuss, G. (1997) The K nuclear shuttling domain: a novel signal for nuclear import and nuclear export in the hnRNP K protein. *Embo J* **16,** 3587-3598.

Miller, M., Reddy, B.A., Kloc, M., Li, X.X., Dreyer, C., and Etkin, L.D. (1991) The nuclear-cytoplasmic distribution of the Xenopus nuclear factor, xnf7,

coincides with its state of phosphorylation during early development. *Development* **113,** 569-575.

Mitchison, J.M. and Creanor, J. (1971) Further measurements of DNA synthesis and enzyme potential during cell cycle of fission yeast Schizosaccharomyces pombe. *Exp.Cell Res.* **69,** 244-247.

Mo, Y.Y., Wang, C., and Beck, W.T. (2000) A novel nuclear localization signal in human DNA topoisomerase I. *J.Biol.Chem.* **275,** 41107-41113.

Mohammadi, M., Honegger, A.M., Rotin, D., Fischer, R., Bellot, F., Li, W., Dionne, C.A., Jaye, M., Rubinstein, M., and Schlessinger, J. (1991) A tyrosine-phosphorylated carboxy-terminal peptide of the fibroblast growth factor receptor (Flg) is a binding site for the SH2 domain of phospholipase C-gamma 1. *Mol.Cell Biol.* **11,** 5068-5078.

Moll, T., Tebb, G., Surana, U., Robitsch, H., and Nasmyth, K. (1991) The role of phosphorylation and the CDC28 protein kinase in cell cycle- regulated nuclear import of the S. cerevisiae transcription factor SWI5. *Cell* **66,** 743-758.

Moreland, R.B., Nam, H.G., Hereford, L.M., and Fried, H.M. (1985) Identification of a nuclear localization signal of a yeast ribosomal protein. *Proc.Natl.Acad.Sci.U.S.A* **82,** 6561-6565.

Moroianu, J. (1998) Distinct nuclear import and export pathways mediated by members of the karyopherin beta family. *J Cell Biochem.* **70,** 231-239.

Moroianu, J. (1999) Nuclear import and export: transport factors, mechanisms and regulation. *Crit.Rev.Eukaryot.Gene Expr.* **9,** 89-106.

Moroianu, J., Blobel, G., and Radu, A. (1995a) Previously identified protein of uncertain function is karyopherin alpha and together with karyopherin beta docks import substrate at nuclear pore complexes. *Proc Natl Acad Sci U S A* **92,** 2008-2011.

Moroianu, J., Hijikata, M., Blobel, G., and Radu, A. (1995b) Mammalian karyopherin alpha 1 beta and alpha 2 beta heterodimers: alpha 1 or alpha 2 subunit binds nuclear localization signal and beta subunit interacts with peptide repeat-containing nucleoporins. *Proc Natl Acad Sci U S A* **92,** 6532-6536.

Mosialos, G., Hamer, P., Capobianco, A.J., Laursen, R.A., and Gilmore, T.D. (1991) A protein kinase-A recognition sequence is structurally linked to

transformation by p59v-rel and cytoplasmic retention of p68c-rel. *Mol.Cell Biol.* **11,** 5867-5877.

Nakai, K. and Horton, P. (1999) PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. *Trends Biochem.Sci.* **24,** 34-36.

Newport, J. and Kirschner, M. (1982a) A major developmental transition in early Xenopus embryos: I. characterization and timing of cellular changes at the midblastula stage. *Cell* **30,** 675-686.

Newport, J. and Kirschner, M. (1982b) A major developmental transition in early Xenopus embryos: II. Control of the onset of transcription. *Cell* **30,** 687-696.

Nieuwkoop,P. and J.Faber. (1967) *Normal Table of Xenopus laevis The stages of Xenopus embryonic development.* Daudin, Amsterdam, The Netherlands, Garland Publishing Inc, New York.

Nigg, E.A. (1997) Nucleocytoplasmic transport: signals, mechanisms and regulation. *Nature* **386,** 779-787.

Nothias, J.Y., Majumder, S., Kaneko, K.J., and DePamphilis, M.L. (1995) Regulation of gene expression at the beginning of mammalian development. *J.Biol.Chem.* **270,** 22077-22080.

Ong, S.H., Guy, G.R., Hadari, Y.R., Laks, S., Gotoh, N., Schlessinger, J., and Lax, I. (2000) FRS2 proteins recruit intracellular signaling pathways by binding to diverse targets on fibroblast growth factor and nerve growth factor receptors. *Mol.Cell Biol.* **20,** 979-989.

Ornitz, D.M. and Itoh, N. (2001) Fibroblast growth factors. *Genome Biol.* **2,** REVIEWS3005.

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

Paterno, G.D., Li, Y., Luchman, H.A., Ryan, P.J., and Gillespie, L.L. (1997) cDNA cloning of a novel, developmentally regulated immediate early gene activated by fibroblast growth factor and encoding a nuclear protein. *J.Biol.Chem.* **272,** 25591-25595.

Paterno, G.D., Mercer, F.C., Chayter, J.J., Yang, X., Robb, J.D., and Gillespie, L.L. (1998) Molecular cloning of human er1 cDNA and its

differential expression in breast tumours and tumour-derived cell lines. *Gene* **222,** 77-82.

Pawson, T. (1995) Protein modules and signalling networks. *Nature* **373,** 573-580.

Pawson, T. and Saxton, T.M. (1999) Signaling networks--do all roads lead to the same genes? *Cell* **97,** 675-678.

Peng, C.Y., Graves, P.R., Thoma, R.S., Wu, Z., Shaw, A.S., and Piwnica-Worms, H. (1997) Mitotic and G2 checkpoint control: regulation of 14-3-3 protein binding by phosphorylation of Cdc25C on serine-216. *Science* **277,** 1501-1505.

Peterson, C.L. and Tamkun, J.W. (1995) The SWI-SNF complex: a chromatin remodeling machine? *Trends Biochem.Sci.* **20,** 143-146.

Picard, D., Salser, S.J., and Yamamoto, K.R. (1988) A movable and regulable inactivation function within the steroid binding domain of the glucocorticoid receptor. *Cell* **54,** 1073-1080.

Post, J.N., Gillespie, L.L., and Paterno, G.D. (2001) Nuclear localization signals in the Xenopus FGF embryonic early response 1 protein. *Febs Lett.* **502,** 41-45.

Powers, C.J., McLeskey, S.W., and Wellstein, A. (2000) Fibroblast growth factors, their receptors and signaling. *Endocr.Relat Cancer* **7,** 165-197.

Radtke, T., Schmalz, D., Coutavas, E., Soliman, T.M., and Peters, R. (2001) Kinetics of protein import into isolated Xenopus oocyte nuclei. *Proc.Natl.Acad.Sci.U.S.A* **98,** 2407-2412.

Ramsdell, A.F. and Yost, H.J. (1998) Molecular mechanisms of vertebrate left-right development. *Trends Genet.* **14,** 459-465.

Rexach, M. and Blobel, G. (1995) Protein import into nuclei: association and dissociation reactions involving transport substrate, transport factors, and nucleoporins. *Cell* **83,** 683-692.

Ribbeck, K. and Gorlich, D. (2001) Kinetic analysis of translocation through nuclear pore complexes. *EMBO J.* **20,** 1320-1330.

Richards, S.A., Carey, K.L., and Macara, I.G. (1997) Requirement of guanosine triphosphate-bound ran for signal-mediated nuclear protein export [see comments]. *Science* **276,** 1842-1844.

Richards, S.A., Lounsbury, K.M., Carey, K.L., and Macara, I.G. (1996) A nuclear export signal is essential for the cytosolic localization of the Ran binding protein, RanBP1. *J.Cell Biol.* **134,** 1157-1168.

Richardson, W.D., Mills, A.D., Dilworth, S.M., Laskey, R.A., and Dingwall, C. (1988) Nuclear protein migration involves two steps: rapid binding at the nuclear envelope followed by slower translocation through nuclear pores. *Cell* **52,** 655-664.

Robbins, J., Dilworth, S.M., Laskey, R.A., and Dingwall, C. (1991) Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence. *Cell* **64,** 615-623.

Roberts, B. (1989) Nuclear location signal-mediated protein transport. *Biochim.Biophys.Acta* **1008,** 263-280.

Roberts, B.L., Richardson, W.D., and Smith, A.E. (1987) The effect of protein context on nuclear location signal function. *Cell* **50,** 465-475.

Robinson, M.J. and Cobb, M.H. (1997) Mitogen-activated protein kinase pathways. *Curr.Opin.Cell Biol.* **9,** 180-186.

Rout, M.P. and Aitchison, J.D. (2001) The nuclear pore complex as a transport machine. *J.Biol.Chem.* **276,** 16593-16596.

Rout, M.P. and Blobel, G. (1993) Isolation of the yeast nuclear pore complex. *J.Cell Biol.* **123,** 771-783.

Roux, P., Blanchard, J.M., Fernandez, A., Lamb, N., Jeanteur, P., and Piechaczyk, M. (1990) Nuclear localization of c-Fos, but not v-Fos proteins, is controlled by extracellular signals. *Cell* **63,** 341-351.

Rupp, R.A., Snider, L., and Weintraub, H. (1994) Xenopus embryos regulate the nuclear localization of XMyoD. *Genes Dev.* **8,** 1311-1323.

Sano, H., Forough, R., Maier, J.A., Case, J.P., Jackson, A., Engleka, K., Maciag, T., and Wilder, R.L. (1990) Detection of high levels of heparin binding growth factor-1 (acidic fibroblast growth factor) in inflammatory arthritic joints. *J.Cell Biol.* **110,** 1417-1426.

Schmidt-Zachmann, M.S., Dargemont, C., Kuhn, L.C., and Nigg, E.A. (1993) Nuclear export of proteins: the role of nuclear retention. *Cell* **74,** 493-504.

Schoorlemmer, J. and Goldfarb, M. (2001) Fibroblast growth factor homologous factors are intracellular signaling proteins. *Curr.Biol.* **11,** 793-797.

Schwab, M.S. and Dreyer, C. (1997) Protein phosphorylation sites regulate the function of the bipartite NLS of nucleolin. *Eur.J Cell Biol* **73,** 287-297.

Schwamborn, K., Albig, W., and Doenecke, D. (1998) The histone H1(0) contains multiple sequence elements for nuclear targeting. *Exp.Cell Res.* **244,** 206-217.

Scorilas, A., Kyriakopoulou, L., Yousef, G.M., Ashworth, L.K., Kwamie, A., and Diamandis, E.P. (2001) Molecular cloning, physical mapping, and expression analysis of a novel gene, BCL2L12, encoding a proline-rich protein with a highly conserved BH2 domain of the Bcl-2 family. *Genomics* **72,** 217-221.

Shapiro, H.M. (1981) Flow cytometric estimation of DNA and RNA content in intact cells stained with Hoechst 33342 and pyronin Y. *Cytometry* **2,** 143-150.

Shieh, M.W., Wessler, S.R., and Raikhel, N.V. (1993) Nuclear targeting of the maize R protein requires two nuclear localization sequences. *Plant Physiol.* **101,** 353-361.

Shiokawa, K., Kurashima, R., and Shinga, J. (1994) Temporal control of gene expression from endogenous and exogenously-introduced DNAs in early embryogenesis of *Xenopus laevis. Int.J.Dev.Biol.* **38,** 249-255.

Shitaka, Y., Matsuki, N., Saito, H., and Katsuki, H. (1996) Basic fibroblast growth factor increases functional L-type Ca2+ channels in fetal rat hippocampal neurons: implications for neurite morphogenesis in vitro. *J.Neurosci.* **16,** 6476-6489.

Shou, W., Li, X., Wu, C., Cao, T., Kuang, J., Che, S., and Etkin, L.D. (1996) Finely tuned regulation of cytoplasmic retention of Xenopus nuclear factor 7 by phosphorylation of individual threonine residues. *Mol Cell Biol* **16,** 990-997.

Siomi, H. and Dreyfuss, G. (1995) A nuclear localization domain in the hnRNP A1 protein. *J Cell Biol* **129,** 551-560.

Slack, J.M., Darlington, B.G., Heath, J.K., and Godsave, S.F. (1987) Mesoderm induction in early Xenopus embryos by heparin-binding growth factors. *Nature* **326,** 197-200.

Smallwood, P.M., Munoz-Sanjuan, I., Tong, P., Macke, J.P., Hendry, S.H., Gilbert, D.J., Copeland, N.G., Jenkins, N.A., and Nathans, J. (1996) Fibroblast growth factor (FGF) homologous factors: new members of the FGF family implicated in nervous system development. *Proc.Natl.Acad.Sci.U.S.A* **93,** 9850-9857.

Smith, L.D., Xu, W.L., and Varnold, R.L. (1991) Oogenesis and oocyte isolation. *Methods Cell Biol.* **36,** 45-60.

Solari, F., Bateman, A., and Ahringer, J. (1999) The Caenorhabditis elegans genes egl-27 and egr-1 are similar to MTA1, a member of a chromatin regulatory complex, and are redundantly required for embryonic patterning. *Development* **126,** 2483-2494.

Sparks, A.B., Rider, J.E., Hoffman, N.G., Fowlkes, D.M., Quillam, L.A., and Kay, B.K. (1996) Distinct ligand preferences of Src homology 3 domains from Src, Yes, Abl, Cortactin, p53bp2, PLCgamma, Crk, and Grb2. *Proc.Natl.Acad.Sci.U.S.A* **93,** 1540-1544.

Speir, E., Sasse, J., Shrivastav, S., and Casscells, W. (1991) Culture-induced increase in acidic and basic fibroblast growth factor activities and their association with the nuclei of vascular endothelial and smooth muscle cells. *J.Cell Physiol* **147,** 362-373.

Stancheva, I., El-Maarri, O., Walter, J., Niveleau, A., and Meehan, R.R. (2002) DNA methylation at promoter regions regulates the timing of gene activation in *Xenopus laevis* embryos. *Dev.Biol* **243,** 155-165.

Standiford, D.M. and Richter, J.D. (1992) Analysis of a developmentally regulated nuclear localization signal in Xenopus. *J.Cell Biol.* **118,** 991-1002.

Sterner, D.E., Wang, X., Bloom, M.H., and Berger, S.L. (2002) The SANT domain of Ada2 is required for normal acetylation of histones by the yeast SAGA complex. *J.Biol.Chem.*

Stewart, M., Baker, R.P., Bayliss, R., Clayton, L., Grant, R.P., Littlewood, T., and Matsuura, Y. (2001) Molecular mechanism of translocation through nuclear pore complexes during nuclear protein import. *Febs Lett.* **498,** 145-149.

Szebenyi, G. and Fallon, J.F. (1999) Fibroblast growth factors as multifunctional signaling factors. *Int.Rev.Cytol.* **185,** 45-106.

Takemoto, Y., Tashiro, S., Handa, H., and Ishii, S. (1994) Multiple nuclear localization signals of the B-myb gene product. *Febs Lett.* **350,** 55-60.

Tata, J.R. (1996) Metamorphosis: an exquisite model for hormonal regulation of post- embryonic development. *Biochem.Soc.Symp.* **62,** 123-136.

Tata, J.R. (1999) Amphibian metamorphosis as a model for studying the developmental actions of thyroid hormone. *Biochimie* **81,** 359-366.

Tolwinski, N.S. and Wieschaus, E. (2001) Armadillo nuclear import is regulated by cytoplasmic anchor Axin and nuclear anchor dTCF/Pan. *Development* **128,** 2107-2117.

Torok, I., Strand, D., Schmitt, R., Tick, G., Torok, T., Kiss, I., and Mechler, B.M. (1995) The overgrown hematopoietic organs-31 tumor suppressor gene of Drosophila encodes an Importin-like protein accumulating in the nucleus at the onset of mitosis. *J.Cell Biol.* **129,** 1473-1489.

Tratner, I. and Verma, I.M. (1991) Identification of a nuclear targeting sequence in the Fos protein. *Oncogene* **6,** 2049-2053.

Turner, D.L. and Weintraub, H. (1994) Expression of achaete-scute homolog 3 in Xenopus embryos converts ectodermal cells to a neural fate. *Genes Dev.* **8,** 1434-1447.

Umbhauer, M., Penzo-Mendez, A., Clavilier, L., Boucaut, J., and Riou, J. (2000) Signaling specificities of fibroblast growth factor receptors in early Xenopus embryo. *J.Cell Sci.* **113 (Pt 16),** 2865-2875.

Vandromme, M., Gauthier-Rouviere, C., Lamb, N., and Fernandez, A. (1996) Regulation of transcription factor localization: fine-tuning of gene expression. *Trends Biochem.Sci.* **21,** 59-64.

Veenstra, G.J., Mathu, M.T., and Destree, O.H. (1999) The Oct-1 POU domain directs developmentally regulated nuclear translocation in Xenopus embryos. *Biol.Chem.* **380,** 253-257.

Wang, B., Zou, J.X., Ek-Rylander, B., and Ruoslahti, E. (2000) R-Ras contains a proline-rich site that binds to SH3 domains and is required for integrin activation by R-Ras. *J.Biol.Chem.* **275,** 5222-5227.

Wang, J.K., Xu, H., Li, H.C., and Goldfarb, M. (1996) Broadly expressed SNT-like proteins link FGF receptor stimulation to activators of Ras. *Oncogene* **13,** 721-729.

Weighardt, F., Biamonti, G., and Riva, S. (1995) Nucleo-cytoplasmic distribution of human hnRNP proteins: a search for the targeting domains in hnRNP A1. *J Cell Sci* **108,** 545-555.

Weis, K., Mattaj, I.W., and Lamond, A.I. (1995) Identification of hSRP1 alpha as a functional receptor for nuclear localization sequences. *Science* **268,** 1049-1053.

Wen, W., Meinkoth, J.L., Tsien, R.Y., and Taylor, S.S. (1995) Identification of a signal for rapid export of proteins from the nucleus. *Cell* **82,** 463-473.

White, E.M., Allis, C.D., Goldfarb, D.S., Srivastva, A., Weir, J.W., and Gorovsky, M.A. (1989) Nucleus-specific and temporally restricted localization of proteins in Tetrahymena macronuclei and micronuclei. *J.Cell Biol.* **109,** 1983-1992.

Wiedlocha, A., Falnes, P.O., Madshus, I.H., Sandvig, K., and Olsnes, S. (1994) Dual mode of signal transduction by externally added acidic fibroblast growth factor. *Cell* **76,** 1039-1051.

Williamson, M.P. (1994) The structure and function of proline-rich regions in proteins. *Biochem.J.* **297 (Pt 2),** 249-260.

Wolpert,L., R.Beddington, J.Brockes, T.Jessell, P.Lawrence, and E.Meyerowitz. (1998) *Principles of Development.* Current Biology Ltd., London, U.K.

Yasuda, G.K. and Schubiger, G. (1992) Temporal regulation in the early embryo: is MBT too good to be true? *Trends Genet.* **8,** 124-127.

You, A., Tong, J.K., Grozinger, C.M., and Schreiber, S.L. (2001) CoREST is an integral component of the Co. *Proc.Natl.Acad.Sci.U.S.A* **98,** 1454-1458.

Zabel, U., Henkel, T., Silva, M.S., and Baeuerle, P.A. (1993) Nuclear uptake control of NF-kappa B by MAD-3, an I kappa B protein present in the nucleus. *EMBO J.* **12,** 201-211.

Zhang, F., White, R.L., and Neufeld, K.L. (2000) Phosphorylation near nuclear localization signal regulates nuclear import of adenomatous polyposis coli protein. *Proc.Natl.Acad.Sci.U.S.A* **97,** 12577-12582.

Phosphorylation site prediction for sequence

Sequence identifier:
Number of residues: **493**
Full sequence:

```
MAEPSLRTASPGGSAASDDHEFEPSADMLVHEFDDEQTLEEEEMLEGEVNFTSEIEHLERESEMPIDELLR
LYGYGSTVP
LPGEEDEEDMDNDCNSGCSGEIKDEAIKDSSGQEDETQSSNDDPTPSFTCRDVREVIRPRRCKYFDTNHEI
EEESEDDED
YVPSEDWKKEIMVGSMFQAEIPVGICKYRETEKVYENDDQLLWNPEYVMEERVIDFLNEASRRTCEERGLD
AIPEGSHIK
DNEQALYEHVKCNFDTEEALRRLRFNVKAAREELSVWTEEECRNFEQGLKAYGKDFHLIQANKVRTRSVGE
CVAFYYMWK
KSERYDFFAQQTRFGKKKYNLHPGVTDYMDRLLDESESATSSRAPSPPPTTSNSNTSQSEKEDCTASNNTQ
NGVSVNGPC
AITAYKDEAKQGVHLNGPTISSSDPSSNETDTNGYNRENVTDDSRFSHTSGKTDTNPDDTNERPIKRQRMD
SPGKESTGS
SEFSQEVFSHGEV
```

The prediction is based on the consensus sequence motifs. For the list of the used motifs see here.

The consensus sequences for 10 different protein kinases are scanned for potential phosphorylation sites.

Potential phosphorylation sites for protein kinase CaMII:

- S-10 RTASPGG
- S-308 RTRSVGE
- S-366 RAPSPPP
- S-472 RMDSPGK

Potential phosphorylation sites for protein kinase CKI (N-terminal S/T must be prephosphorylated):

- T-8 SLRTASP
- S-14 SPGGSAAS
- S-99 SGCSGEI
- S-120 TQSSNDD
- T-129 TPSFTCRD
- T-224 SRRTCEE
- T-278 SVWTEEE
- T-360 SESATSSR
- S-366 SRAPSPPP
- S-374 TTSNSNTS
- S-379 TSQSEKE

- T-390 SNNTQNG
- S-423 TISSSDPS
- T-430 SSNETDTN
- S-444 TDDSRFS
- S-450 SHTSGKT
- S-481 STGSSEFS

Potential phosphorylation sites for protein kinase CKII:

- T-38 DEQTLEE
- S-53 NFTSEIE
- S-111 KDSSGQE
- S-119 ETQSSND
- T-129 PSFTCRD
- T-147 YFDTNHE
- S-155 EEESEDD
- T-224 SRRTCEE
- T-278 SVWTEEE
- S-308 RTRSVGE
- S-377 SNTSQSE
- S-421 PTISSSD
- S-426 SDPSSNE
- T-455 KTDTNPD
- S-489 EVFSHGE

Potential phosphorylation sites for protein kinase GSK3 (C-terminal +4 S must be prephosphorylated):

- S-10 RTASPGGS
- S-358 ESESATSS
- S-362 ATSSRAPS
- T-370 PPPTTSNS
- T-419 NGPTISSS
- S-423 ISSSDPSS
- S-477 GKESTGSS

No phosphorylation sites for protein kinase MLCK found.

Potential phosphorylation sites for protein kinase p34cdc2:

- S-472 RMDSPGK

Potential phosphorylation sites for protein kinase p70s6k:

- S-308 KVRTRSVGE
- S-472 RQRMDSPGK

Potential phosphorylation sites for protein kinase PKA:

- S-10 RTASPGG
- S-62 RESEMP
- T-191 RETEKV
- T-224 RRTCEE
- S-308 RTRSVGE
- S-366 RAPSPPP
- S-447 RFSHTS
- S-472 RMDSPGK

Potential phosphorylation sites for protein kinase PKC:

- S-5 AEPSLRT
- T-129 PSFTCRD
- T-191 YRETEKV
- S-221 NEASRRT
- S-322 WKKSERY
- S-361 SATSSRA
- S-379 TSQSEKE
- S-450 SHTSGKT

No phosphorylation sites for protein kinase PKG found.

NetPhos 2.0 Prediction Results

```
  493 Sequence
MAEPSLRTASPGGSAASDDHEFEPSADMLVHEFDDEQTLEEEEMLEGEVNFTSEIEHLERESEMPIDELLR
LYGYGSTVP      80
LPGEEDEEDMDNDCNSGCSGEIKDEAIKDSSGQEDETQSSNDDPTPSFTCRDVREVIRPRRCKYFDTNHEI
EEESEDDED      160
YVPSEDWKKEIMVGSMFQAEIPVGICKYRETEKVYENDDQLLWNPEYVMEERVIDFLNEASRRTCEERGLD
AIPEGSHIK      240
DNEQALYEHVKCNFDTEEALRRLRFNVKAAREELSVWTEEECRNFEQGLKAYGKDFHLIQANKVRTRSVGE
CVAFYYMWK      320
KSERYDFFAQQTRFGKKKYNLHPGVTDYMDRLLDESESATSSRAPSPPPTTSNSNTSQSEKEDCTASNNTQ
NGVSVNGPC      400
AITAYKDEAKQGVHLNGPTISSSDPSSNETDTNGYNRENVTDDSRFSHTSGKTDTNPDDTNERPIKRQRMD
SPGKESTGS      480
SEFSQEVFSHGEV
560
.........S......S.....................T.................S........S.........
.........      80
.................S..............SS.......SS........T..............Y.......
...S.....      160
Y.................................T...Y...........Y.................T.......
.........      240
......Y.....................................................................S...
.........      320
.S..Y......T.................Y.......S.S..S....S....TS....S.S............
...S.....      400
....Y................S....SS.....................S..S.........T............
S....S..S      480
S..S....S....
560

Phosphorylation sites predicted:       Ser: 31 Thr: 7  Tyr: 8
```

Serine predictions

Name	Pos	Context (v)	Score	Pred
Sequence	5	MAEPSLRTA	0.073	.
Sequence	10	LRTASPGGS	0.993	*S*
Sequence	14	SPGGSAASD	0.250	.
Sequence	17	GSAASDDHE	0.614	*S*
Sequence	25	EFEPSADML	0.020	.
Sequence	53	VNFTSEIEH	0.804	*S*
Sequence	62	LERESEMPI	0.942	*S*
Sequence	77	YGYGSTVPL	0.042	.
Sequence	96	NDCNSGCSG	0.804	*S*
Sequence	99	NSGCSGEIK	0.248	.

Sequence	110	AIKDSSGQE	0.705	*S*
Sequence	111	IKDSSGQED	0.991	*S*
Sequence	119	DETQSSNDD	0.967	*S*
Sequence	120	ETQSSNDDP	0.986	*S*
Sequence	127	DPTPSFTCR	0.280	.
Sequence	155	IEEESEDDE	0.996	*S*
Sequence	164	DYVPSEDWK	0.035	.
Sequence	175	IMVGSMFQA	0.027	.
Sequence	221	LNEASRRTC	0.059	.
Sequence	237	IPEGSHIKD	0.028	.
Sequence	275	REELSVWTE	0.179	.
Sequence	308	VRTRSVGEC	0.995	*S*
Sequence	322	MWKKSERYD	0.609	*S*
Sequence	356	LLDESESAT	0.975	*S*
Sequence	358	DESESATSS	0.869	*S*
Sequence	361	ESATSSRAP	0.993	*S*
Sequence	362	SATSSRAPS	0.301	.
Sequence	366	SRAPSPPPT	0.996	*S*
Sequence	372	PPTTSNSNT	0.562	*S*
Sequence	374	TTSNSNTSQ	0.413	.
Sequence	377	NSNTSQSEK	0.996	*S*
Sequence	379	NTSQSEKED	0.997	*S*
Sequence	387	DCTASNNTQ	0.003	.
Sequence	395	QNGVSVNGP	0.728	*S*
Sequence	421	GPTISSSDP	0.820	*S*
Sequence	422	PTISSSDPS	0.056	.
Sequence	423	TISSSDPSS	0.072	.
Sequence	426	SSDPSSNET	0.996	*S*
Sequence	427	SDPSSNETD	0.982	*S*
Sequence	444	VTDDSRFSH	0.329	.
Sequence	447	DSRFSHTSG	0.997	*S*
Sequence	450	FSHTSGKTD	0.950	*S*
Sequence	472	QRMDSPGKE	0.997	*S*
Sequence	477	PGKESTGSS	0.975	*S*
Sequence	480	ESTGSSEFS	0.959	*S*
Sequence	481	STGSSEFSQ	0.886	*S*
Sequence	484	SSEFSQEVF	0.969	*S*
Sequence	489	QEVFSHGEV	0.973	*S*

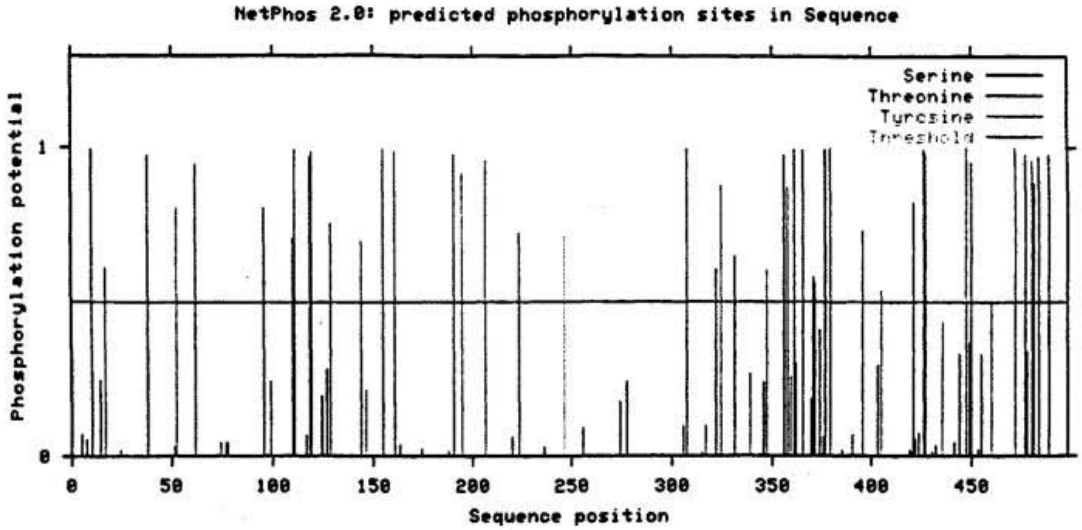
Threonine predictions

Name	Pos	Context	Score	Pred
Sequence	8	PSLRTASPG	0.056	.
Sequence	38	DDEQTLEEE	0.977	*T*
Sequence	52	EVNFTSEIE	0.030	.
Sequence	78	GYGSTVPLP	0.041	.
Sequence	117	QEDETQSSN	0.070	.
Sequence	125	NDDPTPSFT	0.197	.
Sequence	129	TPSFTCRDV	0.752	*T*
Sequence	147	KYFDTNHEI	0.216	.

Sequence	191	KYRETEKVY	0.972	*T*
Sequence	224	ASRRTCEER	0.725	*T*
Sequence	256	CNFDTEEAL	0.095	.
Sequence	278	LSVWTEEEC	0.247	.
Sequence	306	NKVRTRSVG	0.100	.
Sequence	332	FAQQTRFGK	0.652	*T*
Sequence	346	HPGVTDYMD	0.237	.
Sequence	360	SESATSSRA	0.259	.
Sequence	370	SPPPTTSNS	0.181	.
Sequence	371	PPPTTSNSN	0.580	*T*
Sequence	376	SNSNTSQSE	0.059	.
Sequence	385	KEDCTASNN	0.020	.
Sequence	390	ASNNTQNGV	0.069	.
Sequence	403	PCAITAYKD	0.297	.
Sequence	419	LNGPTISSS	0.020	.
Sequence	430	SSNETDTNG	0.013	.
Sequence	432	NETDTNGYN	0.037	.
Sequence	441	RENVTDDSR	0.043	.
Sequence	449	RFSHTSGKT	0.366	.
Sequence	453	TSGKTDTNP	0.018	.
Sequence	455	GKTDTNPDD	0.330	.
Sequence	460	NPDDTNERP	0.512	*T*
Sequence	478	GKESTGSSE	0.336	.

Tyrosine predictions

Name	Pos	Context	Score	Pred
Sequence	73	LLRLYGYGS	0.006	.
Sequence	75	RLYGYGSTV	0.042	.
Sequence	144	RRCKYFDTN	0.701	*Y*
Sequence	161	DDEDYVPSE	0.986	*Y*
Sequence	188	GICKYRETE	0.018	.
Sequence	195	TEKVYENDD	0.911	*Y*
Sequence	207	WNPEYVMEE	0.954	*Y*
Sequence	247	EQALYEHVK	0.709	*Y*
Sequence	292	GLKAYGKDF	0.006	.
Sequence	316	CVAFYYMWK	0.013	.
Sequence	317	VAFYYMWKK	0.101	.
Sequence	325	KSERYDFFA	0.876	*Y*
Sequence	339	GKKKYNLHP	0.270	.
Sequence	348	GVTDYMDRL	0.610	*Y*
Sequence	405	AITAYKDEA	0.537	*Y*
Sequence	435	DTNGYNREN	0.435	.


NetPhos 2.0: predicted phosphorylation sites in Sequence
Serine
Threonine
Tyrosine
Threshold
Phosphorylation potential
1
0
0
50
100
150
200
250
300
350
400
450
Sequence position

CONSTRUCT	PRIMER	SEQUENCE
EGFP-ER1 1-493	Er14	gaagatctcaaacatggcggagccttcactcag
	Er13	gaagatctttaaacctctccatgtg
EGFP-ER1 1-137	Er14	
	Er26	ctagatctgattacttctcgtacatctc
EGFP-ER1 1-144	Er14	
	Er28	ctagatctcttgcaccgacgtggacggattac
EGFP-ER1 145-250	Er29	agagatctattttgatacaaatcatgaaatagaagagg
	Er30	ctagatcttacatgttcatatagggcctgc
EGFP-ER1 251-319	Er31	agagatcttaaaatgcaattttgacacagaagaggc
	Er25	gaagatctgtagattatacttctttttttcc
EGFP-ER1 251-340	Er31	
	Er32	ctagatctccacatgtagtagaatgccacacattc
EGFP-ER1 339-493	Er27	gaagatctcgtataatctacatcctggtgtaacgg
	Er13	
EGFP-ER1 1-461	Er14	
	Er34	ggtcgaccgtttgtatcatctggatttgtgtcag
EGFP-ER1 1-473	Er14	
	Er35	ggtcgaccagggctgtccatacgttgcc
EGFP-ER1 339-461	Er27	
	Er34	
EGFP-ER1 339-473	Er27	
	Er35	

EGFP-NLS-1	Er23K	gatctgcgtaatccgtccacgtcggtgcaaga
	Er33	gatctcttgcaccgacgtggacggattacgca
EGFP-NLS-2	Er36s	Gatctaggcattgagaagactaagatttaatgtcaaagccgccag agaagaag
	Er36a	tcgacttcttctctggcggctttgacattaaatcttagtcttctcaatgcct a
EGFP-NLS-3	Er24	agagatctactacatgtggaaaaaatcagaacg
	Er25	
EGFP-NLS-4 (a)	Er37s	gatctcaaacgaaaggccaataaaaaggcaacgtatggacg
	Er37as	tcgacgtccatacgttgccttttattggcctttcgtttga
EGFP-NLS-4 (b)	Er38	gagatctatccagatgatacaaacg
	Er39	cagatctcttcccagggctgtcc
Egfp-NLS-4 (stop, c)	Er38	
	Er52	gagatctacttcccagggctgtccatacg
SPGK → NPGK (erf.I)	Er40K	ggcaacgtatggacaaccctgggaaggaaag
	Er40R	ctttccttcccagggttgtccatacgttgcc
SPGK → APGK (erfI)	Er41K	ggcaacgtatggacgcccctgggaaggaaag
	Er41R	ctttccttcccaggggcgtccatacgttgcc
SPGK → DGPK (erfI)	Er53K	ggcaacgtatggacgaccctgggaaggaaagtac
	Er53R	gtactttccttcccagggtcgtccatacgttgcc
SPGK → APGK (NLS4B)	ER45K	ggcaacgtatggacgcccctgggaagagatc
	ER45R	gatctcttcccaggggcgtccatacgttgcc

SPGK → DPGK (nls4b)	Er46K	ggcaacgtatggacgaccctgggaagagatctcg
	Er46R	cgagatctcttcccagggtcgtccatacgttgcc
EGFP-ER446-493	Er50	caagatctgatttagtcatacaagtgg
	Er13	
ΔPIKR	Er51K	ccagatgatacaaacgaaaggcaacgtatggacagccctggg
	Er51R	cccagggctgtccatacgttgcctttcgtttgtatcatctgg
ΔNLS1	Xen1s	gtacgagaagtaatctgcaagtattttgatac
	Xen1as	gtatcaaaatacttgcagattacttctcgtac
ΔNLS3	Xnls2s	ctacatgtggtataatctacatcctggtg
	Xnls2as	ggatgtagattataccacatgtagtagaatgccac
ΔKKK (bipartite)	ERN2-k3s	gcccaacaaacacgatttggatataatctacatcctgg
	ERN2-k3a	ccaggatgtagattatatccaaatcgtgtttgttgggc
KK→AA (bipartite)	Er21K	ggcattctactacatgtgggcagcatcagaacgttatgacttc
	Er21R	gaagtcataacgttctgatgctgcccacatgtagtagaatgcc
For cloning into pT7Ts	Gfp-s	gctctagacgctagcgctaccggtc
	Gfp-as	gctctagaggctgattatgatcag
For cloning into nβ-gal-CS2+ and cβ-gal-CS2+	PEGFP-5'	gctctagaccggactcagatc
For colony screening and sequencing of EGFP-C_2 plasmids	PCZ-5'	cactctcggcatggacgagc
	PC2-3'	cccgcggtaccgtcgactgc

Xnf7 bipartite NLS	Xnf7s	tcgagactccccagaagagaaagattgaggaacctgagcctgaaccgaagaaagcaaagc
	Xnf7as	tcgagctttgctttcttcggttcaggctcaggttcctcaatctttctcttctgggagtc
β-gal-xnf7NLS-ER1-282	Er54	gcctctagacatggcggagccttcactcagg
	Er55	cggtctagactacattcttcttcagtcc
β-gal-xnf7NLS-ER267-493	Er56	gcctctagacgtcaaagccgccagagaag
	Er57	cggtctagattaaacctctccatgtgaaaac