

CHARACTERISATION OF THE MURINE CLK1 DUAL-SPECIFICITY KINASE

Peter I. Duncan

Thesis submitted to the Department of Biochemistry in partial fulfillment of the
requirements for the degree of Doctor of Philosophy

University of Ottawa
Ottawa, Ontario, Canada

October, 1996

© Peter I. Duncan, Ottawa, Canada, 1996



National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-612-19950-9

Canada



UNIVERSITÉ D'OTTAWA
UNIVERSITY OF OTTAWA

ABSTRACT

Murine Clk1 (also known as Sty) was identified as a dual-specificity kinase capable of phosphorylating serine, threonine and tyrosine residues when expressed in bacteria. Expression of *Clk1* is developmentally regulated at the level of RNA expression. Embryonic cells express two mRNAs of similar size, whereas differentiated cells express two additional transcripts of larger size. Little is known about the biochemical and biological activity of Clk1 in mammalian cells.

We demonstrate that the two embryonically expressed mRNAs are derived through alternative splicing of a unique exon (exon EB) within the *Clk1* precursor mRNA. These mRNAs encode full-length catalytically active Clk1 and a truncated inactive polypeptide, Clk1^T. Larger incompletely spliced *Clk1* transcripts accumulate in differentiated cells.

When expressed in mammalian cells Clk1 possesses dual-specificity kinase activity and is capable of forming complexes with other molecules of Clk1 and Clk1^T. The regions involved in binding map to the amino-terminal non-catalytic domain of Clk1. Phosphorylation sites map to the amino acids encoded by the alternatively spliced exon EB. Clk1^T and catalytic mutants of Clk1 co-localise with splicing factors in intranuclear speckles, whereas catalytically active Clk1 causes a redistribution of these factors within the nucleus. This activity requires the presence of amino acids encoded by exon EB. These results suggest a role for Clk1 and Clk1^T in the regulation of RNA splicing.

Splicing of a *Clk1* mini-gene, encompassing exon EB, in vivo is regulated by Clk1 and Clk1^T. Catalytically active Clk1 stimulates exclusion of EB leading to the production of *Clk1^T* mRNA. In contrast, Clk1^T promotes EB inclusion leading to production of *Clk1* mRNA.

Two Clk1-related human kinases, hClk2 and hClk3, also exhibit dual-specificity kinase activity and cause the redistribution of nuclear splicing factors. Similar to Clk1^T, the hClk truncated isoforms, hClk2^T and hClk3^T, co-localise with splicing factors in nuclear speckles. hClk2 and hClk3 are able to influence the splicing pattern of a murine *Clk1* mini-gene in vivo, indicating that they can also regulate precursor mRNA splicing. Taken together these results imply a role for the Clk family of kinases in the regulation of gene expression at the level of RNA processing.

-II-

in memory of Iain K. Duncan (Greenock 1939-1990)

ACKNOWLEDGEMENTS

I would like to express my gratitude to some of the many who have put up with me over the years:

- In particular I would like to thank Dr. John Bell who after repeated pestering took me on as a summer student in 1987, and who is now finally rid of me. He was a never ending source of ideas, optimism and laughs. And you think I'm cynical.
- Drs. Doug Gray and Mike McBurney for the early WIPings and continuing interest in my work.
- To Beth Mason for her cheerful "Hellos" and wonderful administrative skills; Oh, has she saved me a few headaches.
- To the "originals" Elizabeth Douville, Danny Afar and Brian Howell for welcoming me into the lab at McGill and again in Ottawa. The last of the novice art collectors is leaving town. A special thanks to Liz for her friendship over the years.
- Ricardo Marius for introducing me to the thrills of the Gatineau Park's Meech Lake trail and for the dust I swallowed following him down it. Those early morning jaunts with Liz and Danny were always an adventure.
- Ninan Abraham, for keeping me entertained with his arguments on music and politics with all willing (pr)opponents. We've been benchmates for so long I'm not sure how I'll handle the divorce. May the force be with you when you come to defend (y)our "territory".
- Emma Tibbo, for introducing me to the world of aerobics and laughing at my attempts at coordinated moves; I'll see you on the volleyball court wee lassie!
- Patrons of the Friday night beer club (Alan, Christine, Dave, Mike, Nat, Simon, Tim, and others) for dinners, dancing, the 'hot-seat' and other oft-missed events.
- To Paul Waqué for beer and cider creations and long chats.
- A special thanks to my parents, Margaret and Iain, who have encouraged and supported me throughout my education.

PREFACE

The work presented in this thesis has been or will be submitted for publication as described below:

CHAPTER 2

Duncan, P.I., Howell, B.W., Marius, R.M., Drmanic, S., Douville, E.M.J. and Bell, J.C. (1995) Alternative splicing of STY, a nuclear dual specificity kinase. *J. Biol. Chem.* **270**, 21524-21531.

CHAPTER 3

Colwill, K., Pawson, T., Andrews, B., Prasad, J., Manley, J.L., Bell, J.C. and Duncan, P.I. (1996) The Clk/Sty protein kinase phosphorylates SR splicing factors and regulates their intranuclear distribution. *EMBO J.* **15**, 265-275.

CHAPTER 4

Duncan, P.I., Stojdl, D.F., Marius, R.M. and Bell, J.C.
Splicing of the SR Protein Kinase Clk1 by an Autoregulatory Mechanism.
Manuscript in preparation.

CHAPTER 6

Duncan, P.I., Stojdl, D.F., Marius, R.M. and Bell, J.C.
The Clk2 and Clk3 Dual-Specificity Protein Kinases Regulate the Intranuclear Distribution of SR proteins and Influence Pre-mRNA Splicing. Manuscript in preparation.

The following people have contributed to the experimental work presented in this thesis: Dr. D. A. Gray (Department of Medicine, University of Ottawa) provided the human liver cDNA and mouse ES cell genomic libraries; Suzana Drmanic sequenced the Clk1^T cDNA; David F. Stojdl performed all RT-PCR analyses; confocal microscopy was performed at the CLSM Facility (Department of Biology, University of Ottawa) with the technical assistance of Andrew Vaillant; Ricardo M. Marius synthesised the Clk1 peptide and provided technical assistance throughout.

TABLE OF CONTENTS

ABSTRACT	I
ACKNOWLEDGEMENTS	III
PREFACE	IV
TABLE OF CONTENTS	V
LIST OF FIGURES AND TABLES	VIII
LIST OF ABBREVIATIONS	XI
 CHAPTER 1 - GENERAL INTRODUCTION	 1
(I) Protein Kinases	2
The Sty/Clk Kinase	5
The Sty/Clk Family of Kinases	6
hClk1, hClk2 and hClk3	6
Doa	7
AFC1, AFC2 and AFC3	8
KNS1	9
(II) RNA SPLICING	9
Requirements for Splicing	11
SR Proteins	13
Spliceosome regulation through phosphorylation	16
Splicing factor localisation	18
THESIS WORK	19
 CHAPTER 2 - Alternative Splicing of Clk1, a Nuclear Dual Specificity Kinase . . .	 20
SUMMARY	21
INTRODUCTION	21
MATERIALS AND METHODS	23
Construction of cDNA library	23
Isolation of cDNA clones	23
Sequencing of cDNA clones	24
Northern hybridization analysis	24
RT-PCR analysis and cloning of <i>Clk1^T</i>	25
Sequencing of RT-PCR products	25
Identification of Additional Introns	26
Construction of cDNA Expression Vectors and Tagged Proteins . . .	26
COS-1 cell culture and transfection	27
Immunoprecipitation/Immunoblot analysis	28
Immunofluorescence of M-Clk1 in COS-1 cells	29
GST-Clk1 binding columns	29
Nucleotide sequence accession numbers	29

RESULTS	30
Sequence analysis of the developmentally induced <i>Clk1</i>	
transcripts reveals introns	30
<i>Clk1</i> -3.2 and <i>Clk1</i> -5.6kb transcripts are nuclear	32
<i>Clk1</i> is alternatively spliced	34
Expression of Clk1 protein in mammalian cells	38
Clk1 protein exhibits dual specificity kinase activity	
<i>in vitro</i> and <i>in vivo</i> .	40
Nuclear localization of Clk1 protein	42
Clk1 protein can dimerize	44
DISCUSSION	46
 CHAPTER 3 - The Clk1 Protein Kinase Regulates the Intranuclear	
Distribution of SR Splicing Factors	49
SUMMARY	50
INTRODUCTION	50
MATERIALS AND METHODS	53
Immunofluorescence Assay	53
Triton X-100 Extraction	53
RESULTS	55
Subnuclear localization of Clk1 and SR proteins <i>in vivo</i>	55
DISCUSSION	60
Clk1 interacts with RNA-Binding Proteins	60
Phosphorylation of ASF/SF2	61
Subnuclear Localization of Clk	61
 CHAPTER 4 - Splicing of the SR Protein Kinase Clk1 by	
an Autoregulatory Mechanism	63
SUMMARY	64
MATERIALS AND METHODS	65
Construction of Plasmids	65
COS-1/293T Transfections	67
RT-PCR Analysis	68
Immunoblot Analysis	68
Immunofluorescence	69
RESULTS AND DISCUSSION	70
 CHAPTER 5 - Dissection of the Regulatory Domain of Clk1	78
INTRODUCTION	79
MATERIALS AND METHODS	81
Plasmid Constructions - Clk1 Deletions	81
Junction Fragments	83
GST-Clk1 Deletions	84

COS-1 Cell Expression of Clk1 Polypeptides	85
GST GST-Clk1 Purification	85
GST GST-Clk1 Binding Assay	86
Peptide Synthesis and Phosphorylation	86
RESULTS	88
Phosphorylation Sites	88
Tyrosine 160 Phosphorylation	95
Identification of Clk1:Clk1 Interaction Domains	98
Binding of K190R mutants	102
Subnuclear "Specklising" Domain	104
SR "De-Specklising" Domain	106
DISCUSSION	110
CHAPTER 6 - The Clk2 and Clk3 Dual-Specificity Protein Kinases Regulate the Intranuclear Distribution of SR proteins and Influence Pre-mRNA Splicing	
SUMMARY	114
INTRODUCTION	115
MATERIALS AND METHODS	119
Construction of Plasmids	119
Cell Culture and Transfection	120
Immunoprecipitation/Immunoblot Analysis	121
Phosphoamino acid analysis	121
Immunofluorescence Assay	122
RT-PCR Analysis	122
RESULTS	124
Expression of hClk2 and hClk3 Proteins in Mammalian Cells	124
hClk2 and hClk3 Proteins Exhibit Dual-Specificity Kinase Activity <i>in Vitro</i> and <i>in Vivo</i>	124
Truncated hClk proteins co-localise with splicing factors	128
The hClk kinases initiate the redistribution of SR proteins	130
Clk2 and Clk3 regulate splicing <i>in vivo</i>	130
DISCUSSION	134
CONCLUSION	137
REFERENCES	144

LIST OF FIGURES AND TABLES

CHAPTER 1

Figure 1	
Examples of MAPK signalling cascades in vertebrates and yeast.	4
Figure 2	
Pre-mRNA splicing mechanism and the sequence requirements	12
Figure 3	
SR protein-mediated interactions in spliceosome assembly.	17

CHAPTER 2

Figure 1	
Nucleotide and predicted amino acid sequences of the <i>Clk1</i> -5.6 partial cDNA and <i>Clk1^T</i> cDNA.	31
Figure 2	
Schematic representation of <i>Clk1</i> transcripts and analysis of their subcellular localization	33
Table 1	
Intron and exon boundaries found in the <i>Clk1</i> -5.6 cDNA.	35
Figure 3	
Schematic representation of the alternative splicing of exon B	35
Figure 4	
RT-PCR analysis of <i>Clk1</i> transcripts demonstrates alternative splicing.	37
Figure 5	
Expression and catalytic activity of wild-type and mutant Clk1 proteins in COS-1 cells	39
Figure 6	
Clk1 protein is phosphorylated on tyrosine <i>in vivo</i>	41
Figure 7	
Subcellular localization of wild-type and mutant Clk1 proteins.	43
Figure 8	
Clk1 can dimerize <i>in vitro</i>	45

CHAPTER 3

Figure 1	
Nuclear localization of Clk1 and its effects on the distribution of SR proteins	56
Figure 2	
Effect of Clk1 on the distribution of Sm proteins.	58
Figure 3	
Active and inactive Clk/STY kinases display differential solubility in non-ionic detergent	59

CHAPTER 4

Figure 1	
Catalytic activity of Clk1 kinase regulates splicing of <i>Clk1</i> pre-mRNA <i>in vivo</i>	71
Figure 2	
Catalytically active Clk1 promotes exon skipping of <i>Clk1</i> pre-mRNA <i>in vivo</i>	73
Figure 3	
Clk1 ^T promotes exon inclusion of <i>Clk1</i> pre-mRNA <i>in vivo</i>	75
Figure 4	
The peptide encoded by the <i>Clk1</i> alternatively spliced exon is necessary for exon skipping activity	77

CHAPTER 5

Figure 1	
Schematic Representation of Clk1 Regulatory Domain Deletions.	89
Figure 2	
Expression and catalytic activity of Clk1 deletion mutants in COS-1 cells.	90
Figure 3	
Bacterial expression and catalytic activity of Clk1 ^{Δ62-160}	92
Figure 4	
Phosphorylation of Clk1 peptide by bacterially expressed Clk1 proteins.	94
Figure 5	
Phosphorylation of Clk1 peptide by Clk1 proteins expressed in COS-1 cells	96
Figure 6	
Catalytic activities of Clk1 Y160 mutants in COS-1 cells.	97
Figure 7	
Binding of Clk1 deletion mutants to GST-Clk1 <i>in vitro</i>	99
Figure 8	
Binding of Clk1 deletion mutants to GST-Clk1 <i>in vitro</i>	101
Figure 9	
Binding of Clk1 ^{K190R} deletion mutants to GST-Clk1 <i>in vitro</i>	103
Figure 10	
Binding of Clk1 ^{K190R} deletion mutants to GST-Clk1 <i>in vitro</i>	105
Figure 11	
Nuclear localization of Clk1 deletion mutants and their effects on SR protein distribution.	107
Figure 12	
Nuclear localization of Clk1 deletion mutants and their effects on SR protein distribution.	108
Figure 13	
Amino acids encoded by the alternatively spliced exon are necessary for SR protein redistribution	109

Figure 14	
Summary of effects of Clk1 deletions on Clk1 binding and	
SR protein redistribution	112

CHAPTER 6

Figure 1	
Expression and catalytic activity of hClk2 and hClk3 proteins in COS-1 cells	125
Figure 2	
hClk2 and hClk3 are phosphorylated on tyrosine in vivo.	127
Figure 3	
Nuclear Localization of hClk2 and hClk2 ^T and their effects	
on the distribution of SR proteins.	129
Figure 4	
Nuclear Localization of hClk3 and hClk3 ^T and their effects	
on the distribution of SR proteins.	131
Figure 5	
hClk2 and hClk3 regulate alternative splicing <i>in vivo</i>	133

CONCLUSION

Figure 1	
Schematic representation of Clk pre-mRNA splicing and	
translated protein products.	138
Figure 2	
A model for mClk1-mediated splicing regulation of the <i>mClk1</i>	
pre-mRNA through a splicing repressor element	140
Figure 3	
A model for mClk1-mediated splicing regulation of the <i>mClk1</i>	
pre-mRNA through an exon splicing enhancer	141

LIST OF ABBREVIATIONS

<i>AFC</i>	<i>Arabidopsis thaliana fus3</i> -complementing gene
ASF	Alternative splicing factor
bp	Base pair
cDNA	Complementary DNA
Chap.	Chapter
Clk	CDC28/ <i>cdc2</i> ⁺ -like kinase
CMV	Cytomegalovirus promoter
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleoside triphosphate
<i>Doa</i>	<i>Darkener of apricot</i>
DSK	Dual-specificity kinase
E	Exon
EB	Exon B
EDTA	Ethylenediaminetetraacetic acid
ERK	Extracellular-Regulated Kinase
FITC	Fluorescein isothiocyanate
GST	Glutathione-S-transferase
h	Hour
HA	Haemagglutinin epitope

hClk	Human Clk
HEPES	N-2-hydroxyethylpiperazine-N'-2-ethanesulphonic acid
hnRNP	Heterogenous nuclear RNP
I	Intron
IB	Immunoblot or Intron B
Ig	Immunoglobulin
IP	Immunoprecipitation
kb	kilobase pair
kDa	kiloDalton
<i>KNS1</i>	<i>Kinase Next to Spa2</i>
mAb	Monoclonal antibody
MAPK	Mitogen-Activated Protein Kinase
mClk	Murine Clk
MEK	MAPK/ERK Kinase
MEM	Minimal essential medium
mRNA	Messenger RNA
NLS	Nuclear localization signal
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
pre-mRNA	Precursor mRNA
pSer or pS	Phosphoserine
PSK	Protein serine/threonine kinases

-XIII-

pThr or pT	Phosphothreonine
pTyr or pY	Phosphotyrosine
PVDF	Polyvinylidene difluoride
PYK	Protein tyrosine kinases
RBD	RNA binding domain
RNA	Ribonucleic acid
RNP	Ribonucleoprotein particle
RS domain	Arginine/serine-rich domain
RT-PCR	Reverse transcriptase coupled with PCR
SAPK	Stress-Activated Protein Kinase
SDS	Sodium dodecyl sulfate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
SF2	Splicing factor 2
SH2 domain	Src homology 2 domain
snRNP	Small nuclear ribonucleoprotein particle
SR protein	Serine/arginine-rich protein
Sty kinase	Serine, threonine and tyrosine kinase
U	Unit
U2AF	U2 snRNP auxiliary factor
UTR	Untranslated region

CHAPTER 1

GENERAL INTRODUCTION

(I) Protein Kinases

Eukaryotic cells are continually responding to environmental cues during their growth and development. To capture and transduce these signals, cells have evolved complex signalling mechanisms, with protein phosphorylation, one of the most frequently observed post-translational modifications in eukaryotic cells, frequently being central to their action. Phosphorylation is achieved through the enzymatic activity of protein kinases and occurs predominantly on the three hydroxy-amino acids, serine, threonine and tyrosine.

Kinases have historically been divided into two families, protein tyrosine kinases (PYK) and protein serine/threonine kinases (PSK), based upon both amino acid sequence conservation within the catalytic domain and upon the amino acid(s) phosphorylated by that catalytic domain (1). Protein tyrosine kinases constitute approximately one third of all protein kinases identified (2,3), however, phosphotyrosine accounts for less than 0.05% of phosphorylated hydroxy-amino acids in eukaryotic cells (4). Such low levels of phosphotyrosine may be reflective of the exquisite control of PYK activity within the cell. This control underlies the prominence of PYKs within vital signal transduction pathways.

Within the last five years a new family of kinases, the dual-specificity kinases (DSK), has gained prominence. Members of this family have the ability to phosphorylate serine/threonine and/or tyrosine residues (reviewed in (5,6)). The majority of lower eukaryotes do not contain genes encoding DSKs or PYKs, thus it has been suggested that DSKs and PYKs evolved from an ancestral PSK (7). There are now approximately fifteen members of the DSK family in mammalian cells (5,6), however, the physiological

substrates of many remain unknown. A common feature of several DSKs, with known substrates, is their participation in protein kinase cascade signalling pathways.

Signalling through kinase cascades is typified by the MAPK (**Mitogen-Activated Protein Kinase**, also known as ERK (**Extracellular-Regulated Kinase**)) network having a core of three sequentially activated kinases which can amplify and transmit the extracellular signal (recently reviewed in (8,9), Fig. 1). Two similar pathways have been identified and well characterised. The first is stimulated by extracellular growth factors (MAPK pathway) while the second relies on cytokines or agents causing cellular stress (ie. radiation, osmotic shock) and is now known as the SAPK (**Stress-Activated Protein Kinase**) pathway (10). The MAPKs and SAPKs are themselves related proteins as are the kinases preceding them in the cascades.

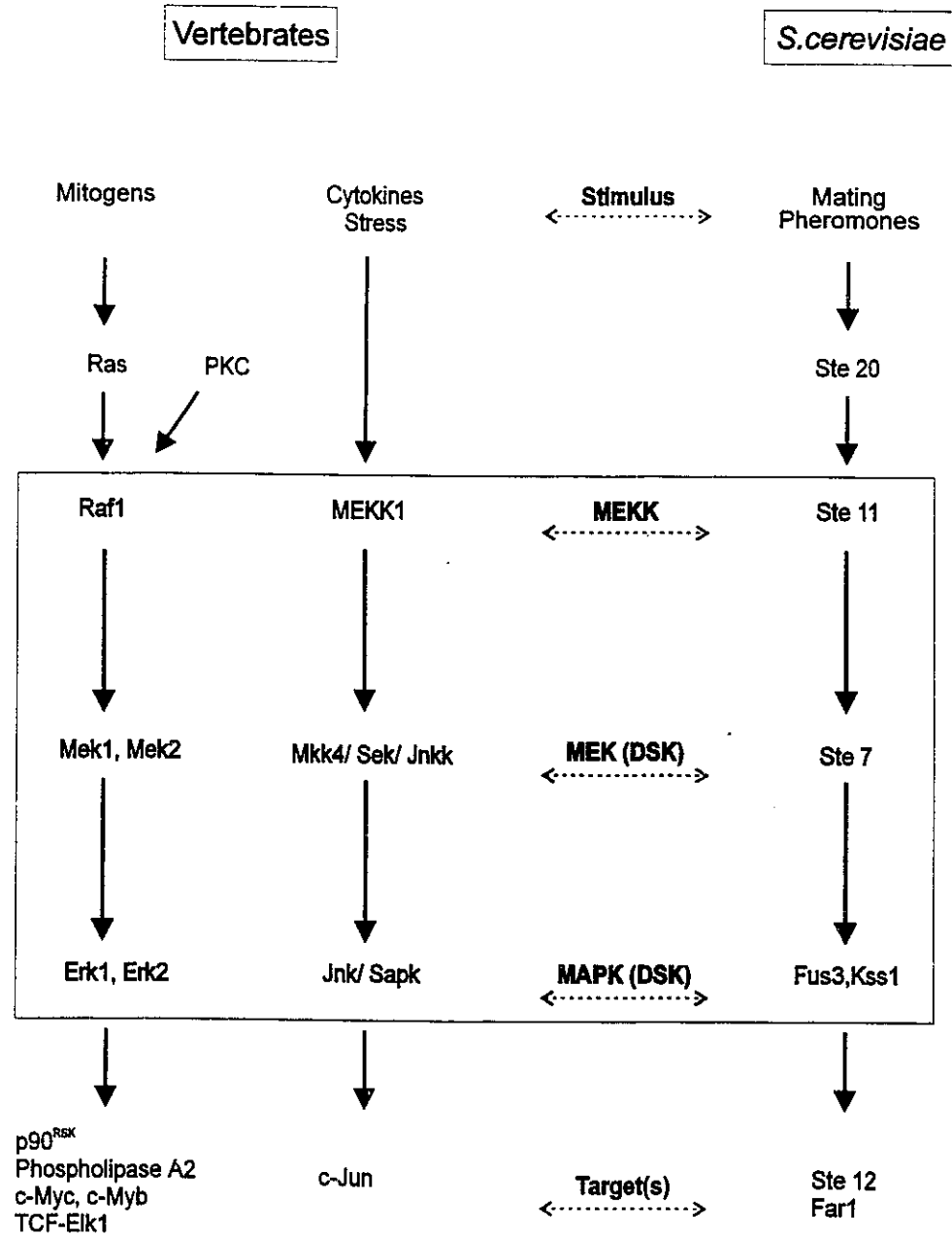
MAPKs are activated by one of the MEKs (**MAPK/ERK Kinase**) which are in turn activated by the Raf family of kinases (in this naming system Raf is a MEK Kinase or MEKK). Both MAPK and MEK have been shown to be dual-specificity kinases (11-13). MAPK is predominately a broad spectrum serine/threonine kinase which autophosphorylates to a low stoichiometry on tyrosine (14). The regulatory nature of this tyrosine phosphorylation is unclear. The MEKs were the first subfamily of DSKs for which the physiological substrates were shown to be phosphorylated with dual-specificity. The MEKs phosphorylate MAPKs with threonine/tyrosine specificity (12,13) and are themselves activated through serine/threonine phosphorylation by Raf kinases (15).

Figure 1

Examples of MAPK signalling cascades in vertebrates and yeast.

Relationship of the signalling molecules in the MAPK pathway stimulated by mitogens, the SAPK pathway stimulates by cytokines and stress, and the yeast, *Saccharomyces cerevisiae*, mating pheromone stimulated pathway. The conserved protein kinase cascade is outlined with a grey box. Dashed arrows indicate the related kinases. DSK - dual-specificity kinase.

Figure 1



It has become clear that many enzymes require phosphorylation on both serine/threonine and tyrosine residues (16,17). This is exemplified, as mentioned above, by the obligatory phosphorylation of MAPKs on both tyrosine and threonine required for kinase activation (16). These dual phosphorylations can be achieved through the action of two protein kinases: a PSK and a PYK. Alternatively, a single DSK could achieve the same result. The evolutionary conservation of DSKs within the MAPK pathways, from yeast to humans, suggests that there has been selective pressure for their maintenance.

The Sty/Clk Kinase

The Sty/Clk kinase was isolated by virtue of its tyrosine-phosphorylating activity but upon sequencing the catalytic domain was found to be similar to protein kinases of the serine/threonine family (18,19). Sty/Clk can be divided into two major domains; a carboxyl-terminal catalytic domain (of approximately 320 amino acids) which has highest homology to members of the CDC28/*cdc2* family of serine/threonine kinases (hence the name Clk, CDC28/*cdc2*-like kinase), and an amino-terminal non-catalytic domain (approximately 160 amino acids) which has little homology to other known proteins. Strongest similarity is observed between the catalytic domain and the yeast *FUS3* gene product (~30% amino acid homology (18,19)). The amino-terminal non-catalytic region contains a putative nuclear localisation signal (NLS) reminiscent of the NLS of the simian vacuolating virus 40 large T antigen, suggesting that Sty/Clk may be localised to the nucleus.

Further analysis revealed that when expressed and isolated from bacteria *Sty/Clk* could autophosphorylate on the three hydroxy-amino acids, Serine, threonine and tyrosine (*Sty*), placing it in the DSK family. *In vitro* *Sty/Clk1* autophosphorylates on serine and tyrosine residues to approximately equal stoichiometry and on threonine to a lesser extent (18,19).

Sty/Clk mRNAs show what appears to be a ubiquitous pattern of expression in adult mouse tissues, however, expression of *Sty/Clk* RNA is developmentally regulated. Embryonic stem cells express two mRNA species approximately 1.7 and 1.8-kb in size, while differentiation of embryonic cells *in vitro* leads to the expression of two additional *Sty/Clk* transcripts approximately 3.2 and 5.6-kb in size (18). RNA derived from various tissues of the adult mouse appear to contain the four species of *Sty/Clk* mRNA (18,19). Interestingly, the leukaemic cell lines SP10 (myeloma) and P388 (pre-B cell), like embryonic stem cells, express only the smaller (1.7/1.8-kb) transcripts (18).

The *Sty/Clk* Family of Kinases

Since embarking on the work described in this thesis several other protein kinases have been cloned that share considerable sequence similarity to that of *Sty/Clk*.

hClk1, hClk2 and hClk3

Johnson and Smith (20) isolated a human homologue, also named Clk (*cdc2*-like kinase; herein named hClk1 for human Clk1)), using a polymerase chain reaction (PCR) based approach with degenerate oligonucleotides designed to identify *cdc2*-related kinases.

This kinase is 87% identical to *Sty/Clk*, with the major difference being the presence of an additional thirty amino acids adjacent to the catalytic domain in *Sty Clk*. Kinase activity, although predicted for hClk1, has not been demonstrated so it is unclear whether it is itself a dual-specificity kinase.

Hanes *et al.* independently cloned hClk1 and two additional related kinases (hClk2 and hClk3 with 56 and 51% identity to hClk1, respectively) (21), however, once again kinase activity has not been demonstrated. The genes for these kinases display a wide pattern of expression, similar to that of *Sty/Clk* and interestingly, *hClk1*, 2 and 3 show an RNA splicing pattern similar to that of *Sty/Clk* ((21) and see Chap. 2).

Doa

The *Doa* (*Darkener of apricot*) gene was identified in *Drosophila melanogaster* as a dosage-sensitive modifier of the *apricot* allele (w^a) of the *white eye* (*w*) locus (22). The w^a allele is caused by the insertion of the copia retrotransposon into the second intron of the *w* gene. Phenotypically this results in flies, normally displaying red eyes, having “apricot-coloured” eyes. *Doa* mutations are dominant in their suppression of the w^a phenotype, thus, flies expressing mutations in the *Doa* gene have eyes which are darker than “apricot-coloured”. *Doa* interacts specifically with the copia element, with mutant alleles elevating levels of copia transcripts. Mutations of *Doa* result in defects in segmentation, and eye and nervous system development. Most recessive mutations are lethal, supporting the idea that the *Doa* gene product is vital for fly development. *Doa* is expressed throughout development, and both stage and tissue-specific transcripts are

produced. The *Doa* gene product possesses intrinsic protein kinase activity although the amino acid specificity is currently unknown.

The involvement of Doa in the regulation of copia expression, and its predicted nuclear localisation raise the possibility that it and other Sty/Clk family members may phosphorylate nuclear proteins regulating gene expression.

AFC1, AFC2 and AFC3

Bender and Fink (23) isolated *AFC1* (*Arabidopsis thaliana fus3*-complementing gene 1) in a screen to identify genes in higher plants that could complement the yeast *Saccharomyces cerevisiae* with mutations in their MAPK signalling pathways. Haploid yeast with mutations in both MAPK homologues, *fus3* and *kss1*, are sterile and cannot mate as these kinases are unable to phosphorylate and activate the STE12 transcription factor, a requirement for the expression of mating-specific genes. *AFC1* is able to complement the *fus3/kss1* double mutant resulting in mating-specific gene expression in haploid yeast and mating of haploid yeast to yield diploids. Surprisingly, *AFC1* is not a MAPK homologue but shows similarity to Sty/Clk family members (41% identity to Sty/Clk kinase domain (23)). Bender and Fink have also isolated two genes related to *AFC1* called *AFC2* (75% identity to *AFC1*) and *AFC3* (68% identity) (23), and as with *AFC1* the amino acid specificity of these kinases is unknown.

KNS1

KNS1 (*Kinase Next to Spa2*) was cloned from *Saccharomyces cerevisiae* (24) and displays approximately 40% amino acid identity to mClk1 within the catalytic domain (22). Little is known about the function of *KNS1*. When deleted or over expressed in yeast *KNS1* displays no overt phenotype suggesting that, unlike that of *Doa*, it is a non-essential gene (24).

Thus, the Sty/Clk family of kinases has been evolutionarily conserved from yeast to man arguing that these kinases perform function(s) which confer a selective advantage to the organism.

For simplicity in this thesis the mouse and human kinases will be referred to as mClk and hClk, respectively. Sty/Clk, being the founding member of this family has been given the designation mClk1 with its corresponding human homologue called hClk1.

(II) RNA SPLICING

The control of gene expression is critical for normal cellular growth and differentiation. There are four possible steps where this regulation can occur: 1) The structure of the DNA encoding and surrounding the gene: expressed genes are found in euchromatin which can facilitate the accessibility of the transcription machinery; 2) transcriptionally, through the coordinated regulation of transcription factors (both integral components of the transcription machinery as well as *trans*-acting factors); 3) post-

transcriptionally at the RNA level, and 4) post-translationally on the protein stage. I will focus here on post-transcriptional regulation through RNA splicing.

Following transcription in eukaryotes the primary RNA transcript of a gene must undergo a complex series of modifications before it becomes a mature, readily translatable messenger RNA (mRNA). These modifications include capping of the 5' end with 7-methyl-guanosine, polyadenylation of the 3' end, and splicing to remove the intervening non-coding sequences between exons. With the successful completion of these processes the mRNA must be transported from the nucleus to the cytoplasm where the translation machinery resides and turnover of the mRNA can occur. All these processes are potential sites for the regulation of gene expression and are controlled by a multitude of interactions between RNAs and RNA binding proteins.

RNA splicing has been recognized as a means whereby higher eukaryotes can regulate expression and generate isoform diversity within individual genes (25). Precursor mRNAs (pre-mRNAs) can undergo alternative splicing to generate diverse functional mRNAs. Alternative splicing can involve the use of one splice site (5' or 3') and multiple complementary splice sites, or exon/intron "skipping", where entire exons are removed or introns retained. In some instances, what may be perceived as regulation through alternative splicing may actually result from differential promoter or polyadenylation site usage (26). Regulated RNA processing at this level may involve more than one of the above mentioned processes which are not necessarily mutually exclusive.

Requirements for Splicing

Figure 2 outlines the sequence requirements within pre-mRNAs for splicing to occur, as well as a general scheme of the splicing reaction. Cis-splicing elements (5' splice site (5'ss), 3'ss, polypyrimidine tract and the branch point sequence) are not highly conserved in mammalian RNAs and mutation within these sequences often does not completely inhibit splicing but can affect the efficiency of splicing and/or activate cryptic splice sites (25).

The splicing reaction requires the coordinated action of a large number of protein and RNA-protein complexes combined as a macromolecular enzyme termed the spliceosome. The major sub-units of the spliceosome are the U1, U2, U4, U5 and U6 RNA-protein complexes termed snRNPs (small nuclear ribonucleoprotein particles). The snRNPs contain a common set of polypeptides referred to as the Sm proteins, as well as proteins unique to the particular snRNP (27). Both RNA- and protein-interactions between the snRNPs themselves, and the intron are believed to impart the correct structure permitting the catalysis of splicing to occur.

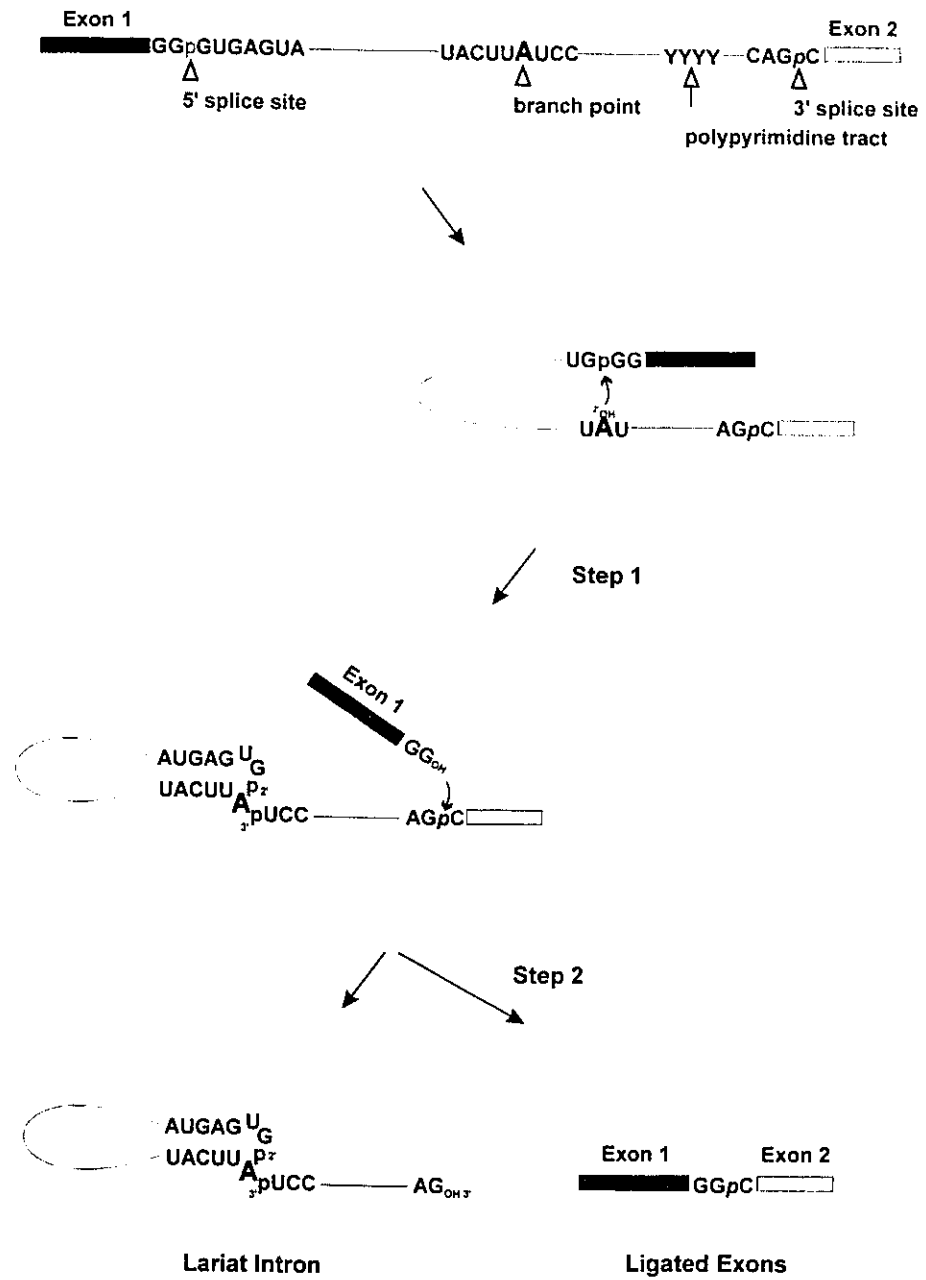
In addition to the snRNPs, spliceosomes also contain a group of non-snRNP splicing factors. Many of these factors contain conserved protein motifs which confer functions that are consistent with their involvement in the splicing reaction. These motifs include the RNA binding domain (RBD) and RGG box (both of which are RNA binding domains), the DEAD/DEAH box (characteristic of a family of RNA helicases) and arginine/serine-rich (RS) domains (consisting of repeats of arginine-serine dipeptides). Many of the non-snRNP splicing factors contain one or more of these motifs (28). For

Figure 2

Pre-mRNA splicing mechanism and the sequence requirements.

Splicing of precursor mRNA requires 5' and 3' splice sites, a polypyrimidine tract and a branch point. The branch point adenosine necessary for catalysis is highlighted. The pre-mRNA/spliceosome complex creates the necessary structure to allow the two transesterification reactions (Step 1 and 2) to occur. In Step 1 the 2' hydroxyl (2'OH) group of the branch point adenosine initiates cleavage at the 5' splice site producing a 2'-5' phosphodiester bond between the adenosine and the 5' end of the intron. Step 2 requires the free 3' hydroxyl of Exon 1 to initiate cleavage at the 3' splice site and ligation to Exon 2. Splicing yields two products: the ligated exons and the excised intron in the form of a lariat. The fate of the phosphate moieties at the splice sites is indicated through highlighted p's. Y - pyrimidine
(Adapted from Moore *et al.* (29))

Figure 2



example, U2 snRNP auxiliary factor (U2AF), which is required for the stable binding of U2 snRNP at the branch point in pre-mRNAs, is a heterodimer consisting of U2AF³⁵ (which contains an RS domain) and U2AF⁶⁵ (which contains three RBDs and an RS domain). U2AF binds to the polypyrimidine tract of pre-mRNAs through the RBDs of U2AF⁶⁵ which then facilitates binding of U2 snRNP to the nearby branch point sequence (30,31).

SR Proteins

RS domains have been found in the U1 snRNP-specific protein U1-70k and in the *Drosophila melanogaster* splicing regulators *transformer* (*tra*), *transformer-2* (*tra-2*) and *suppressor of white apricot* (*su(w^a)*). However, RS domains are typified in a series of proteins known as SR (serine/arginine-rich) proteins which are immunoreactive with the monoclonal antibody 104 (32,33). Eight SR proteins have been identified in humans and all are modular in structure consisting of an amino-terminal RBD and a carboxyl-terminal RS domain (34,35). The main features differentiating the individual SR proteins are the number of RBDs (one or two) and the length of the RS domain. Biochemical evidence indicates that the RS domain can mediate protein-protein interactions (12,36), influence RNA binding (31,37) or act as a signal directing localisation of the SR protein to nuclear substructures (38,39).

The proto-typical SR protein, ASF/SF2 (also known as SRp30a), was discovered by two separate groups using two different assays. SF2 (splicing factor 2) was purified, and subsequently cloned, as a factor essential for splicing *in vitro* (40,41). ASF (alternative

splicing factor) was identified as an activity that could influence the selection of alternative 5' splice sites in a viral pre-mRNA (42). It has subsequently been shown that all of the SR proteins can activate constitutive splicing *in vitro* (33) and can modulate alternative splicing (41,43,44). Complete redundancy amongst the SR proteins is unlikely and more recent work has started to shed some light on the differences in activities of the individual proteins (45,46).

The presence of RBDs in SR proteins suggests that sequence-specific RNA binding is probably important for their function. Experiments performed by Zuo and Manley (1994) have shown that ASF/SF2 can recognize 5' splice sites and that mutations within the splice site reduce or abolish binding (47). Tacke and Manley (1995) have gone on to show, using iterative selection from a pool of random RNA sequences, that ASF/SF2 and the SR protein SC35 (48) bind purine-rich RNA sequences 8-10 bases in length (49). ASF/SF2 and SC35 proteins with their RS domains removed displayed the same RNA binding specificity as the full-length proteins suggesting that the RS domain is not involved in determining this specificity. The selected sequences were distinct for the two polypeptides but interestingly both were found in some 5' splice sites.

Both ASF/SF2 and SC35 have been shown to regulate alternative 5' splice site choice in a concentration dependent manner (41,43,44). At high concentrations of the SR protein the 5' splice site proximal to the 3' splice site is chosen whereas with decreasing concentration of SR protein more distal 5' splice sites are used. This alternative splice site choice is antagonised by another spliceosomal-associated protein, hnRNP A1 (heterogenous nuclear RNP A1) (44,50).

Another of the ASF/SF2 selected RNA-binding sequences is found in a number of purine-rich splicing enhancers (49). Splicing enhancers are most often found within exons and facilitate splicing of the upstream intron (51-54). This is believed to occur through the strengthening of weak 5' or 3' splice sites (weak being defined as a sequence not conforming to the consensus) through the recruitment of other spliceosomal factors. An additional feature of exons containing splicing enhancers is that they are often regulated through alternative splicing (52-54).

SR proteins and purine-rich splicing enhancers can also function negatively to repress splicing. Kanopka *et al.* have demonstrated that a purine-rich sequence located next to a regulated 3' splice site in an adenovirus pre-mRNA represses splicing to that site in the presence of excess SR proteins (55). Thus, the sequence is acting as a splicing repressor, with binding of SR proteins to this site likely preventing the recruitment of U2 snRNP to the branch point sequence through steric hindrance. Interestingly, when placed in the downstream exon, this purine-rich sequence acts as a splicing enhancer (55), thus, SR proteins and the sequences they bind can influence splice site choices either positively or negatively depending upon their location.

As previously mentioned RS domains can facilitate protein-protein interactions and evidence suggests that this is required for RS domain-containing proteins to exert their effects. Interactions have been demonstrated between the RS domains of individual SR proteins, as well as between SR proteins and U2AF and U1-70k (12,36).

Figure 3A indicates how RS domain-containing proteins might enhance splicing through the stimulation of spliceosome assembly. ASF/SF2 can bind the U1 snRNP-

specific protein U1-70k, thus, binding of ASF/SF2 to the 5' splice site could enhance the subsequent binding of U1 snRNP to this site through the protein-protein binding of the ASF/SF2 and U1-70k RS domains (12,47). Furthermore, SC35 can mediate specific interactions between U1 and U2 snRNPs at the 3' splice site by binding U1-70k and U2AF (which stimulates U2 snRNP binding) (36,56). Functionally, this may bring the 5' splice site, branch point, and 3' splice site sequences in proximity to each other, creating a structure necessary for cleavage and ligation reactions to occur.

Similarly, SR protein binding to exon splicing enhancers could stimulate recognition of a weak 3' splice site through RS domain mediated binding of U2AF bound to the polypyrimidine tract (Fig. 3B). SR protein/exon enhancer-regulated recognition of 3' splice sites is a simple model for an alternative splicing mechanism. Accumulating evidence suggests that SR proteins function through their combined abilities of RNA binding and protein-protein complex formation mediated through RS domains.

Spliceosome regulation through phosphorylation

Although a considerable amount is known about the factors involved in the splicing reaction little is known about the mechanisms which exist to regulate their activity. A dynamic post-translational modification potentially involved in splicing factor regulation is protein phosphorylation. Mermoud *et al.* have used protein phosphatase inhibitors to demonstrate that dephosphorylation is required for both catalytic steps of the splicing reaction to occur (57) and that the addition of an excess of protein phosphatase inhibits spliceosome assembly (58). Further evidence for the role of phosphorylation in splicing

Figure 3

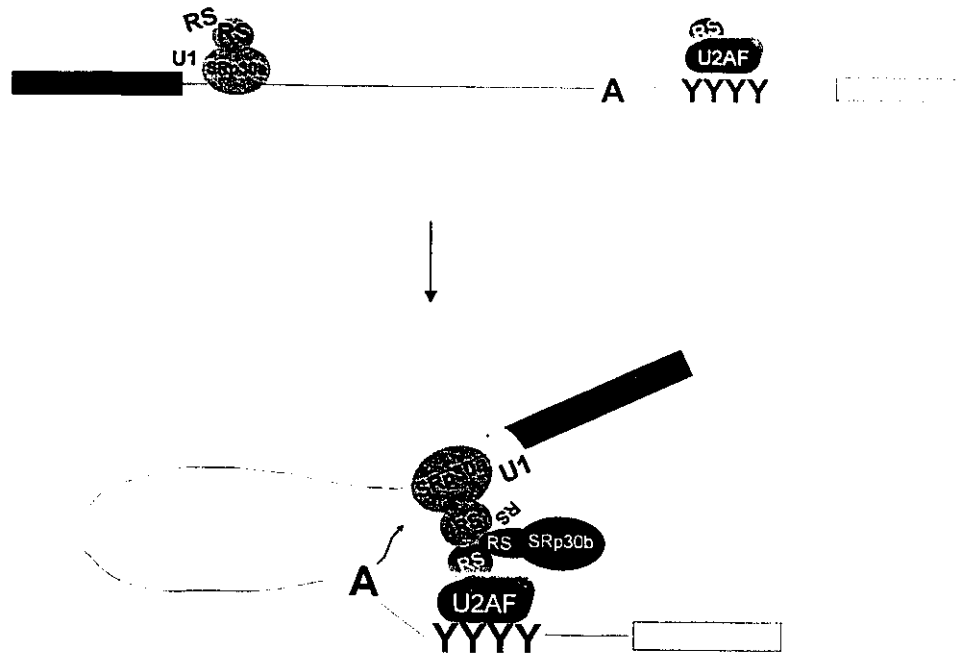
SR protein-mediated interactions in spliceosome assembly.

Protein-protein interactions between the RS domains (RS) of various splicing factors may aid spliceosome assembly. **(A)** The binding of the U1 snRNP (U1) to the 5' splice site is promoted through SRp30a (ASF/SF2) binding to the same 5' splice site. Binding of U2AF to the polypyrimidine tract stimulates binding of U2 snRNP to the branch point and promote interaction with SRp30a bound at the 5' splice site. The SRp30a/U2AF interaction may require SRp30b (SC35). These RS domain-mediated interactions may facilitate the required interaction of the branch point adenosine and the 5' splice site through structural changes in the pre-mRNA. **(B)** SR proteins bound to splicing enhancers may stimulate recognition of weak 3' splice sites by enhancing binding of U2AF to the polypyrimidine tract. The pre-mRNA is represented as in Figure 2.

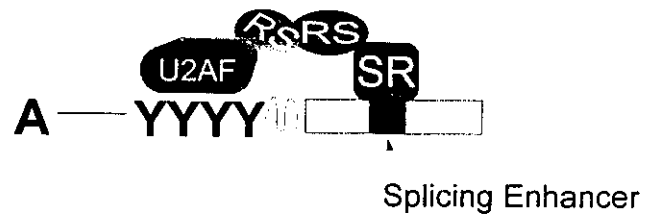
(Adapted from Valcárcel and Green (34))

Figure 3

A



B



is that thiophosphorylation (to prevent dephosphorylation) of the U1-70k protein inhibits splicing but does not prevent spliceosome assembly (59). RS domain containing polypeptides are potential targets for regulation through phosphorylation.

SR proteins constitute a family of phosphoproteins whose activity is likely regulated by dynamic phosphorylation (58,60). Kinases which can phosphorylate these epitopes *in vitro* have been identified (43,61-64). Phosphorylation of proteins within their RS domain has the potential to directly regulate protein-protein interactions and directly or indirectly regulate the other distinct functions of these polypeptides (ie. RNA binding, subcellular localisation).

Splicing factor localisation

Accumulating evidence indicates that pre-mRNA splicing occurs on nascent transcripts at sites of transcription (65-67). Much of this data is based upon the co-localisation of unspliced and spliced RNAs at sites of transcription. Splicing factors have been found distributed throughout the nucleoplasm (68) as well as being concentrated in discrete sites within the nucleus (60,68,69). These subnuclear sites can be differentiated by electron microscopy into peri-chromatin fibrils, interchromatin granules (also called "speckles", based upon their speckled appearance following immunolocalisation) and coiled bodies (70-72). RNA with a rapid turnover has been localised to peri-chromatin fibrils suggesting that these are the sites of active pre-mRNA splicing. What then is the function of interchromatin granules, which represent the majority of speckles visualised as containing spliceosomal components (48,68)? It has also been proposed that these act

as storage sites for unassembled splicing factors or that they are sites for the assembly of spliceosomal components (70,71); two possibilities that are not necessarily mutually exclusive. Further experimentation is required to resolve these issues.

THESIS WORK

Kinases play an important role in the regulation of signal transduction pathways. Clk1 was cloned during a screen to identify novel kinases potentially involved in such pathways. The work described in this thesis was aimed at elucidating a role for this kinase in mammalian cells. To this end we have characterised the RNA gene products and initiated a characterization of the protein products.

Chapter 2 is devoted to characterising the RNAs derived from the *clk1* gene and initial biochemical analysis of the *clk1* gene products expressed in mammalian cells.

In Chapter 3 data is presented which indicates a potential role for the *clk1* gene products in pre-mRNA splicing. Chapter 4 continues in this theme with evidence implicating the *clk1* gene products in the autoregulation of *clk1* gene expression through pre-mRNA splicing.

To further characterise the Clk1 polypeptide we have attempted to identify regulatory regions contained within the non-catalytic domain of Clk1 (Chapter 5).

Finally, a preliminary biochemical characterisation of the human Clks, hClk2 and hClk3, is presented in Chapter 5, the results of which suggest that, there is functional similarity amongst these related kinases.

CHAPTER 2

Alternative Splicing of Clk1, a Nuclear Dual Specificity

Kinase

SUMMARY

The LAMMER subfamily of kinases has been conserved throughout evolution and its members are thought to play important roles in the regulation of cellular growth and differentiation programs. Clk1 is a murine LAMMER kinase which has been implicated in the control of PC12 cell differentiation. Multiple transcripts are derived from the *Clk1* gene and their relative abundance is developmentally regulated. Alternative splicing of the primary *Clk1* transcript generates mRNAs encoding full length catalytically active (Clk1), and truncated, kinase deficient, polypeptides. Both Clk1 and its truncated isoform, Clk1^T, are localised in the nucleus, and are capable of heterodimerizing. We also demonstrate that Clk1 functions as a dual specificity kinase in mammalian cells.

INTRODUCTION

The *Clk1* gene encodes a member of the recently discovered family of dual specificity kinases (18,19). To date, at least eighteen distinct genes encoding dual specificity kinases have been identified in the genomes of yeast and mammals (for reviews see Douville *et al.* (6) and Lindberg *et al.* (5)). These kinases, expressed either as bacterial products or isolated from mammalian cells, have the ability to autophosphorylate on serine, threonine and tyrosine residues. Two of the best studied of these enzymes are the yeast *wee1*⁺ kinase, a regulator of progression through the cell cycle (73), and MEK (MAPK/ERK kinase), believed to be a key molecule in mitogen stimulated signalling pathways (74). We have previously reported the cloning of the murine *Clk1* cDNA which when expressed in bacteria encodes a protein capable of phosphorylating serine, threonine

and tyrosine residues *in vitro* (18). The Clk1 kinase contains an amino acid motif, LAMMER, found in kinase subdomain X (22), a feature which is shared with at least eight other dual specificity kinases expressed in humans, mice, plants and insects (21-23). The LAMMER motif containing protein kinases appear to be conserved throughout evolution suggesting that these enzymes may play important roles in the control of cellular growth and differentiation. Indeed, the DOA kinase, a *Drosophila* gene product, is critical to the development of the fly embryo and affects eye differentiation in the adult (22,75). The AFC1 kinase gene of *Arabidopsis thaliana* can complement yeast signal transduction mutants via activation of the transcription factor STE12 (23). In mammalian cells, the physiological function of the LAMMER kinases is largely unknown although overexpression of the Clk1 kinase in PC12 cells appears to initiate their differentiation into neural derivatives possibly through activation of a protein kinase cascade (76). As has been observed with other members of the LAMMER kinase family, the *Clk1* gene appears to express several differentially processed transcripts. To gain a full understanding of the physiological role of the Clk1 kinase we felt it important to identify and biochemically characterize all its gene products. We report here the cloning and sequencing of three novel *Clk1* cDNAs. Two of these cDNAs are derived from incompletely processed transcripts which accumulate in the nucleus in a developmentally regulated fashion. The third cDNA encodes a truncated polypeptide (Clk1^T) which like the full length Clk1 kinase is found in the nucleus and contains domains which facilitate dimerization. Furthermore, we present evidence that the Clk1 kinase is able to phosphorylate serine, threonine and tyrosine residues when expressed in mammalian cells.

MATERIALS AND METHODS

Construction of cDNA library

An L1210 leukaemia cell line cDNA library was constructed in λ gt10 from size selected poly(A)⁺ RNA. Poly(A)⁺ RNA was fractionated through a 2-10% sucrose gradient in 85% formamide, 10mM Tris·HCl, pH 7.5, 1mM EDTA and 0.2% SDS. Gradients were run at 200 000xg for 20h at 20°C. Following centrifugation, 0.4ml fractions were collected and aliquots analyzed by Northern blotting using a *Clk1* cDNA probe. Fractions containing RNA larger than 1.8-kb were pooled and precipitated with ethanol. cDNA was generated with 5 μ g size-selected poly(A)⁺ RNA using the SuperScript cDNA synthesis system (Gibco BRL) and was subsequently size selected for products larger than 1kb by gel filtration (Sephacose 4B, Pharmacia).

Isolation of cDNA clones

The λ gt10 library was screened using a full-length *Clk1* cDNA probe. From a screen of 200 000 plaques 18 positives were identified. Six were chosen at random for further analysis. To determine the size of the inserts PCR was employed to amplify cDNA using λ gt10 and *Clk1* specific primers. Briefly, pairs of λ gt10/*Clk1* primers were used to amplify DNA from the primary plaques of the library. Primer pairs were λ gt10 forward (5'-agcaagttcagcctggttaag^{3'}) and *Clk1* 619F (5'-tccaccacttttccgaaagca^{3'}) or λ gt10 reverse (5'-cttatgagtatttctccagggtta^{3'}) and *Clk1* 619F. 50 μ l PCR reactions were performed with 50 μ mol each primer, 0.2mM each dNTP, 1.25U Taq DNA polymerase in 50mM KCl, 10mM Tris·HCl, pH 8.4, 2.5mM MgCl₂ and 0.2mg/ml gelatin. PCR conditions were

94°C, 40s; 50°C, 60s; 72°C, 2min for 30 cycles. Amplified products were resolved by agarose gel electrophoresis and ethidium bromide staining. Two clones yielding inserts of approximately 2.6-kb were plaque purified by standard procedures (77).

Sequencing of cDNA clones

The cDNA inserts were subcloned into the *KpnI* site of the plasmid pGEM-4 (Promega). Double-stranded DNA sequencing was carried out using the dideoxy chain termination method (78). Full length sequence of the cDNA insert was achieved using oligonucleotide primers.

Northern hybridization analysis

Total RNA was prepared from cells or tissues as described by Chirgwin *et al.* (79). Poly(A)⁺ RNA was selected by passage of total RNA through oligo(dT)-cellulose columns as described by Jacobson (80). To isolate nuclear and cytoplasmic RNA NIH 3T3 cells (3-5x10⁷) were resuspended in 1ml lysis buffer (10mM Tris·HCl, pH 8.0, 140mM KCl, 5mM MgCl₂, 1mM dithiothreitol, 0.5% NP40) and lysed in a Dounce homogenizer. Nuclei were removed from the cytoplasmic fraction by centrifugation for 5min at 1200xg. Total nuclear and total cytoplasmic RNA was isolated from these fractions as outlined above. Aliquots of total RNA (20µg) were electrophoresed through 1% agarose gels containing 19% formaldehyde, 40mM morpholinepropanesulfonic acid (MOPS), 10mM sodium acetate, 1mM EDTA, transferred to Hybond N membrane (Amersham) and UV crosslinked as described by the manufacturer. Hybridization was carried out in 50% formamide, 6XSSC

(0.9M NaCl, 0.09M sodium citrate), 5X Denhardt's solution (1% [w/v] Ficoll, 1% [w/v] polyvinylpyrrolidone, 1% [w/v] bovine serum albumin), 0.5% SDS, 150 μ g/ml sheared herring sperm DNA, with a random-primed [³²P]-labelled *Clk1* cDNA probe at 42 °C. Blots were washed in 0.2XSSC, 0.1% SDS at 65°C. The positions of 18S and 28S RNA were determined by ethidium bromide staining of the agarose gel.

RT-PCR analysis and cloning of *Clk1*^T

Total RNA from P19 cells or mouse tissues was used to synthesize oligo (dT) primed first strand cDNA with SuperScript RNase H⁻ reverse transcriptase (Gibco BRL) as per the manufacturer's instructions. 1-2 μ l of a 20 μ l cDNA synthesis reaction was subsequently used in a PCR amplification. 50 μ l reactions were performed containing 20 μ mol primer *Clk1* 619F, 20 μ mol primer *Clk1* 93R (5'cgatccatgagacattcaaagagaact3') and 1.5U Taq DNA polymerase for 30-40 cycles under the conditions described above. *Clk1*^T cDNA was generated by RT-PCR using primer *Clk1* 7R (5'catcgtcgtaatcgtttgc3') and primer *Clk1* 1539F (5'cgtatgccttttaagtgg3'). Amplified products were resolved by electrophoresis through a 1.5% agarose gel and ethidium bromide staining.

Sequencing of RT-PCR products

Amplified PCR products were gel purified and cloned directly into the pCRII plasmid (Invitrogen) (7R/1539F product) or digested with *Hind*III and *Sac*I restriction endonucleases (93R/619F product) and subcloned into the plasmid pTZ19R (Pharmacia) following conventional procedures (77). Sequencing was performed as described above.

Identification of Additional Introns

Additional introns were identified within the *Clk1* gene by sequencing of genomic fragments isolated from a D3 embryonic stem cell genomic library (gift of Dr. D.A.Gray, University of Ottawa) and subcloned into pGEM-4 (Promega).

Construction of cDNA Expression Vectors and Tagged Proteins

pECE/*Clk1* was generated by insertion of *Clk1* cDNA into the mammalian expression vector pECE (81). The 1.7-kb *Clk1* fragment was isolated from pTZ19R/*Clk1* by PCR amplification using the following oligonucleotides: 5'cggtccatgagacattcaaagagaact3' (including a *Bam*HI site at the 5' end) and reverse sequencing primer (5'agcggataacaatttcacacagga3'). This fragment was subsequently digested with *Eco*RI and subcloned into pECE previously digested with *Sma*I and *Eco*RI.

The Clk1 polypeptide was tagged at the amino terminus with a six-repeat human Myc epitope by subcloning the Myc epitope isolated from Bluescript KS⁺ MTG (60) (*Sal*II/*Pst*I blunted fragment) into the pECE/*Clk1* plasmid (*Kpn*I blunted). The Myc epitope in pECE/M-*Clk1* was confirmed to be in frame by sequencing.

pECE/M-*Clk1*^{K190R} and pECE/M-*Clk1*^T were generated by PCR directed mutagenesis using the pECE/M-*Clk1* plasmid as template and the following oligonucleotides: 5'gcatagttaaaaatgtggatag3' and 5'gcactgctacacgtctac3' for pECE/M-*Clk1*^{K190R} or 5'cccgtgtgaatggtgctg3' and 5'atgaaattgttgatactttaggt3' for pECE/M-*Clk1*^T. 50µl reactions were performed containing 20ng plasmid, 25pmol each oligonucleotide, 0.2mM each dNTP, 2U Vent(exo⁻) DNA Polymerase (New England BioLabs) in buffer supplied by the

manufacturer. Amplification conditions were 95°C, 1min; 56°C, 1min; 72°C, 5min for twenty cycles. The reaction product was phosphorylated with polynucleotide kinase and recircularized with T4 DNA ligase.

An amino-terminal deletion mutant, pECE/M-*Clk1*^{ΔXhoII}, lacking the first sixty amino acids of Clk1, was created by substitution of the *Bam*HI/*Bam*HI *Clk1* fragment from pECE/M-*Clk1* with the *Xho*II/*Bam*HI *Clk1* fragment generated by subsequent digestion of the *Bam*HI/*Bam*HI fragment with *Xho*II. All mutations were confirmed by sequencing.

The Myc epitope in these constructs is immunoreactive with the monoclonal antibody (mAb) 9E10 in immunoprecipitation/immunoblot analysis (82).

Clk1 was also subcloned into the bacterial expression vector pGEX-3X (Pharmacia) to generate the fusion protein GST-Clk1. The 1.7-kb *Clk1* fragment was isolated from pTZ19R/*Clk1* by PCR amplification as described above. This blunt/*Eco*RI fragment was subcloned into pGEX-3X (*Bam*HI, blunted, followed by *Eco*RI digestion). *Clk1* was confirmed to be in frame with GST by sequencing.

COS-1 cell culture and transfection

COS-1 cells were maintained in α -minimal essential medium (MEM) supplemented with 10% calf serum. For transient transfection COS-1 cells were trypsinized, counted and resuspended in serum-free α -MEM at $2-3 \times 10^6$ cells in 0.5ml. 20 μ g of the appropriate plasmid DNA was added and the cells were chilled on ice for 10 minutes. Cells were electroporated using a Gene Pulser (Bio-Rad) at 220V and 960 μ F following which the cells were left at room temperature for 15 minutes. Transfectants were then plated in α MEM

supplemented with 10% calf serum in the absence or presence of 50 μ M sodium vanadate. Cells were harvested 24 hours later.

Immunoprecipitation/Immunoblot analysis

Transfected COS-1 cells were lysed in immunoprecipitation (IP) buffer (20mM Tris·HCl, pH 7.5, 150mM NaCl, 2mM EDTA, 1% Triton X-100, 2mM NaF, 2mM sodium pyrophosphate, 500 μ M sodium vanadate, 200 μ g/ml phenylmethylsulfonyl fluoride, 2 μ g/ml aprotinin, 5 μ g/ml leupeptin) for 30 minutes on ice. Lysates were cleared by centrifugation and the supernatants were immunoprecipitated with either mAb 9E10 or anti-phosphotyrosine mAb 4G10 (Upstate Biotechnology Incorporated). Immunoprecipitates were assayed for kinase activity in 20mM HEPES, pH 7.1, 10mM MgCl₂, 2mM MnCl₂ containing 10 μ Ci [γ ³²P]-ATP for 30 minutes at room temperature. Reactions were arrested by the addition of Laemmli's sample buffer followed by boiling. Samples were resolved by 10% SDS-polyacrylamide gel electrophoresis (PAGE), transferred to nitrocellulose membrane and exposed to Kodak XAR-5 X-ray film.

For immunoblots membranes were blocked in 5% Blotto in TBST (150mM NaCl, 10mM Tris·HCl, pH 7.5, 0.05% Tween-20). Anti-Myc immunoblotting was performed using ammonium sulfate precipitated 9E10 hybridoma culture supernatant at 1:50-1:100. Immunoblots were visualized with horseradish peroxidase-conjugated goat anti-mouse antibody followed by enhanced chemiluminescence (Amersham) and exposure to Kodak XAR-5 X-ray film.

Phosphoamino acid analysis was performed as previously described (4).

Immunofluorescence of M-Clk1 in COS-1 cells

Following transfection of pECE constructs into COS-1 cells approximately 7.5×10^4 cells were plated onto gelatin coated coverslips. 22-24h post-transfection cells were washed with PBS and fixed in -20°C methanol for 10min at -20°C . Following rehydration of cells in PBS, anti-Myc mAb (1:50 in PBS containing 0.3% Triton X-100) was added followed by fluorescein isothiocyanate (FITC)-conjugated anti-mouse secondary antibody (1:20, DAKO).

GST-Clk1 binding columns

GST-Clk1 binding reactions were carried out essentially as described (83). Briefly, bacterially expressed GST or GST-Clk1 was coupled to Glutathione-Sepharose 4B beads (Pharmacia) in IP buffer. Approximately $10\mu\text{g}$ of GST or GST-Clk1 coupled to beads was used in each binding experiment. COS-1 cell lysates containing M-Clk1, M-Clk1^T or M-Tik were mixed with GST or GST-Clk1 beads. Incubations were carried out for 2h at 4°C on a rotating platform. The beads were then washed with IP buffer and proteins eluted with sample buffer. Bound proteins were analyzed by SDS-PAGE, transferred to nitrocellulose and subsequently immunoblotted with anti-Myc mAb.

Nucleotide sequence accession numbers

The Genbank accession numbers are U11054 and U21209 for the *Clk1-5.6* and *Clk1^T* cDNAs respectively.

RESULTS

Sequence analysis of the developmentally induced *Clk1* transcripts reveals introns

Northern blot analysis revealed that the L1210 cell line expressed three distinct species of *Clk1* mRNA corresponding in size to the developmentally regulated mRNA species identified previously in P19 cells and other cell lines (data not shown, (18) and Fig. 2). In order to determine if these transcripts encode different isoforms of the Clk1 kinase we cloned the larger of these mRNA species. Using size selected mRNA isolated from L1210 cells we prepared cDNAs from mRNA species larger than 1.8-kb (see Materials and Methods). Our initial screen of 200 000 recombinants resulted in 18 positive clones one of which (2.6-kb in size) was sequenced fully. This 2.6-kb cDNA clone contained regions of sequence identity with the *Clk1* 1.8-kb embryonic transcript (hereafter referred to as *Clk1-1.8*), including the 5' end, interspersed with regions of non-identity. The intervening sequences did not, however, maintain the open reading frame as they contained stop codons in all reading frames (see Fig. 1). Further analysis revealed 5' and 3' splice sites at the ends of the intervening regions suggesting the presence of intronic sequences within an incompletely spliced transcript. The 2.6-kb cDNA did not contain the entire predicted *Clk1* open reading frame, but rather contained four exons (including an exon with the predicted initiating methionine), three complete introns (with 5' and 3' splice sites) and one partial intron (with a 5' splice site). To determine from which mRNA species the cDNA had been cloned, northern blots were probed with the cDNA regions encompassing individual introns (data not shown). The results of these experiments are depicted schematically in Figure 2. Probes derived from intron A (IA), or intron B (IB),

Figure 1

Nucleotide and predicted amino acid sequences of the *Clk1-5.6* partial cDNA and *Clk1^T* cDNA.

The amino acid sequence is derived from that predicted from the *Clk1-1.8* cDNA (18). *Panel A: Clk1-5.6* partial cDNA. Predicted intron/exon boundaries are indicated in bold face. The alternatively spliced exon is underlined and consensus amino acids found in kinase subdomains are highlighted and indicated by Roman numerals. Amino acids indicated by an asterisk are encoded by codons split between two adjacent exons. Exons (E) and introns (I) are labelled as follows: EA: nucleotides 1-456; IA: nucleotides 457-1326; EB nucleotides 1327-1417; IB: nucleotides 1418-1811; EC: nucleotides 1812-1995; IC: nucleotides 1996-2082; ED: nucleotides 2083-2249; ID: nucleotides 2250-2572. *Panel B: Clk1-1.7 (Clk1^T)* cDNA.

Figure 1

A

gcaagattctt ggcggctggc ttgtaagctt tgccttcgcc ttgcaagctt tctcttca 58
 M R H S K R T Y C P D W D 13
 7721772aaag atg aga cat tca aaq aga act tac tct cct zac tgg gat 108
 E R D W D Y G T W R S S S S H K 29
 7aa aga tgc tgg tac tat tga aca tgg aga agc agc agc agc cac aaa 156
 R K K R S H S S A R E Q K R C R 45
 aga aaq aga tca cat agc agc gcc cgt tag caa aaq cgc tgc agq 204
 Y D H S K T T D S Y Y L E S R S 61
 tac gat cac tcc aaa tgg dca gac agc tat tat ctg gaa atg aga tcc 257
 I N E K A Y H S R R Y V D E Y R 17
 ata aat gag aaa gct tat cat agt cga cgc tat gtt gat gaa tac agt 100
 N D Y H C Y E P G H P Y G E P S 93
 aat gac tac atg gcc tac agc cca agc cat ccc tat gga gaa cct gga 148
 S R Y Q M H S S K S S G R S G R 103
 agc aga tac cag atg cat agt agc aaq tcc tct gct gat gaa agt gga aga 196
 S S Y K S K H R S R H H T S Q H 125
 agc agt tac aaa agt aaa cac agc agc cgc cac cac aca tgc cag cac 444
 H S H G 129
 cat tca cac ggg gtaagagctc tttaaaagct tgcataagct ttatttcgcc 497
 atggagaaca ggaacacgct tctgtcttga gaaatttcca agataaaagt catcacaaact 557
 ctacctctga ggaatttct ttatgatttt tgcatttaca taagaatttc atgaaataag 617
 tgaattttag tctgttaagt ataggaaag tctgttattt gctatttcca gaaatactat 677
 tctgaataat tctcagcttt ttcccaagga attaggtatt ttgtgttag ggttagttat 737
 gactatata catgtatata ttcttctccc ataaattaca gattgaaagt agtcttttca 797
 tctatttatt tggaaattca tccggaatac tttaaaagt tatagtttca tatagctttt 857
 ttcttcttc ccaattttag gctatcattt tatttactat tttaataat attttcttca 917
 ttattttag ccaatttcca atttgcacac aggttactaa agtaataact acattcttca 977
 ttattttag atgaatttga gactatattt ttgattgaat gccacagata agtttttctt 1037
 ttattttag gaaagaaagc aaatttcttt atgttttctt gacttaccac ttgagctatc 1097
 ttattttag atgtttcttt tgcgtttt ttcttttctt aaataataat atgtttcttt 1157
 ttattttag ggaatttga ttgaatttgc tcttcccaac cgttttcttc ttggttttca 1217
 ttattttag tcccttctga atgaatttgc ttatttcttc ttgaatttct cgttttcttc 1277
 gatttcttgc atgatttctgc tcttcttcttc gccccttga cgttcttcag K S 131
 agt agt 1372
 H R R K R S R S V E D D E E G H 147
 cac cga agq aga aga tca agt agt gta agq gat tag agq ggt cag 1380
 L I C Q S G D V L S A R 159
 tga atg tga cag agt gga gac gta cta agt gca aga t gatagaat 1426
 atttttcaac acttatttct ttcttctga acataatttca tatctatata tagattttaa 1486
 ttcttctga ttcttctga ttcttcttct ttcttcttct ttcttcttct ttcttcttct 1546
 ttcttcttct ttcttcttct ttcttcttct ttcttcttct ttcttcttct ttcttcttct 1606
 ttcttcttct ttcttcttct ttcttcttct ttcttcttct ttcttcttct ttcttcttct 1666
 ttcttcttct ttcttcttct ttcttcttct ttcttcttct ttcttcttct ttcttcttct 1726
 ttcttcttct ttcttcttct ttcttcttct ttcttcttct ttcttcttct ttcttcttct 1786
 Y E I V D T L G E 168
 aataataatt atttcttctt tttag at gaa att gtt gat act tta gtt gaa 1837
 G A P G K V V E C I D H K V G G 184
 ggt gct ttc gga aaa gtt gtt gaa tgc atc gat cat aaa gtt gga ggt 1885
 R R V A V K I V K N V D R Y C E 200
 aga gct gta gca gta aaa ata gtt aaa aat gtt gat aga tac tct gaa 1938
 A A Q S E I Q V L H G H L N T T D 216
 gct gct caa tgc gaa ata caa gtt ttc gaa cac ttg aat aca aca gac 1981
 P H S T F 221
 ccc cat agt act tt gtaagttcac attaaactt gaaatagga gttttgaatt 2036
 R C V 224
 gtaagcttca ttatttgaatt gtttttcttt tttaagctt ttatttgc ggt tat gtc 2092
 Q M L E W F E H R G H I C I V F 240
 tag atg ttc gaa ttc gaa cat cga ggt cag att tgc att gtc ttt 2140
 E L L G L S T Y D P I K E N S F 256
 gaa ctt ctc ggg ctt agt act tat gat ttc att aag gaa aac agt ttt 2188
 L P P R M D H I R K H A Y Q I C 272
 ctc ctc ttc cga atg gat cat atc agc agc agc gca tat caa ata tgc 2236
 K S V N 276
 aaa ctt gta aac t gtaagtttctca ttatttcttc aaaaatttga ttgaacta 2291
 atgttgaact aatttcttc tttaaaaaa aaatttcttc ggaacttctg ggaacttctg 2351
 tgaatgacag aaatttcttc ttatttcttc ttcccttctg tacagaagac tagaacttga 2411
 gacttctctt ggaattttaa agcttctctt gaaatctat aaacttctaa ttggttctt 2471
 gacttctctt aatttctctg gcttctctt agtcttctaa agcttctctg gacttctctg 2531
 gacttctctt ccccttctt ttcaagctcag agcttctctg a 2572

B

atcttcttga ttcttctga acttctctgc ggccttctgc taacttcttc 50
 ttcttcttgc taacttctgc ttcaagcttgc gaaag atg aga cat tca 4
 M R H S 97
 K R T Y C P D W D E R D W D 18
 agt aga act tac tct cct gac tgg gat gaa aga gac tgg gat 139
 Y G T W R S S S S S H K R K K 32
 tat gga aca tgg aga agc agc agc agt cac aaa aga aaq agq 181
 R S H S S A R E Q K R C R Y 46
 aga tca cat agc agc gcc cgt tag caa aaq cgc tgc agq tac 223
 D H S K T T D S Y Y L E S R 60
 gat cac tcc aaa acg aca gac agc tat tat ctc gaa agc aga 265
 S I N E K A Y H S R R Y V D 74
 tcc ata aat gag aaa gct tat cat agt cga cgc tat gtt gat 307
 E Y R N D Y H C Y E P G H P 88
 gaa tac agc aat gac tac atg ggc tac gac cga ggg cat ccc 349
 Y G E P G S R Y Q M H S S K 102
 tat gga gaa cct gga agc aga tac cag atg cat agt agc aaq 391
 S S G R S C R S S Y K S K H 116
 tcc tct ggt agc agt gga aga agc agt tac aaa agt aca cac 433
 R S R H H T S Q H H S H G M 130
 agt agt cgc cac cac aca tgc cag cac cat tca cac tgg agc 475
 K L L I L 135
 aaa ttg ttg ata ctt tag gtaagcttgc ttcttcttca ggttcttga 523
 gacttctga taagcttga ggttctga tagacttga aatagttaa 573
 aatgttga gacttctga agcttctga ttcttctga aatgttga 623
 aacttctga taagcttga ttcttctga ttcttctga taagcttga 673
 gacttctga gacttctga ttcttctga ttcttctga taagcttga 723
 aacttctga taagcttga ttcttctga ttcttctga taagcttga 773
 aacttctga taagcttga ttcttctga ttcttctga taagcttga 823
 aacttctga taagcttga ttcttctga ttcttctga taagcttga 873
 aacttctga taagcttga ttcttctga ttcttctga taagcttga 923
 aacttctga taagcttga ttcttctga ttcttctga taagcttga 973
 aacttctga taagcttga ttcttctga ttcttctga taagcttga 1023
 aacttctga taagcttga ttcttctga ttcttctga taagcttga 1073
 aacttctga taagcttga ttcttctga ttcttctga taagcttga 1123
 aacttctga taagcttga ttcttctga ttcttctga taagcttga 1173
 aacttctga taagcttga ttcttctga ttcttctga taagcttga 1223
 aacttctga taagcttga ttcttctga ttcttctga taagcttga 1274

hybridized to the 5.6-kb and 3.2-kb transcripts (hereafter called *Clk1-5.6* and *Clk1-3.2*, respectively), while probes from intron C (IC) and D (ID) hybridized to *Clk1-5.6* but not *Clk1-3.2*. This data indicated that the 2.6-kb cDNA clone represented a partial 5' clone of the *Clk1-5.6* RNA species. Since probes corresponding to introns A and B hybridized to *Clk1-3.2* it is likely that this mRNA species is also an incompletely spliced transcript. Thus both the larger 5.6 and 3.2-kb transcripts contain intron sequences and arise as a result of incomplete and/or alternative splicing. The intron/exon boundaries of the 2.6-kb cDNA clone are indicated in Table 1.

***Clk1-3.2* and *Clk1-5.6*kb transcripts are nuclear**

Although both *Clk1-3.2* and *Clk1-5.6* RNAs contained intron sequences the possibility remained that these RNA species were capable of being translated into truncated Clk1 polypeptides by the cytoplasmic translation machinery. To test this idea, RNA from NIH 3T3 cells was separated into cytoplasmic and nuclear fractions and analyzed by northern blotting. All three *Clk1* transcripts were found in the nuclear fraction whereas only *Clk1-1.8* was found in the cytoplasmic fraction (Fig. 2). This is consistent with the larger transcripts (*Clk1-3.2* and *Clk1-5.6*) being nuclear, immature (incompletely spliced) transcripts. Fractionated Hela cell RNA and L cell RNA showed the same nuclear/cytoplasmic partitioning of *Clk1* transcripts (data not shown).

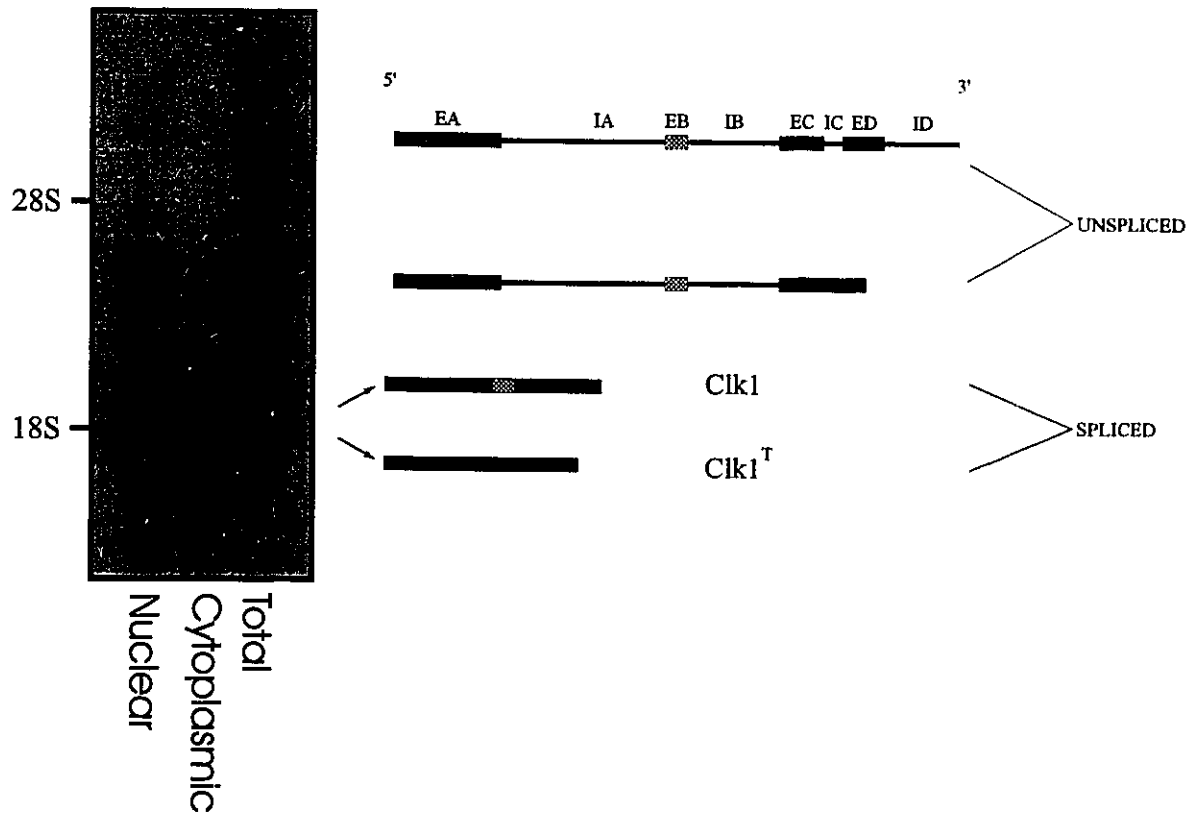
To determine whether *Clk1-5.6* represented the primary transcript from the *Clk1* gene, fragments of the *Clk1* locus were isolated from a D3 embryonic stem cell genomic λ DASHII library (see Materials and Methods). A fragment surrounding the predicted

Figure 2

Schematic representation of *Clk1* transcripts and analysis of their subcellular localization.

Total RNA from NIH 3T3 cells was separated into cytoplasmic and nuclear fractions and subjected to northern blot analysis. Each lane represents the hybridization of the *Clk1*- 1.8 cDNA to 20 μ g total RNA. The approximate sizes of the *Clk1* mRNAs are (from top to bottom) 5.6, 3.2, 1.8 and 1.7-kb. Positions of 18S and 28S RNA were determined by ethidium bromide staining of the agarose gel. Exons (EA-ED) are depicted by thick lines and introns (IA-ID) as thin lines. The alternatively spliced exon (EB) is hatched.

Figure 2



initiating methionine was cloned into the sequencing vector pGEM-4. Portions of this genomic fragment were sequenced using oligonucleotides residing in the *Clk1* cDNA. Two introns were identified in addition to those found in *Clk1-5.6*, demonstrating that *Clk1-5.6* is not the primary transcript but is an incompletely spliced RNA. The intron/exon boundaries and relative positions of these introns are indicated in Table 1.

***Clk1* is alternatively spliced**

Alignment of the amino acid sequence of Clk1 with that of the reported human homologue (hClk, (20)) using the Lipman-Pearson algorithm (84) revealed very high homology (83% identity over 482 amino acid overlap). However, there is a stretch of 30 amino acids contained in Clk1 that is absent within hClk (Fig. 3A). Sequence analysis indicated that these thirty amino acids are encoded within the second exon (exon B) of the 2.6-kb cDNA clone (see Fig. 1 & 3B). This suggested that exon B may be alternatively spliced. The *Clk1* cDNA originally isolated (18) would encode the isoform containing exon B, whereas the *hClk* cDNA (20) would encode the isoform lacking this exon. Consistent with this idea, the *Clk1-1.8* transcript often appears as a doublet upon northern blotting suggesting the presence of a fourth transcript (data not shown, (18) and Fig. 2).

To investigate this possibility we performed reverse transcriptase coupled with PCR amplification (RT-PCR) to amplify the entire coding sequence of *Clk1* (nucleotide position 7-1539, data not shown) from P19 cell derived cDNA. Two DNA species were amplified, the sizes of which were consistent with alternative slicing of exon B (data not shown). To

Table 1

Intron and exon boundaries found in the *Clk1-5.6* cDNA.

Coding sequence is capitalized while noncoding sequence is in small letters. Numbering of nucleotide positions is relative to that of *Clk1-1.8* (18). Intron/exon boundaries indicated by an asterisk were determined by sequencing genomic DNA from the *Clk1* locus. ND - not determined

Figure 3

Schematic representation of the alternative splicing of exon B.

Panel A: Comparison of the amino acid sequence surrounding exon B of Clk1 (mouse) with that of the human homologue (hClk). The sequence encoded by exon B in mouse which is absent in the human cDNA is denoted by the dashed line. Amino acid positions indicated are relative to those predicted for Clk1 (18) and hClk (20). Identical residues are identified by bars and conservative substitutions by dots. *Panel B:* The nucleotide sequence of the intron/exon boundaries surrounding the alternatively spliced exon is shown between the predicted splicing products. Splicing in of exon B would result in production of full-length Clk1. Splicing out of exon B results in a frameshift which introduces a premature stop codon (tag, indicated by a period). The resulting protein (Clk1^T) is truncated prior to the catalytic domain.

Table I

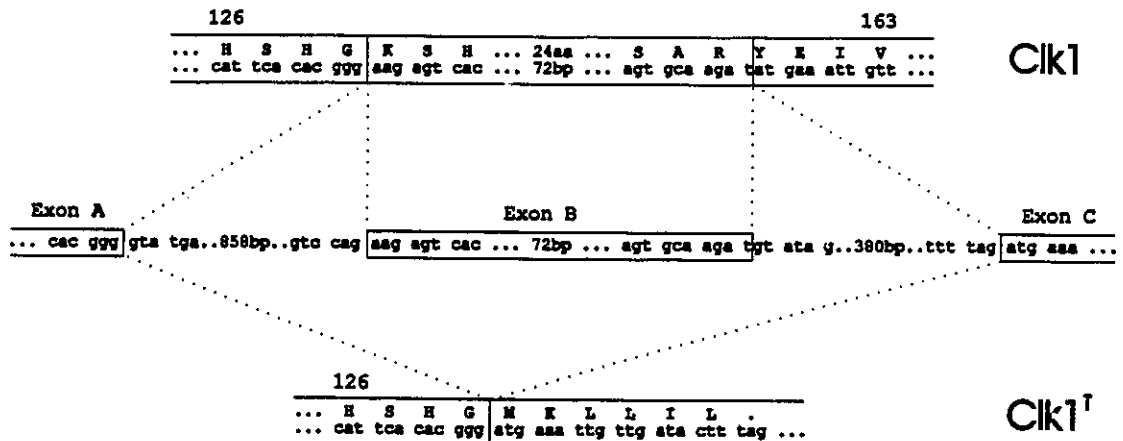
	<u>5' SPLICE SITE</u>	<u>INTRON SIZE</u>	<u>3' SPLICE SITE</u>
Intron A	⁴⁷⁰ GGG:gtatga	--- ~870bp ---	tgcgcgccgcttgacgtgtccag:AAG ⁴⁷³
Intron B	⁶⁶¹ GAT:gtatag	--- 394bp ---	taataattattcctcttttttag:ATG ⁶⁶⁶
Intron C	⁷⁴⁶ TTT:gtaagt	--- 86bp ---	ttgttttttaacgtgtttatag:CCG ⁷⁵⁰
Intron D	⁹¹² ACT:gtaagt	--- >320bp ---	not present
Intron 1*	ND	ND	ctggtaaaattgtctttcatag:ATG ⁸⁸
Intron 2*	²⁴⁴ CAG:gtaaga	--- ~600bp ---	aacaaactttttctcctctag:CTA ²⁴⁹
Consensus	$\begin{smallmatrix} C \\ A \end{smallmatrix}AG:gt\begin{smallmatrix} g \\ a \end{smallmatrix}agt$	--- >60bp ---	$\begin{smallmatrix} t & n & t & g \\ c & c & c & a \end{smallmatrix}a\begin{smallmatrix} t \\ c \end{smallmatrix} - (\begin{smallmatrix} t \\ c \end{smallmatrix})_n - cag:G$

Figure 3

A

	120		166
Mouse	HHTSQHHSHGKSHRRKRSRSVEDDEEGHLICQSGDVLSARYEIVDTL		
Human	HSTSHRRSHG-----DEIVDTL		
	121		137

B



verify the identity of the PCR products, the DNA was isolated, reamplified and subcloned into the pCRII vector. Sequencing revealed that the larger amplification product corresponded to full-length *Clk1* whereas the smaller lacked exon B (Fig. 1). This result was observed with cDNA derived from P19 cells (see P19 lane, Fig. 4) suggesting that the alternatively spliced transcript lacking exon B (*Clk1-1.7*) arises from a mRNA that co-migrates with the 1.8-kb transcript.

Expression of *Clk1-1.7* and *Clk1-1.8*, was assayed by RT-PCR amplification of the region encompassing exon B from cDNA of several mouse tissues (Fig. 4). All tissues tested, contained both *Clk1-1.7* and *Clk1-1.8* transcripts. A PCR product of approximately 1-kb in size was also reproducibly amplified from all tissues, the size of which is consistent with that of a splicing intermediate derived from *Clk1-5.6* (data not shown).

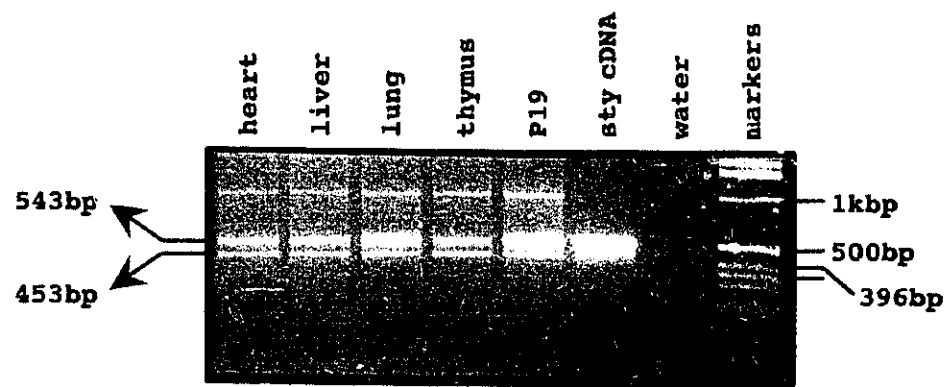
The *Clk1-1.7* mRNA, generated by alternative splicing of exon B, should encode a protein homologous to the hClk protein product. The protein product from the murine *Clk1-1.7* transcript is predicted to be truncated and catalytically inactive (we have termed this product Clk1^T for truncated Clk1) (see Fig. 3B). However, the protein product of *hClk*, which lacks exon B, is predicted to be catalytically active. This apparent contradiction suggested that either the sequence encompassing exon B is absent in the human genome and that in its absence the gene has evolved to maintain a catalytically active molecule or, alternatively, the *hClk* cDNA is the homologue of *Clk1-1.7* and should encode a truncated molecule. To investigate this further a human liver cDNA library (gift of Dr. D.A. Gray, University of Ottawa) was screened with a full-length *Clk1* cDNA probe. From an initial screening of 200,000 recombinants, thirty clones were identified.

Figure 4

RT-PCR analysis of *Clk1* transcripts demonstrates alternative splicing.

Oligonucleotide primers flanking exon B were used to amplify cDNA generated from adult mouse tissues or P19 EC cells as indicated above each lane or to amplify the *Clk1-1.8* cDNA. PCR products were resolved on a 1.5% agarose gel and stained with ethidium bromide. The predicted product sizes of 543-bp (*Clk1*) and 453-bp (*Clk1¹*) are indicated on the left and DNA markers sizes are shown on the right.

Figure 4



Of these, ten were picked at random and analyzed by PCR to determine if they contained the exon A-exon C region. One positive clone was analyzed further and the sequence of this boundary indicated that it was the human homologue of *Clk1-1.7* and would encode a truncated protein product (data not shown). The human cDNA, as published by Johnson and Smith (20), contains an additional G residue at this boundary which would maintain the open reading frame through the catalytic domain. The absence of this residue in our mouse and human cDNAs causes a frameshift to occur which introduces a stop codon (Fig. 3B).

Expression of Clk1 protein in mammalian cells

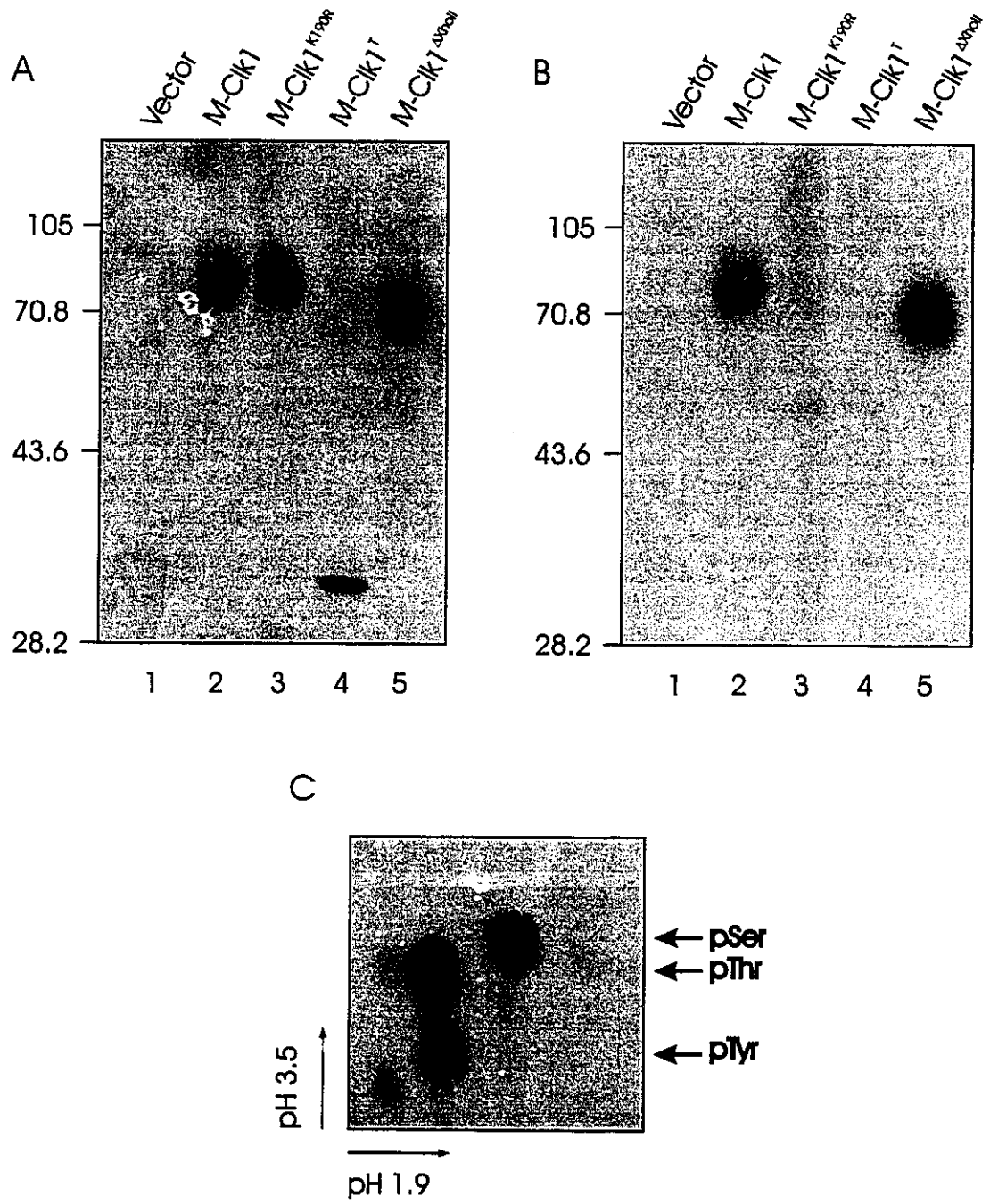
The *Clk1* cDNA encodes a 57-kDa polypeptide, whereas the *Clk1^T* cDNA encodes a 16.3-kDa truncated polypeptide lacking the catalytic domain. In order to detect these protein products we tagged the Clk1 and Clk1^T polypeptides with a six-repeat human Myc epitope and expressed these fusion proteins in COS-1 cells. The epitope tag added 89 amino acids to the amino-terminal end, increasing the molecular mass by approximately 10-kDa. COS-1 cells transfected with the pECE vector alone showed no proteins detectable by anti-Myc immunoblotting (Fig. 5A, lane 1). Transfection of pECE vectors encoding Myc-tagged Clk1 (M-Clk1) or Clk1^T (M-Clk1^T) cDNAs demonstrated anti-Myc immunoreactive proteins of approximately 75-kDa and 34-kDa respectively (Fig. 5A, lanes 2,4) corresponding to full-length and truncated Clk1 kinases.

Figure 5

Expression and catalytic activity of wild-type and mutant Clk1 proteins in COS-1 cells.

COS-1 cells were transfected with either pECE expression vector alone (*lane 1*) or vector encoding Myc epitope-tagged Clk1 (M-Clk1), catalytic mutant Clk1 (K190R), truncated Clk1 (T) or an amino-terminal deletion of Clk1 (Δ XhoII) (*lanes 2-5*, respectively). Myc tagged proteins were immunoprecipitated with the anti-Myc monoclonal antibody and subjected to an *in vitro* kinase assay. Proteins were subjected to SDS-PAGE and transfer to nitrocellulose. Expression levels were determined by immunoblotting with the anti-Myc antibody and enhanced chemiluminescence (*panel A*). Autoradiogram of *in vitro* catalytic activities of wild-type and mutant Clk1 proteins (*panel B*). Phosphoamino acid analysis by two dimensional thin-layer electrophoresis of wild-type M-Clk1 labelled *in vitro* (*panel C*). Positions of molecular mass markers (kDa) are indicated to the left of each panel. The positions of phosphoserine (pSer), phosphothreonine (pThr) and phosphotyrosine (pTyr) are also indicated.

Figure 5



Clk1 protein exhibits dual specificity kinase activity *in vitro* and *in vivo*

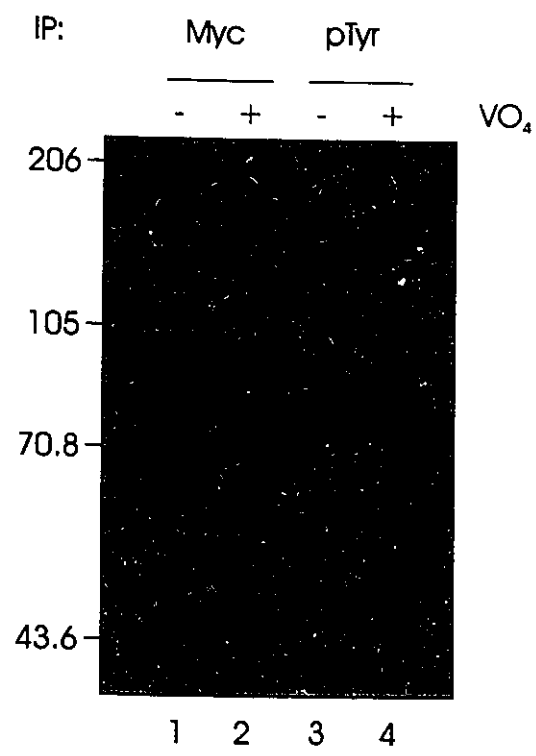
Clk1 expressed in bacteria has dual specificity kinase (DSK) activity (1). To determine if Clk1 expressed in mammalian cells has similar DSK activity we used an immune complex kinase assay. Cells transfected with the M-*Clk1* plasmid expressed a 75-kDa phosphoprotein as detected in an anti-Myc mAb immune complex kinase assay (Fig. 5B, lane 2). Phosphoamino acid analysis of the M-Clk1 phosphoprotein revealed substantial phosphoserine, phosphothreonine and phosphotyrosine (pTyr) (Fig. 5C). It remained possible that a co-precipitating kinase was responsible for phosphorylation of some or all of the sites on M-Clk1. To address this problem a catalytically inactive mutant of M-Clk1, M-Clk1^{K190R} was generated. Oligonucleotide-directed PCR mutagenesis was used to change the invariant lysine residue 190 in subdomain II of the catalytic domain to an arginine residue (K190R). While expression of this mutant could be readily detected by immunoblotting COS-1 cell extracts (Fig. 5A, lane 3), M-Clk1^{K190R} showed no kinase activity or evidence of phosphorylation (Fig. 5B, lane 3). This result demonstrates that Clk1 protein expressed in mammalian cells exhibits DSK activity. To test its *in vivo* phosphorylation state, M-Clk1 was transfected into COS-1 cells, immunoprecipitated with either anti-Myc mAb or anti-pTyr mAb and then subsequently immunoblotted with anti-Myc mAb (Fig. 6). Cells grown in the absence or presence of vanadate, a phosphotyrosine phosphatase inhibitor, showed similar levels of anti-Myc immunoprecipitable M-Clk1 (Fig. 6, lanes 1,2). In the absence of vanadate only a fraction of the M-Clk1 could be immunoprecipitated with anti-pTyr mAb (Fig. 6, lane 3). In the presence of vanadate, however, similar amounts of M-Clk1 could be immunoprecipitated with anti-pTyr mAb

Figure 6

Clk1 protein is phosphorylated on tyrosine *in vivo*.

COS-1 cells were transfected with the pECE vector encoding Myc epitope-tagged Clk1 (M-Clk1) and grown overnight in the absence (*lanes 1,3*) or presence (*lanes 2,4*) of 50 μ M sodium vanadate (VO₄). M-Clk1 was immunoprecipitated (IP) with either the anti-Myc monoclonal antibody (myc, *lanes 1,2*) or an antiphosphotyrosine monoclonal antibody (pTyr, *lanes 3,4*). Immunoprecipitated Myc-tagged proteins were visualized by immunoblot analysis and enhanced chemiluminescence with the anti-Myc antibody. Positions of molecular mass markers (kDa) are indicated on the left.

Figure 6



(Fig. 6, compare lanes 2 and 4). As expected the M-Clk1^{K190R} mutant did not react with the antibody directed against phosphotyrosine (data not shown). These results were corroborated by incubating transfected COS-1 cells with [³²P]-orthophosphate and performing phosphoamino acid analysis on labelled M-Clk1 protein. As with the *in vitro* kinase assay, all three hydroxyamino acids were phosphorylated when M-Clk1 was labelled *in vivo* (data not shown).

Nuclear localization of Clk1 protein

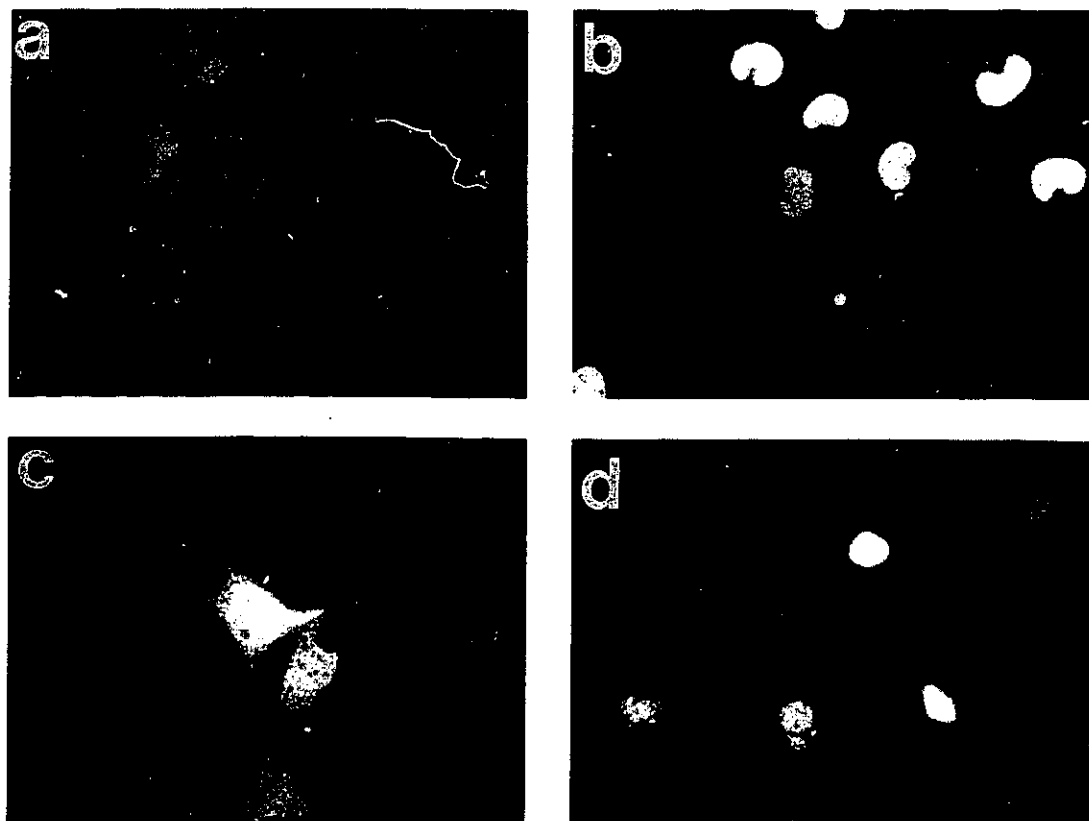
The subcellular localization of M-Clk1 in transfected COS-1 cells was determined by indirect immunofluorescence with the anti-Myc mAb. M-Clk1 expressed in COS-1 cells was nuclear with undetectable cytoplasmic staining (Fig. 7B). Deletion of the first sixty amino acids of Clk1 (M-Clk1^{ΔXhoII}), which contain the predicted nuclear localization signal (NLS), directed widespread cytoplasmic expression of this polypeptide (Fig. 7C). Nuclear staining was not eliminated, suggesting that there may be multiple NLSs in Clk1. M-Clk1^{ΔXhoII} expressed in COS-1 cells as assessed by immunoblotting, retained *in vitro* catalytic activity (Fig. 5A, lane 5, Fig. 5B, lane 5 respectively). The mutants, M-Clk1^T and M-Clk1^{K190R}, were also targetted to the nucleus (Fig. 7D and data not shown) demonstrating that catalytic activity is not required for nuclear compartmentalization of Clk1.

Figure 7

Subcellular localization of wild-type and mutant Clk1 proteins.

COS-1 cells were transfected with the pECE vector alone (*panel a*), or vector encoding Myc epitope-tagged Clk1 (*panel b*), an amino-terminal deletion of Clk1 (Clk1 ^{Δ X_{hoII}}, *panel c*) or truncated Clk1 (Clk1^T, *panel d*). Myc-tagged Clk1 proteins were visualized by indirect immunofluorescence with the anti-Myc monoclonal antibody followed by FITC-conjugated secondary antibody. Cells were photographed at a magnification of 40X with exposure times of 30s (*panels b-d*) and 60s (*panel a*).

Figure 7



Clk1 protein can dimerize

Since both Clk1 and its truncated counterpart, Clk1^T, are found in the nucleus we wished to determine if the two isoforms could physically interact with one another. M-Clk1 expressed in COS-1 cells was incubated *in vitro* with bacterially expressed glutathione-S-transferase (GST) or a GST-Clk1 fusion protein coupled to glutathione-Sepharose 4B beads. M-Clk1 from COS-1 cells did not bind to beads alone or to GST-coupled beads (Fig. 8, lanes 2,3). However, M-Clk1 was able to bind to immobilized GST-Clk1 (Fig. 8, lane 4). To rule out the possibility that the Myc epitope tag was mediating this interaction we tested the ability of another epitope tagged DSK, M-Tik (Abraham and Bell, unpublished), to bind to the GST-Clk1 column. As can be seen in Figure 8 (lanes 5-8) the M-Tik kinase did not bind to either GST or GST-Clk1 columns demonstrating that the Myc epitope was not responsible for the observed binding of M-Clk1 to GST-Clk1.

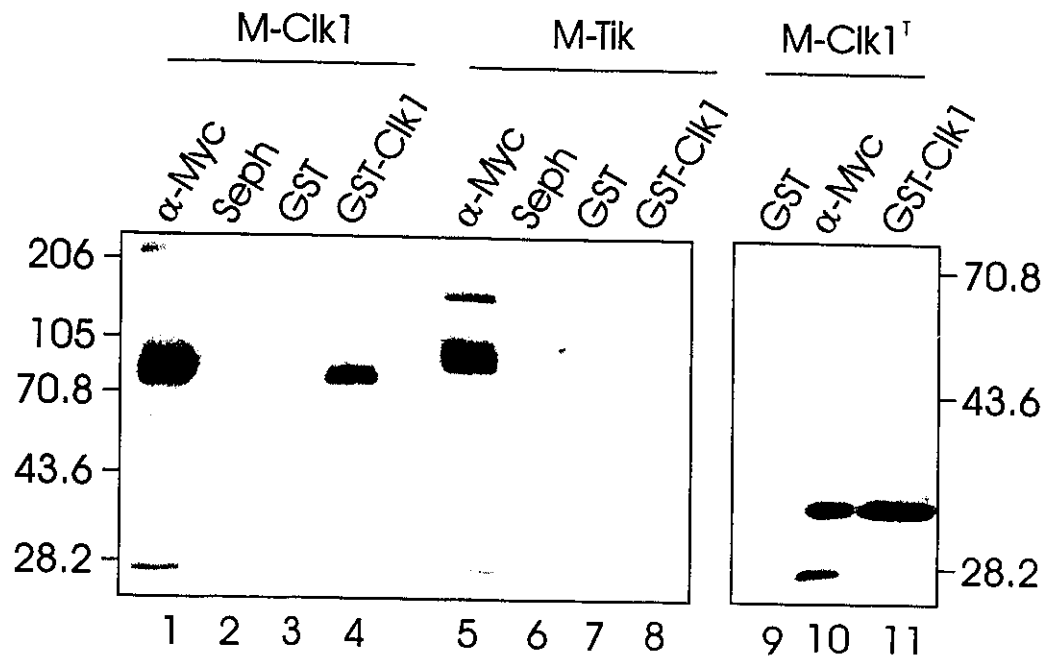
Truncated Clk1 (M-Clk1^T) also formed heterodimers with full-length Clk1 (Fig. 8, lanes 9-11) demonstrating that the Clk1 dimerization domain is contained in the amino-terminal portion of the molecule and is not dependent upon catalytic activity of the kinase. These results, however, do not differentiate between a direct association of Clk1 with itself or an association mediated through an intermediary molecule.

Figure 8

Clk1 can dimerize *in vitro*.

COS-1 cells were transfected with the pECE expression vector encoding Myc epitope-tagged Clk1 (M-Clk1, *lanes 1-4*), Tik (M-Tik, *lanes 5-8*) or truncated Clk1 (M-Clk1^T, *lanes 9-11*). Protein from equivalent fractions of cell lysates was subjected to immunoprecipitation with the anti-Myc monoclonal antibody (*lanes 1,5,10*) or affinity purification with glutathione-Sepharose 4B (Seph, *lanes 2,6*), Gst coupled to Seph (GST, *lanes 3,7,9*) or GST-Clk1 coupled to Seph (GST-Clk1, *lanes 4,8,11*). Purified Myc-tagged proteins were visualized by immunoblot analysis and enhanced chemiluminescence with the anti-Myc antibody. Positions of molecular mass markers (kDa) are indicated on the left and right.

Figure 8



DISCUSSION

Isoforms of several protein kinases have been predicted and identified based on cDNA cloning of multiple RNA transcripts (85-87). For instance, six transcripts, coding for at least two protein isoforms, are derived from the gene encoding the *trkB* neurogenic tyrosine kinase receptor (85). These *trkB* proteins, gp145^{*trkB*} and gp95^{*trkB*}, have identical extracellular and transmembrane domains, but only gp145^{*trkB*} contains the cytoplasmic kinase domain. Similarly, the fibroblast growth factor tyrosine kinase receptor-1 gene generates several transcripts which encode receptor variants (86). One of these, a receptor-like molecule that lacks the transmembrane and kinase domains, is secreted and catalytically inactive.

The expression of the *Clk1* gene is developmentally regulated at the level of transcript processing. Embryonic stem cells express two major transcripts (1.7 and 1.8-kb) whereas differentiated cells express two additional partially spliced mRNAs (3.2 and 5.6-kb). While the two larger mRNAs are sequestered in the nucleus and are unavailable for protein translation, the 1.7 and 1.8-kb mRNAs are capable of directing synthesis of a truncated and full length Clk1 protein respectively. In general, unspliced primary transcripts are relatively short lived with half lives on the order of minutes (88). However, several examples of developmentally regulated mRNAs exist which reflect developmental changes in expression of factors involved in the splicing process. In *Drosophila*, differentially expressed splicing factors control the expression of the sex lethal and transformer RNAs during sexual determination (89). In the developing pituitary, increased transcription of the proopiomelanocortin (POMC) gene results in accumulation of

immature transcripts possibly due to limiting amounts of specific splicing factors (90). Immature T cell antigen mRNA species appear to accumulate at specific stages of T cell development due to the presence of a cycloheximide sensitive splicing inhibitor protein (91). It is unknown at this time which, if any, of these mechanisms accounts for the accumulation of *Clk1* mRNA precursors. It is of interest to note, however, that two leukaemic cell lines, P388 (pre-B) and SP10 (myeloma), like embryonic stem cells, do not accumulate these larger transcripts (18). Thus the 1.7/1.8-kb embryonic transcripts may be favoured in rapidly proliferating cells such as embryonic or tumour cells.

Amino acid comparison of Clk1 with its human homologue, hClk, revealed a thirty amino acid insertion in Clk1. Isolation and sequencing of the *Clk1-5.6* partial cDNA clone revealed that this thirty amino acid segment was contained within a single exon. Alternative splicing of this exon would generate either the full-length product, Clk1 (containing the exon), or a truncated polypeptide (Clk1^T, lacking this exon) due to a frameshift which introduces a stop codon (see Figure 3B). Johnson and Smith (20) suggested that the *hClk* cDNA, which lacks the exon, coded for an active kinase. We have isolated a human *Clk* cDNA and sequenced it in this region. Our sequence indicated that the G residue at position 546 in the *hClk* cDNA is not present, and as a result an mRNA with this 91-bp exon absent would encode a truncated protein as predicted by the mouse sequence. One simple interpretation of this data is that the original human cDNA contained a sequencing error and that Clk1^T and Clk1 isoforms are encoded in both the human and mouse genomes. A similar conclusion has been reached by Hanes *et al.* (21) who have recently reported the sequence of two new human kinase cDNAs with significant sequence

homology to both *Clk1* and *hClk*. As we have shown here for *Clk1*, the primary transcripts of these two kinases (*hClk2* and *hClk3*) undergo alternative splicing to generate mRNAs predicted to encode full length and truncated, catalytically inactive kinases (21). This striking conservation of amino acid sequence and splicing pattern for three separate kinase genes suggests that they may represent a subfamily of enzymes with related physiological targets. It is of interest to note that the DOA and AFC1 LAMMER kinases have been implicated in the control of transcription. The nuclear localization of Clk1 is consistent with it having a role in the regulation of transcription and/or RNA processing. With respect to this latter idea, we and others have recently found that Clk1 and Clk1^T can interact with members of the SR family of splicing factors (62). We are testing the possibility that hClk2, hClk3 and their truncated derivatives also associate with splicing factors.

One commonly accepted paradigm of kinase regulation is that dimerization is required for kinase activation and subsequent signal transduction (92,93). Evidence for this idea includes the demonstration that catalytically inactive kinase mutants dominantly suppress wild type kinase activity (94). Clk1 and Clk1^T show identical patterns of expression, subcellular localization and can form heterodimers *in vitro*. We suggest that a dynamic interaction between Clk1 and Clk1^T may be involved in the regulation of the biological properties of Clk1 and related kinases.

CHAPTER 3

**The Clk1 Protein Kinase Regulates the Intranuclear
Distribution of SR Splicing Factors**

SUMMARY

Mammalian Clk1 is the prototype for a family of dual specificity kinases (termed LAMMER kinases) that have been conserved in evolution, but whose physiological substrates are unknown. In a yeast two-hybrid screen, the Clk1 kinase specifically interacted with RNA-binding proteins, particularly members of the serine/arginine (SR)-rich family of splicing factors. Clk1 itself has an SR-rich non-catalytic amino-terminal region which is important for its association with SR splicing factors. *In vitro*, Clk1 efficiently phosphorylated the SR family member ASF/SF2 on serine residues located within its SR-rich region (the RS domain). Tryptic phosphopeptide mapping demonstrated that the sites on ASF/SF2 phosphorylated *in vitro* overlap with those phosphorylated *in vivo*. Immunofluorescence studies showed that catalytically inactive forms of Clk1 co-localise with SR proteins in nuclear speckles. Over expression of the active Clk1 kinase caused a redistribution of SR proteins within the nucleus. These results suggest that the Clk1 kinase directly regulates the activity and compartmentalization of SR splicing factors.

INTRODUCTION

Clk1 is a mammalian protein kinase with the unusual property of phosphorylating both serine/threonine and tyrosine residues (reviewed in (5,6)). Clk1 was isolated in an anti-phosphotyrosine antibody screen designed to isolate cDNAs for enzymatically active tyrosine kinases (18,19). The protein is composed of a carboxyl-terminal kinase domain,

most closely related to serine/threonine-specific kinases, preceded by an amino-terminal non-catalytic region of 156 residues. When isolated from bacteria or mammalian cells, Clk1 has been shown to autophosphorylate on all three hydroxyamino acids (6,18,19).

Clk1 is the prototypic member of the LAMMER family of protein kinases, which has been conserved through evolution (22). This family includes: KNS1 from *Saccharomyces cerevisiae* (24), AFC1, AFC2, AFC3 from *Arabidopsis thaliana* (23), DOA from *Drosophila melanogaster* (22), human Clk1 (hClk)(20), hClk2 (21,95), and hClk3 (21) as well as mouse Clk1. In addition to a strong degree of homology across the kinase domain, these proteins share a highly conserved motif, "EHLAMMERILGPLP", in subdomain X of the kinase domain. This region is not conserved among kinases in general (1) and may be involved in dictating substrate specificity (22). The best characterised member of this family from a genetic perspective is *Drosophila* DOA. The *doa* gene was isolated as a dosage-sensitive modifier of the *apricot* allele at the *white eye* locus (75). Homozygous loss-of-function *doa* mutants die at the early larval stage indicating that this gene is critical for development (75). Stage-specific and tissue-specific *doa* transcripts have been identified in *Drosophila* indicating that expression of the LAMMER family may be regulated at the post-transcriptional level (22). Consistent with this notion, the expression of Clk1 is developmentally regulated at the level of RNA processing, as illustrated by the generation of two alternatively spliced transcripts encoding a catalytically active kinase (Clk1) and a truncated protein lacking the kinase domain (Clk1^T) (96). These two transcripts are observed in rapidly proliferating cells and undifferentiated stem cells (18,19,96). In contrast, in adult mouse tissues or upon stem cell

differentiation, larger transcripts are detected (18,19). The biological functions of mammalian Clk1 are unknown, although a possible role in PC12 differentiation has been suggested (76). No *in vivo* substrates of Clk1, or other LAMMER kinases, have been identified. To address this question, a yeast two hybrid screen was employed to isolate Clk1-binding proteins (97). In this assay Clk1 was shown to interact with five RNA-binding proteins: hnRNP G (98), RNP S1 (99), and three members of the SR (serine-arginine rich) protein family; ASF/SF2 (also known as SRp30a) (41,43), X16 (also known as SRp20) (100), and SRp55 (K. Colwill, T. Pawson, unpublished). The five cloned proteins contain an amino-terminal RNA binding domain and a carboxyl-terminal RS domain. Catalytic activity was not required for this binding as kinase-inactive Clk1^{K190R} bound to the RNA binding proteins with similar efficiency as wild-type. The kinase domain alone showed no interaction with these proteins. In contrast, the amino-terminal non-catalytic region could interact with RNP S1, ASF/SF2 and X16 but not with hnRNP G or SRp55. The isolated RS or RNA binding domains of X16 showed little interaction with Clk1 suggesting that more extensive contacts along both binding proteins are required for efficient association.

Clk1 can phosphorylate ASF/SF2 *in vitro* at sites within the RS domain that are also phosphorylated *in vivo*. Although Clk1 displays dual-specificity kinase activity ASF/SF2 was only phosphorylated on serine and threonine residues (62).

We show here that Clk1 regulates the nuclear distribution of SR splicing factors *in vivo*. We propose that SR splicing factors are physiologically relevant substrates for the Clk1 kinase.

MATERIALS AND METHODS

Immunofluorescence Assay

COS-1 cells were transfected with plasmids encoding M-Clk1 or M-Clk1^{K190R} (96), both of which contain the Myc epitope. Cells were fixed, and probed as described (96). For anti-Myc/anti-SR double staining by indirect immunofluorescence, fixed cells were incubated with mAb 9E10 (α -Myc, 25 μ l of culture supernatant) (82), followed by secondary fluorescein-conjugated goat-anti-mouse IgG (1:100, Jackson ImmunoResearch Laboratories) specific for mouse IgG and subsequently incubated with mAb 104 (α -SR, 25 μ l culture supernatant) (32) followed by secondary biotin-labelled α -mouse IgM (1:100, Amersham) and streptavidin coupled to Texas Red (1:100, Amersham). Anti-Myc/anti-Sm staining was performed using mAb 9E10 (25 μ l of culture supernatant) and a human anti-Sm antisera (1:300, gift of Dr. X.-D. Fu, University of California, San Diego) followed by rhodamine-conjugated goat-anti-mouse IgG (1:50, Tago, Inc.) and FITC-conjugated rabbit-anti-human IgG/IgA/IgM (1:500, gift of Dr. M. Paulin-Levasseur, University of Ottawa). Cells were viewed by confocal microscopy using an upright Leica Confocal Laser Scanning Microscope equipped with a 55 mW Krypton/Argon air-cooled laser and a 25X or 63X Plan Apo oil immersion lens.

Triton X-100 Extraction

Transfected cells were extracted on coverslips (for immunofluorescence) or in microcentrifuge tubes (for immunoblotting) with extraction buffer (20mM Tris-HCl, pH 7.5, 150mM NaCl, 2mM EDTA, 1% Triton X-100, 2mM NaF, 2mM sodium

pyrophosphate, 500 μ M sodium vanadate, 200 μ g/ml phenylmethylsulfonyl fluoride, 2 μ g/ml aprotinin, 5 μ g/ml leupeptin) for 30 minutes at 4°C. For immunofluorescence coverslips were washed with PBS, and fixed and processed as above. For immunoblotting lysates were cleared by centrifugation. Supernatants (Triton extracted fraction) and pellets (SDS extracted fraction) were boiled in Laemmli's SDS buffer and samples resolved by 10% SDS-PAGE and transferred to nitrocellulose membrane. Membranes were probed with anti-Myc mAb as described previously (Chap. 2).

RESULTS

Subnuclear localization of Clk1 and SR proteins *in vivo*

If Clk1 interacts with SR proteins in mammalian cells then these polypeptides might be anticipated to have a shared subcellular localization. During interphase, pre-mRNA splicing factors including the SR proteins, are found localised in nuclear "speckles" (48,69,101). To investigate this point, COS-1 cells were transiently transfected with plasmids encoding either full length Clk1 (M-Clk1) or kinase inactive Clk1 (M-Clk1^{K190R}) (96). These proteins were fused to the Myc epitope to allow detection by mAb 9E10 (82). Cells were subsequently probed with mAb 104 which recognizes SR proteins (32) and mAb 9E10 to detect transfected Clk1 proteins and analysed by indirect immunofluorescence followed by confocal microscopy imaging. Representative fields are shown in figure 1. Clk1^{K190R} was found predominantly in nuclear speckles where they co-localised with endogenous SR proteins (Fig. 1B,C). In contrast, cells expressing catalytically active Clk1 displayed a diffuse nucleoplasmic staining pattern for both the kinase and SR proteins (Fig. 1D). Untransfected cells adjacent to those expressing Clk1 displayed a normal distribution of SR proteins in nuclear speckles (Fig. 1A). Identical results were obtained using an antibody specific for a single SR protein family member, SC35 (48).

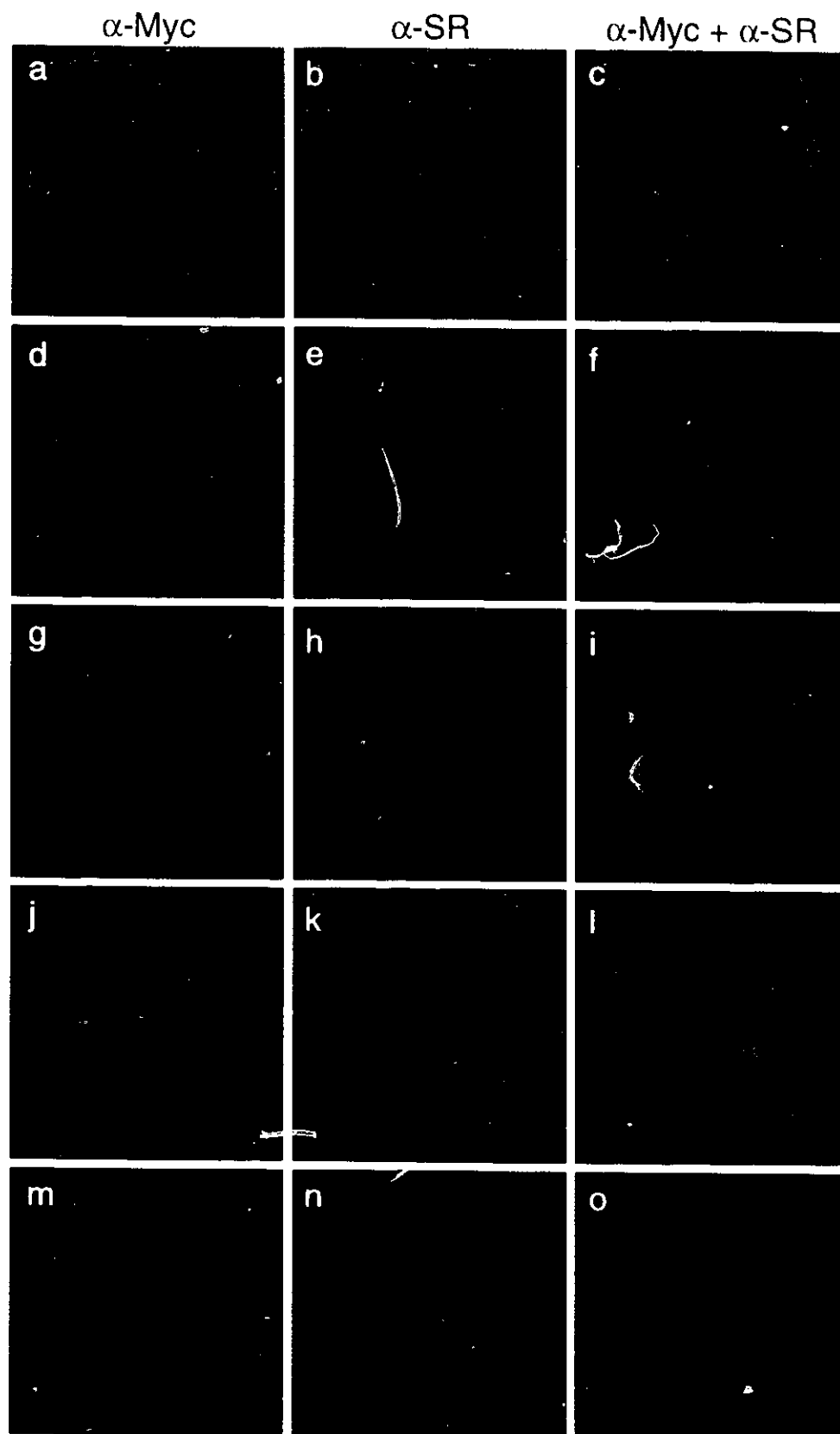
Since anti-SC35 and anti-SR mAb 104 recognize phospho-epitopes, it is conceivable that the dispersion of the immunofluorescence signal representing the SR family seen when Clk1 is over expressed may be due to phosphorylation of SR members

Figure 1

Nuclear localization of Clk1 and its effects on the distribution of SR proteins.

Indirect immunofluorescent staining and confocal microscopy of transfected COS-1 cells using the anti-Myc mAb (α -Myc, panels *a,d,g,j* and *m*), anti-SR mAb 104 (α -SR, panels *b,e,h,k* and *n*), or an overlay of the two signals (α -Myc + α -SR, panels *c,f,i,l* and *o*). Alignment of green and red signals appears yellow. Both low power magnification (25X) of M-Clk1^{K190R} transfected COS-1 cells (panels *d-f*) and M-Clk1 transfected cells (panels *j-l*) and high power magnification (63X) of untransfected COS-1 cells (panels *a-c*), M-Clk1^{K190R} (panels *g-i*) and M-Clk1 (panels *m-o*) transfected cells is shown.

Figure 1



outside of the speckle once Clk1 is released from the speckles. To ensure that over expression of Clk1 caused disassembly of speckles and redistribution of SR family members rather than phosphorylation of SR members outside the speckles, the anti-Sm antibody, which recognizes components of the snRNPs, was used (69). When Clk1^{K190R} is over expressed, the Sm proteins are co-localised with Clk1^{K190R} in speckles (Fig. 2). Over expression of Clk1 causes a redistribution of the Sm proteins similar to the SR family, indicating that Clk1 does indeed cause disassembly of nuclear speckles and subsequent redistribution of speckle proteins.

The apparent contrasting subnuclear localization of Clk1 and Clk1^{K190R} proteins as detected by immunofluorescent staining was supported by the observation that these two proteins display differential solubility in a non-ionic detergent. COS-1 cells transfected with either M-Clk1 or M-Clk1^{K190R} were separated into Triton X-100 soluble and insoluble fractions as described in Materials and Methods and analysed by immunoblotting (Fig. 3). While essentially all of the Clk1 protein was soluble in Triton X-100, a significant portion of the Clk1^{K190R} protein remained insoluble. It is of interest to note that the active kinase, Clk1, migrates as a broad band while the inactive point mutant, Clk1^{K190R}, migrates as a discrete band following SDS-PAGE. This banding is suggestive of multiple phosphorylation states of the active kinase, as has been observed for other kinases such as MAPK and Wee1 (102,103).

When Triton X-100 treated COS-1 cells were subjected to analysis by immunofluorescence, the speckled pattern of Clk1^{K190R} remained intact (Fig. 3) while the wild type protein staining pattern was lost (data not shown). These results are consistent

Figure 2

Effect of Clk1 on the distribution of Sm proteins.

Indirect immunofluorescent staining and confocal microscopy of transfected COS-1 cells using the anti-Myc mAb (α -Myc, panels *a, d, g* and *j*), anti-Sm antisera (α -Sm, panels *b, e, h* and *k*), or an overlay of the two signals (α -Myc + α -Sm, panels *c, f, i* and *l*). Alignment of green and red signals appears yellow. Both low power magnification (25X, panels *a-c, g-i*) and high power magnification (63X, panels *d-f, j-l*) are shown.

Figure 2

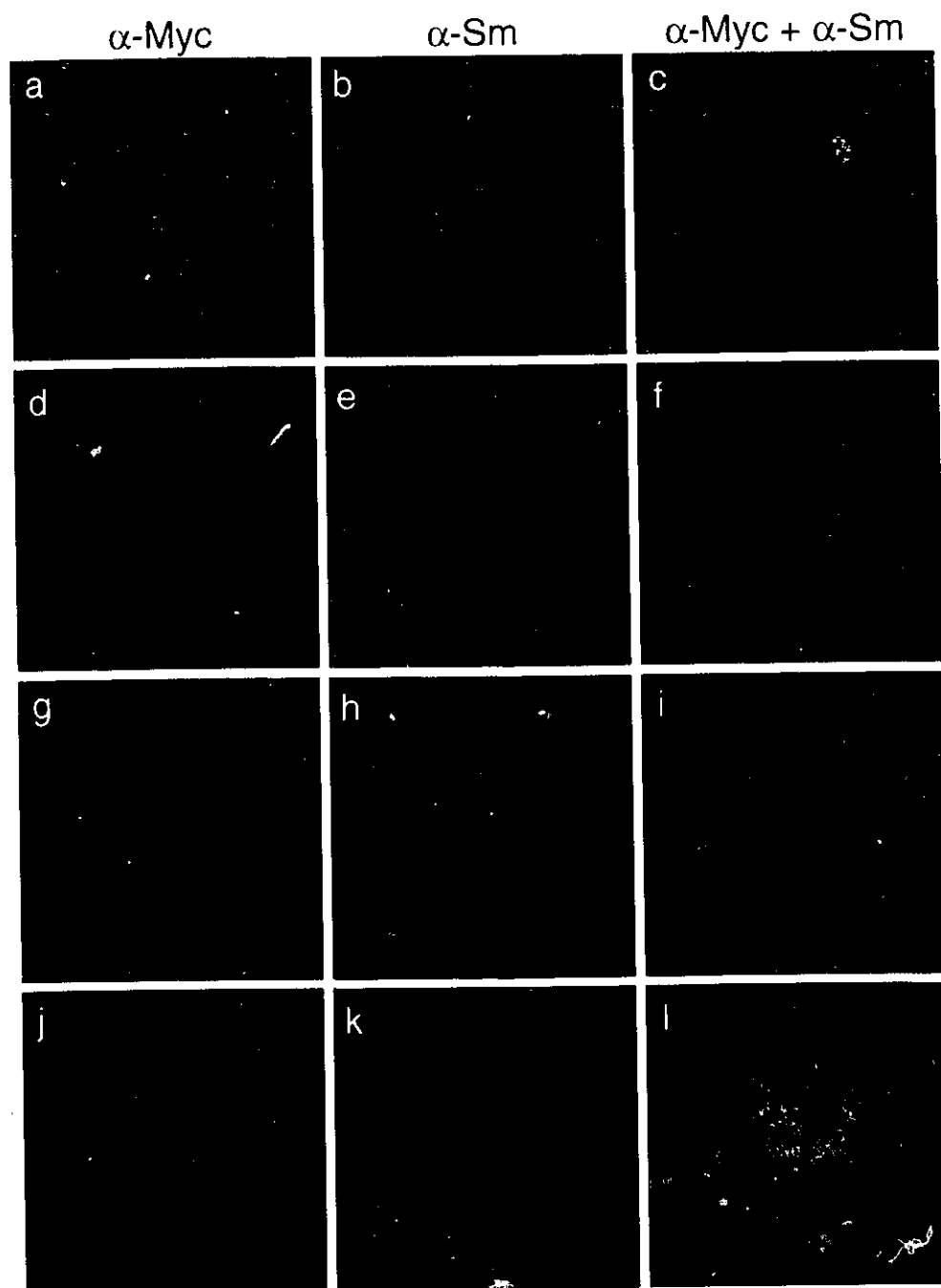
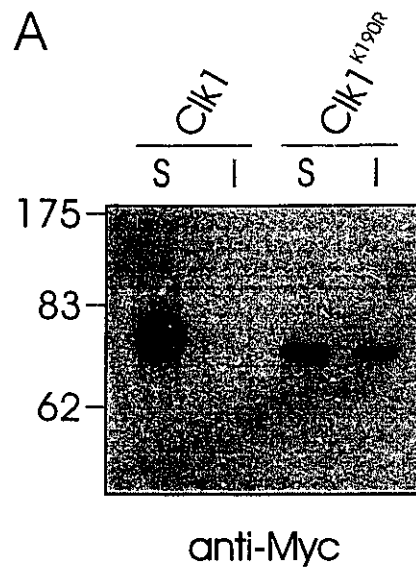


Figure 3

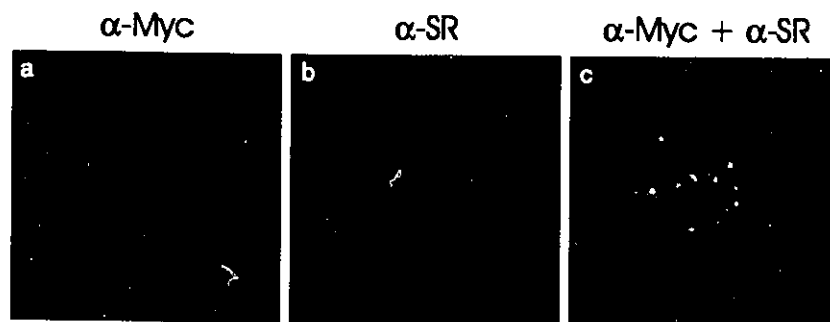
Active and inactive Clk/STY kinases display differential solubility in non-ionic detergent.

(A) COS-1 cells transfected with pECE expression vectors encoding Myc epitope tagged Clk1 (M-Clk1) or a catalytic mutant M-Clk1^{K190R} were lysed in Triton X-100 containing buffer (see Materials and Methods) to yield a Triton-soluble fraction (S) and a Triton-insoluble fraction (I). Extracts were analysed by SDS-PAGE and immunoblotting with the anti-Myc antibody as described. Positions of molecular mass markers (kDa) are indicated to the left. (B) The Triton X-100 insoluble Clk1^{K190R} co-localises with SR splicing proteins. COS-1 cells transfected with M-Clk1^{K190R} and grown on coverslips were extracted with Triton X-100-containing buffer. Following extraction cells were fixed and processed for indirect immunofluorescence by confocal microscopy using the anti-Myc antibody (α -Myc) and the anti-SR antibody (α -SR). (a) α -Myc, (b) α -SR, (c)overlay of the α -Myc and α -SR signals. Alignment of the green and red signals appears yellow.

Figure 3



B



with the idea that the catalytic activity of Clk1 is involved in the regulation of its subnuclear localization.

DISCUSSION

Clk1 interacts with RNA-Binding Proteins

Potential binding partners for the Clk1 protein kinase have been isolated by screening a T cell cDNA library using the yeast two hybrid system (62). Five of the clones isolated are RNA-binding proteins and three are members of the SR family of essential splicing factors. While the Clk1 kinase domain was not sufficient for a positive interaction with these RNA-binding proteins, the isolated amino-terminal RS region of Clk1 associated with ASF/SF2, X16, and RNP S1. This result is consistent with previous data implicating RS domains in protein-protein interactions (12,36,104). A characteristic of previously described RS domains is the presence of repeating RS/SR dipeptides (reviewed in (34,35)). Although Clk1 does not contain long stretches of RSRS repeats, it contains ten RS/SR dipeptides and one RSRS motif. This amino-terminal domain is also sufficient to target Clk1 to nuclear speckles *in vivo* (see Chap. 4), a feature that has been shown to exist for other RS domains (38,39).

The amino-terminal domain of Clk1 therefore possesses many of the features of known RS domains, including an involvement in nuclear targeting and an ability to interact with RS proteins.

Phosphorylation of ASF/SF2

Phosphorylation/dephosphorylation of splicing factors may be an important aspect of splicing regulation (57,105), and SR protein function *in vitro* can be affected by phosphorylation (58,106). *In vitro*, Clk1 efficiently phosphorylated ASF/SF2, predominantly in the ASF/SF2 RS domain on serine residues. Phosphothreonine was only seen in the RNA-binding domain and not in the RS domain, therefore, it is likely that the RNA-binding domain of ASF/SF2 was phosphorylated to a minor extent by Clk1. The *in vitro* sites of phosphorylation were also found to be phosphorylated *in vivo* suggesting that ASF/SF2 is a physiological substrate of Clk1. There are additional sites phosphorylated *in vivo* that cannot be accounted for by Clk1 phosphorylation *in vitro*, suggesting that Clk1 is not the only kinase that phosphorylates SR proteins. Indeed, two other kinases have been identified that can phosphorylate ASF/SF2 *in vitro*, SRPK1 (106) and an uncloned U1 70K kinase (61). In contrast to Clk1, neither of these kinases are able to phosphorylate ASF/SF2 outside of the RS domain.

Subnuclear Localization of Clk

It has been shown previously that Clk1 is localised to the nucleus (96). In this study, we show that a kinase inactive mutant of Clk1 is localised to discrete speckles within the nucleus where SR proteins are located. The truncated Clk1^T protein is encoded by an alternatively spliced Clk1 transcript, and contains only the RS region. At present, the function and localization of the truncated Clk1^T polypeptide is unclear. However, the tight association of the Clk1 RS domain with SR proteins suggests that Clk1^T could be

involved in the regulation of SR protein compartmentalization. This could occur by the formation of inactive Clk1:Clk1^T heterodimers or by sequestering of SR proteins by Clk1^T (93,94). In contrast to the Clk1 kinase-inactive mutant, wild-type Clk1 displayed a more diffuse nuclear distribution. *In vivo*, kinase-inactive Clk1 appears to be hypophosphorylated whereas the wild-type protein appears hyperphosphorylated presumably as a consequence of autophosphorylation. This result indicates that phosphorylation may oppose retention in the speckles. This hypothesis is further borne out by the loss of SR proteins from the speckles when wild-type Clk1 is over expressed. We propose that the SR proteins and Clk1 are targetted to the speckles through their RS domain (38,39) but are released from this site upon phosphorylation.

Nuclear speckles have been proposed to act as storage sites for splicing components (67,70,71,107). There is no evidence that speckles are disassembled prior to the onset of splicing (67) and this may only occur when the kinase responsible for the release of factors is over expressed. One possibility is that at physiological levels of Clk1 only a subset of SR factors are phosphorylated and released from the speckles. The availability of a subset of SR factors may be a significant regulatory mechanism, given the influence of the relative concentration of individual splicing factors on pre-mRNA processing (42,45,46,108). Due to the complexity of factors involved in pre-mRNA processing it seems likely that multiple kinases, perhaps other members of the LAMMER family, regulate the many facets of splicing. We propose that Clk1 and SRPK1 are two members of a family of kinases that regulate SR protein trafficking and activity.

CHAPTER 4

Splicing of the SR Protein Kinase Clk1 by an Autoregulatory Mechanism

SUMMARY

Control over the expression of cellular genes through constitutive and alternative precursor mRNA splicing requires SR (serine/arginine-rich) proteins. Reversible protein phosphorylation is essential for both spliceosome assembly and the catalytic process of splicing and it has been suggested that SR proteins may be one of the key targets of regulated protein kinase activity. The Clk1 kinase phosphorylates SR proteins and causes their redistribution within the nucleus. The primary *Clk1* transcript contains an alternatively spliced exon (EB), yielding mRNAs encoding catalytically active and truncated inactive polypeptides (Clk1 and Clk1^T respectively). We present evidence that Clk1 and Clk1^T proteins autoregulate the splicing of the *Clk1* pre-mRNA *in vivo*. We show that the peptide domain encoded by exon EB is essential for the regulatory activity of the Clk1 kinase. Taken together, our results demonstrate that Clk1 kinase is involved in the control of alternative splicing.

MATERIALS AND METHODS

Construction of Plasmids

The *Clk1* mini-gene (containing introns A and B, (109)) was created in the plasmid pGEM-4 (Promega). pGEM-4/*Clk1* was created by the insertion of the *Clk1* cDNA (*KpnI/XbaI* fragment from pECE/*Clk1* (96) into pGEM-4 previously digested with *KpnI* and *XbaI*. pGEM-4/*Clk1*^{K190R} was generated by PCR-mediated mutagenesis using pGEM-4/*Clk1* plasmid as a template and the following oligonucleotides: 5'gcatagttaaaaatgtggatag3' (663R) and 5'gcactgctacacgtctac3' (659F) as previously described (109).

The mini-genes in pGEM-4 (called pGEM-4/*Clk1*[1-2-3] and pGEM-4/*Clk1*^{K190R}[1-2-3]) were created by substitution of the 236-bp *NsiI* fragment from pGEM-4/*Clk1* or pGEM-4/*Clk1*^{K190R}, respectively, with the 1501-bp *NsiI* fragment from pGEM-4/*Clk1* 5.6 ((109); this fragment encompasses Intron A-Exon B-Intron B). The mini-genes were subsequently subcloned into the pECE expression vector by substitution of the ~1450-bp *SalI/XbaI* fragment from pECE/*M-Clk1* with a ~2.7-kbp *SalI/XbaI* fragment from either pGEM-4/*Clk1*[1-2-3] or pGEM-4/*Clk1*^{K190R} to create pECE/*M-Clk1*[1-2-3] or pECE/*M-Clk1*^{K190R}[1-2-3]. The first exon of the minigene is fused in frame with a human Myc epitope to facilitate identification of protein products (60).

To facilitate future splicing analyses the mini-genes were subcloned into the pcDNA3 expression vector (Invitrogen). The ~3-kb *Clk1* mini-gene was isolated from either pECE/*M-Clk1*[1-2-3] or pECE/*M-Clk1*^{K190R}[1-2-3] by PCR amplification using the oligonucleotides: 5'tattccagaagtagtgagg3' and 5'ggcgcactagtcgtatgccttttaagtgg3' (1539F, which includes a *SpeI* site at the 5' end). 50μl reactions were performed containing 200ng

plasmid, 25pmol each oligonucleotide, 0.2mM each dNTP, 2mM MgSO₄, 1 unit Vent DNA Polymerase (New England Biolabs) in buffer supplied by the manufacturer. Amplification conditions were 94°C, 30sec; 48°C, 30sec; 72°C, 3min for 4 cycles followed by 94°C, 30sec; 60°C, 30sec; 72°C, 3min for 20 cycles. The reaction products were digested with *SpeI* and subcloned into pcDNA3 previously digested with *EcoRV* and *XbaI* to create the plasmid mini-genes CMV/*Clk1* and CMV/*Clk1*^{K190R}. Generation of these plasmids resulted in the loss of *Clk1*'s normal translation stop codon. Twelve new amino acids are added before a stop codon is encountered within the pcDNA3 vector (KHTTRGPYSIVSPKC.).

Mini-gene CMV/*CR-1* was generated by removal of the 709bp *ClaI* fragment from the *Clk1* catalytic domain in CMV/*Clk1*. The resulting expression plasmid generates *Clk1*^T and a truncated version of *Clk1* (amino acids 1-179 of *Clk1*).

CMV/*CR-2* was generated by digestion of CMV/*Clk1* plasmid with *BamHI*, subsequent blunt-ending with DNA polymerase (Klenow fragment) and religation. The outcome of this is to introduce a translation stop codon (TGA) following the Myc epitope tag. CMV/*CR-2* does not encode any *Clk1* polypeptides.

To create pECE/*M-Clk1*^{Δ130-158} the *Clk1* catalytic domain (amino acids 159-483) was PCR amplified using oligonucleotides 5'gagctcgagatatgaaattgttgat^{3'} (567R; including a *XhoI* site at the 5' end) and 5'ggcgcactagtcgtatgcttttaagtgg^{3'} (1539F; including a *SpeI* site at the 5' end). The cDNA fragment encoding the Myc epitope tag and *Clk1* amino acids 1-129 was PCR amplified using oligonucleotides 5'cgactagtgatcgatcagctatg^{3'} (MTG3; including a *SpeI* site at the 5' end) and 5'cccgtgtgaatggtgctg^{3'} (479F). PCR amplifications were

performed as described earlier. The amplified catalytic domain was digested with *XhoI* and *SpeI* and ligated into pBK-RSV (Stratagene) previously digested with *XhoI* and *XbaI* to create pBK-RSV/*Clk1*[*cat.dom.*]. The amplified regulatory domain was digested with *SpeI* and subcloned (*SpeI*/blunt) into pBK-RSV/*Clk1*[*cat.dom.*] prepared by digestion with *XhoI*, blunt ending with T4 DNA polymerase, and subsequent digestion with *SpeI* to create pBK-RSV/*M-Clk1* ^{$\Delta 130-158$} . Ligation of these two fragments introduces a novel serine residue between Clk1 amino acids 129 and 158. To generate pECE/*M-Clk1* ^{$\Delta 130-158$} the *M-Clk1* ^{$\Delta 130-158$} fragment was isolated by *SacI* digestion, blunt ending with T4 DNA polymerase, subsequent digestion with *KpnI* and ligated into the pECE vector (*HindIII* digested, blunt ended with T4 DNA polymerase, followed by *KpnI* digestion). Generation of pECE/*M-Clk1* ^{$\Delta 130-158$} resulted in the loss Clk1's normal translation stop codon. Two new amino acids are added before a stop codon is encountered within the pECE vector (KHTTR.).

COS-1/293T Transfections

Cells (COS-1 (110) and 293T/17 (111))(60-75% confluent 60-mm dish) were transfected with 18 μ l Lipofectamine (Gibco BRL) according to the manufacturer's instructions. Typically, 2.5 μ g CMV/*CR-1* or 0.1 μ g CMV/*CR-2* was cotransfected with *Clk1* or control plasmids totalling 5 μ g. pECE and pcDNA3 DNA was used as filler to maintain constant amounts of transfected DNA for *M-Clk* and *HA-Clk* plasmids, respectively. Cells were harvested after 24 hours in PBS. Protein lysates were generated from one quarter of the cells by the addition of sample buffer. Total RNA was extracted from the remaining three quarters using RNeasy purification (Qiagen) as per the manufacturer's instructions.

RT-PCR Analysis

Total RNA samples (100ng), isolated using RNeasy (Qiagen), were processed for RT-PCR by reverse transcription followed by amplification using the following primers: 5'tggtaggagtgaagaag³' (corresponding to positions 397-414 in the *Clk1* cDNA (18)) and 5'gatggctggcaactagaa³', directed against the Bovine Growth Hormone polyadenylation signal of the pcDNA3 vector. PCR conditions, including number of cycles and template concentrations, were optimized to maintain linearity during amplification. Linearity was monitored by serial dilution of the template cDNAs, demonstrating a corresponding decrease in signal, as well as consistent ratios between bands of interest (EB⁺/EB⁻). PCR products were separated on 1.5% agarose gels stained using SYBR Green I (Molecular Probes), visualized and quantitated using the STORM fluorescence imager and ImageQuaNT software (Molecular Dynamics).

The identity of the spliced *CR-1* mRNAs was confirmed by direct sequencing of the PCR products. As in figure 1, the larger PCR product is derived from spliced RNA containing exon EB and the smaller product corresponds to spliced RNA in which exon EB has been skipped.

Immunoblot Analysis

Aliquots of protein were resolved by 10% SDS-PAGE and transferred to nitrocellulose. Membranes were blocked with 5% skim milk powder in TBST (150mM NaCl, 10mM Tris-HCl, pH 7.5, 0.05% Tween 20) and subsequently probed with a mix of either anti-HA monoclonal antibody (mAb) (BAbCO Inc.) and anti- β -tubulin mAb (112) or anti-Myc

9E10 mAb and anti- β -tubulin mAb. Immunoblots were visualised with Horseradish peroxidase-conjugated goat anti-mouse antibody followed by enhanced chemiluminescence and exposure to Reflection film (Dupont NEN).

Immunofluorescence

COS-1 cells were transfected with pECE/*M-Clk1* expression vectors, fixed and stained as previously described (62). Cells were viewed by confocal microscopy using an upright Leica Confocal Laser Scanning Microscope equipped with a 55 mW krypton/argon air-cooled laser and a 63X Plan Apo oil immersion lens.

RESULTS AND DISCUSSION

We have demonstrated previously that the *Clk1* pre-mRNA is alternatively spliced (96) and interacts with SR proteins (62). To determine if the Clk1 kinase (18,19) can also elicit changes in pre-mRNA splicing we constructed a *Clk1* mini-gene which contains two introns flanking the alternatively spliced exon, referred to in the following as EB (Fig. 1A). Splicing of the mini-gene to retain exon EB (exon inclusion) generates an mRNA encoding full-length catalytically active Clk1 whereas splicing to exclude exon EB (exon skipping) generates an mRNA encoding Clk1^T (96). When expressed in COS-1 cells the Clk1 mini-gene leads to the production of predominately Clk1^T protein as determined by immunoblot analysis (Fig. 1B, lane 2). In contrast, when a point mutation is introduced within the ATP binding fold of the catalytic domain of the Clk1 mini-gene rendering the full-length Clk1 protein catalytically inactive (Clk1^{K190R}), significantly more full-length protein is produced (Figure 1B, lane 3). Analysis of the mRNA produced from the two mini-genes, by reverse transcription and quantitative PCR amplification (RT-PCR), supported the immunoblot results. As can be seen in figure 1 (C and D, lane 2), cells transfected with the wild type *Clk1* mini-gene predominantly express mRNA encoding Clk1^T protein. The catalytically inactive *Clk1*^{K190R} mini-gene construct produced significantly more mRNA encoding full-length Clk1^{K190R} protein (Fig. 1C and D, lane 3). These results are consistent with the idea that the catalytic activity of the Clk1 kinase regulates the splicing of the *Clk1* pre-mRNA.

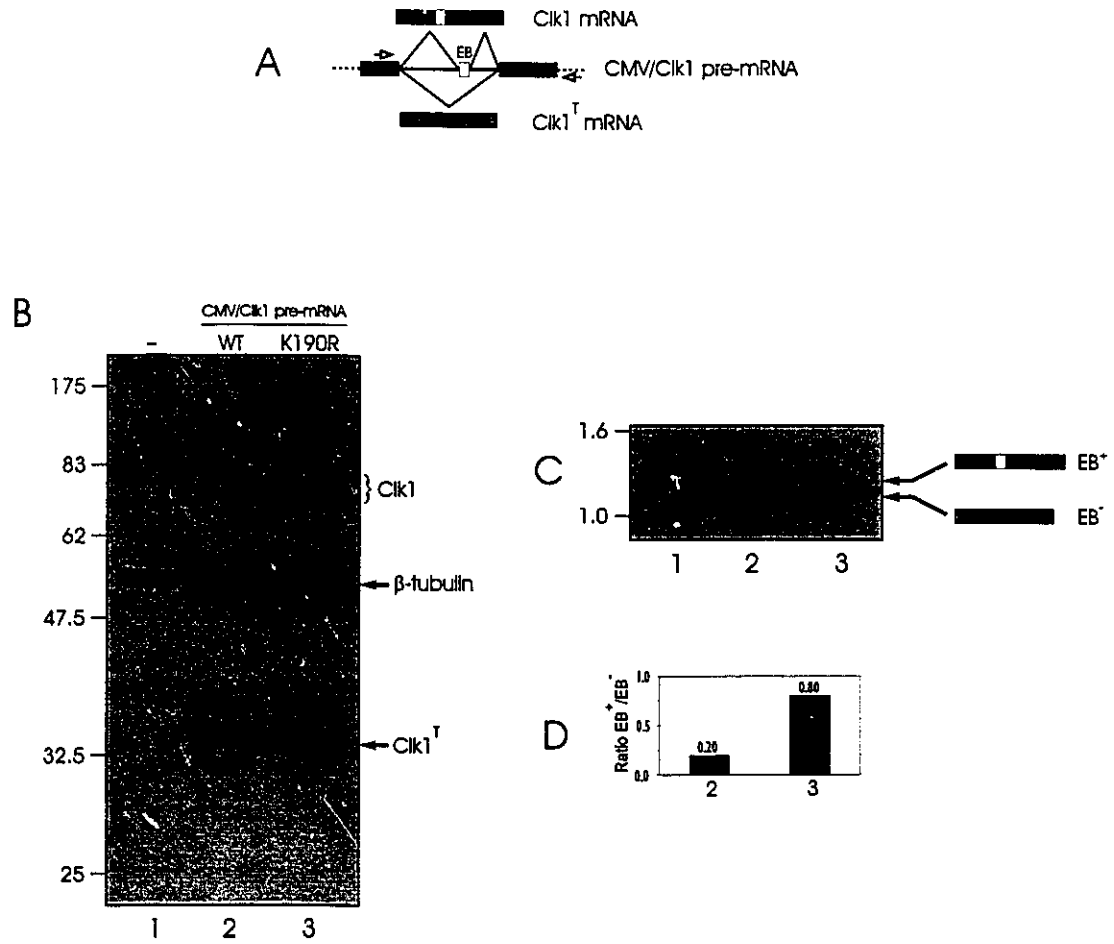
To determine the relative contributions of the Clk1 and Clk1^T proteins to the alternative splicing of *Clk1* pre-mRNA, two smaller mini-gene constructs were created

Figure 1

Catalytic activity of Clk1 kinase regulates splicing of *Clk1* pre-mRNA *in vivo*.

A, Schematic representation of the *Clk1* pre-mRNA mini-gene and the spliced mRNA products. The alternatively spliced exon (EB) and primers used for amplification by RT-PCR are shown. CMV, cytomegalovirus promoter. **B**, Clk1 and Clk1^T proteins expressed in cells following transfection of CMV vector (lane 1), CMV/*Clk1* pre-mRNA (lane 2) and CMV/*Clk1*^{K190R} pre-mRNA. Clk1 proteins were detected by anti-Myc mAb immunoblot. Anti- β -tubulin mAb was used to control for equal loading of proteins. Positions of molecular mass markers (kDa) are indicated to the left of the panel. WT, wild-type kinase. **C**, Detection of *Clk1* mRNAs expressed following RT-PCR amplification; lanes are as in **B**. Positions of molecular size standards (kb) are indicated to the left of the panel. **D**, Ratio of *Clk1* mRNAs shown in **C**; lanes are as in **B**.

Figure 1



(CMV/CR-1 and CMV/CR-2, Fig. 2A). The CR-1 mini-gene contains a large deletion within the catalytic domain and has the capacity to produce only truncated Clk1 proteins and no full length kinase while the CR-2 mini-gene contains a translational stop codon following the Myc epitope tag which prevents the production of any Clk1 related proteins. The CR-1 mini-gene was co-transfected into the human cell line 293T (113) along with increasing amounts of an expression vector encoding a haemagglutinin epitope tagged version of Clk1 (HA-Clk1,(62)). The RNA products of the CR-1 reporter were analysed by quantitative RT-PCR amplification (Fig. 2). As can be seen in figure 2B, lane 2, transfection of the CR-1 reporter on its own, gave rise to mRNA products in a ratio similar to that seen with *Clk*^{K190R} mini-gene (compare Fig. 2B lane 2 and Fig. 1C, lane 3). Co-transfection of the HA-Clk1 expression vector promoted exon EB skipping. The ratio of exon EB inclusion to exon EB skipping decreased as the HA-Clk1 protein level increased (Fig 2, lanes 2-5) This experiment is representative of five separate transfections which gave 6-10 fold decreases in the ratio of exon EB inclusion:exon EB skipping,(EB⁺/EB⁻). Thus, overexpression of Clk1 kinase *in vivo* can promote exon EB skipping in COS-1 and 293T cells.

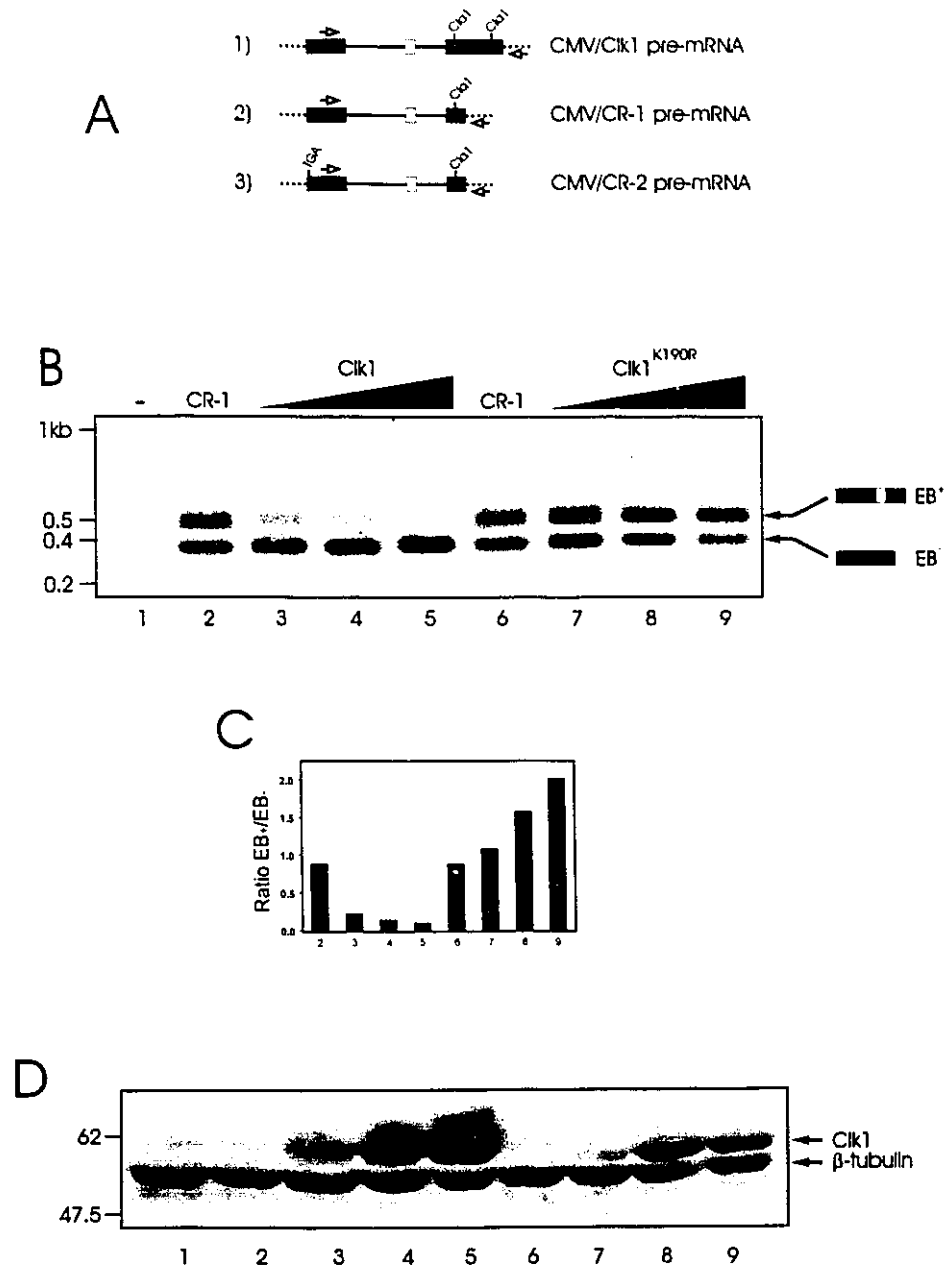
Co-transfection of CR-1 with HA-Clk1^{K190R} resulted in a modest (2-3 fold) but reproducible increase in the ratio of EB⁺/EB⁻ (Fig. 2B, lanes 6-9). One interpretation of this result is that the catalytically inactive Clk1^{K190R} protein can act as a dominant negative inhibitor of the endogenous 293T human Clk1 protein kinase and thus favours exon EB inclusion.

Figure 2

Catalytically active Clk1 promotes exon skipping of *Clk1* pre-mRNA *in vivo*.

A, Schematic representation of *Clk1* mini-genes. The primers used for amplification by RT-PCR are indicated. **B**, Pattern of *Clk1* alternative splicing upon cotransfection of *Clk1* and *Clk1*^{K190R} expression vectors. Lane 1, CMV vector, lanes 2-9 2.5µg CMV/*CR-1* + 0.5, 2.5, 5µg CMV/*HA-Clk* (lanes 3-5) or CMV/*HA-Clk*^{K190R} (lanes 7-9), respectively. Positions of molecular size standards (kb) are indicated to the left of the panel. **C**, Ratio of *Clk1* mRNAs shown in B.; lanes are as in B. **D**, Detection of Clk1 proteins expressed by anti-HA mAb immunoblot; lanes are as in B. Positions of molecular mass markers (kDa) are indicated to the left of the panel.

Figure 2



We have previously suggested that Clk1^T may be a natural antagonist to Clk1 activity (96). Since the CR-1 reporter construct does have the coding capacity to produce Clk1^T, we compared the spliced products generated from CR-1 and CR-2, which produces no Clk1 related proteins. In figure 3A it is evident that increasing amounts of CR-1 transfected per cell favours exon EB inclusion. We believe that this is a result of an increase in Clk1^T protein production per transfected cell. To test this idea a fixed amount of CR-1 was transfected along with increasing amounts of a Clk1^T expression vector and again we observed an increase in exon EB inclusion (Fig. 3C and D). In contrast, increasing amounts of the CR-2 reporter, which cannot produce Clk protein, does not alter the EB⁺/EB⁻ ratio (Fig. 3A and B). Thus, overexpression of catalytically active Clk1 favours skipping of exon EB, while conversely, Clk1^T promotes exon EB inclusion.

We next determined whether regions other than the catalytic domain of Clk1 are involved in the promotion of exon EB skipping. In particular, since portions of exon EB are conserved between the three known Clk family members (21) we created a deletion mutant of Clk1 lacking exon EB (herein referred to as Clk1^{Δ130-158}). The Clk1^{Δ130-158} protein when expressed in bacteria displayed considerable catalytic activity indicating that it retains a functional kinase domain (data not shown). Co-transfection of CR-1 with Clk1^{Δ130-158} in 293T cells resulted in a splicing pattern similar to that seen with Clk1^{K190R} and Clk1^T (Fig. 4A and B).

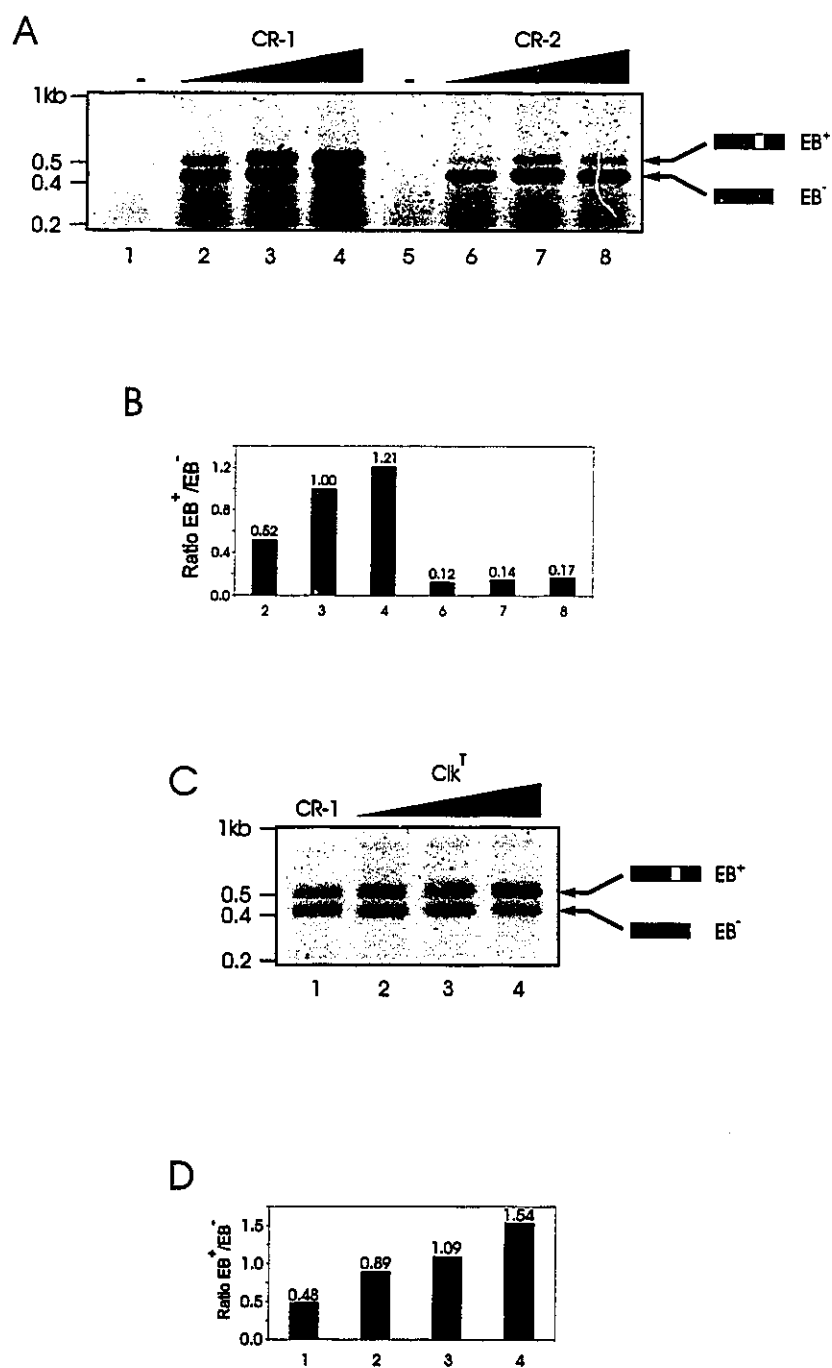
We have shown previously that Clk1 kinase can alter the subnuclear distribution of SR proteins ((62) and Fig. 4D, a-c). Clk1^T co-localises with SR proteins in nuclear speckles but does not affect their distribution (Fig. 4D, d-f). The Clk1^{Δ130-158} protein, which

Figure 3

Clk1^T promotes exon inclusion of *Clk1* pre-mRNA *in vivo*.

A, Pattern of *Clk1* alternative splicing upon cotransfection of Clk1^T expression vector. Lane 1-4, 2.5μg CMV/*CR-1* + 0, 0.5, 2.5, 5μg pECE/*M-Clk1^T* (109), respectively. Positions of molecular size standards (kb) are indicated to the left of the panel. **B**, Ratio of *Clk1* mRNAs shown in A.; lanes are as in A. **C**, Increasing expression of *CR-1* but not *CR-2* *Clk1* pre-mRNA promotes exon inclusion. Cells were transfected with 0 (lanes 1,5), 0.5 (lanes 2,6), 2.5 (lanes 3,7) or 5μg (lanes 4,8) CMV/*CR-1* (lanes 1-4) or CMV/*CR-2* (lanes 5-8). Positions of molecular size standards (kb) are indicated to the left of the panel. **D**, Ratio of *Clk1* mRNAs shown in C.; lanes are as in C.

Figure 3



was predominantly nucleoplasmic, like Clk1, but was also found at a low but detectable level in speckles, did not, however, significantly affect SR protein speckles (Fig. 4D, g,h,i). We conclude that the peptide motif encoded by exon EB is essential for both SR protein re-distribution and exon EB skipping.

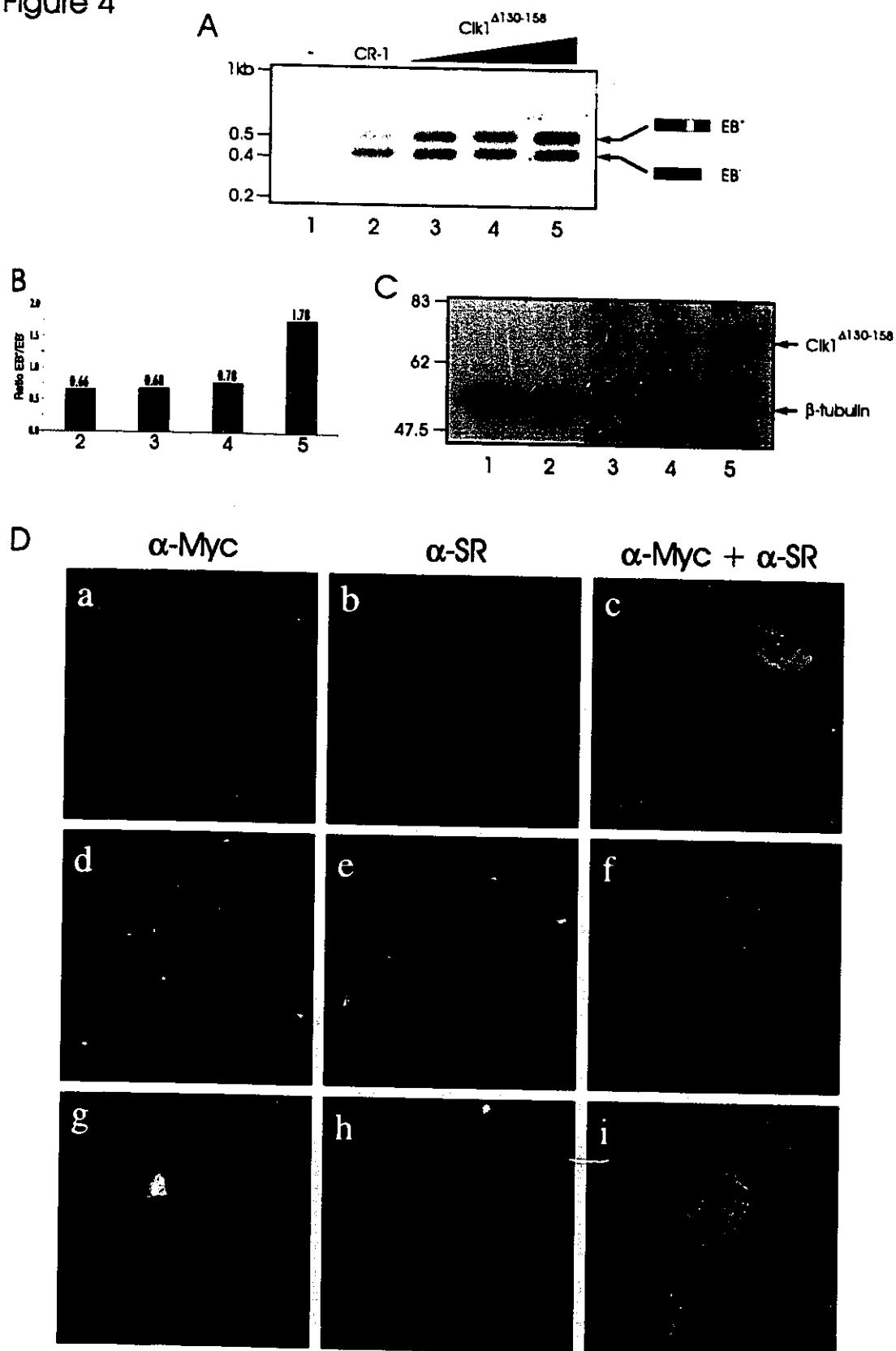
Alterations in SR protein concentration are thought to be critical to the control of alternative splicing (45,46,50,114). Indeed exon skipping and inclusion can be modulated by both the concentration of SR proteins and the placement of the splicing enhancers they bind within the pre-mRNA (52,55). Alternatively, SR protein activity could be modulated in a more rapid and dynamic fashion by post-translational phosphorylation (57-59,106). With this in mind, we propose the following model as a mechanism for the auto-regulation of *Clk1* gene expression. Clk1 kinase phosphorylates an SR protein(s) and *inhibits* its ability to stimulate exon inclusion, while Clk1^T protein forms inactive heterodimers with Clk1, thus reducing SR protein phosphorylation and *promoting* exon inclusion. We suggest that *in vivo*, members of the Clk kinase family (21) can phosphorylate distinct SR proteins thus impinging upon the regulation of alternative splicing programs throughout development. Furthermore, the Clk kinases may act as an interface between signal transduction pathways and the mRNA splicing machinery.

Figure 4

The peptide encoded by the *Clk1* alternatively spliced exon is necessary for exon skipping activity.

A, Pattern of *Clk1* alternative splicing upon cotransfection of $Clk1^{\Delta 130-158}$ expression vector. Lane 1, CMV vector, lanes 2-5 0.1 μ g CMV/*CR-2* + 0, 2.5 and 5 μ g pECE/*M-Clk1* ^{$\Delta 130-158$} , respectively. Positions of molecular size standards (kb) are indicated to the left of the panel. **B**, Ratio of *Clk1* mRNAs shown in A.; lanes are as in A. **C**, Detection of $Clk1^{\Delta 130-158}$ protein expression by anti-Myc mAb immunoblot; lanes are as in A. Positions of molecular mass markers (kDa) are indicated to the left of the panel. **D**, Nuclear localisation of *Clk1* isoforms and their effects on the distribution of SR proteins. Indirect immunofluorescent staining of transfected cells using the anti-Myc mAb (α -Myc, panels a,d,g) and anti-SR mAb (α -SR, panels b,e,h) or an overlay of the two signals (α -Myc + α -SR, panels c,f,i). Cells were transfected with pECE/*M-Clk1* (panels a-c), pECE/*M-Clk1*^T (panels d-f) or pECE/*M-Clk1* ^{$\Delta 130-158$} . Magnification 63X.

Figure 4



CHAPTER 5

Dissection of the Regulatory Domain of Clk1

INTRODUCTION

Successful signal transduction in mammalian cells requires a number of regulated and dynamic transitions. Two of the most important of these are protein- protein interactions and protein phosphorylation (reviewed in (115)). Protein domains implicated in protein-protein binding have been found in numerous diversely related polypeptides. These include Src homology 2 and 3 (SH2 and SH3) domains (116), the phosphotyrosine-binding (PTB) (117), WWP/WW (118,119), and RS domains (34) and 14-3-3 proteins (120). Interestingly, SH2, PTB and 14-3-3 domains mediate their protein-protein interactions through the binding of sequence-specific peptide motifs containing a phosphorylated residue. SH2 and PTB domains bind peptides containing phosphotyrosine (116), whereas 14-3-3 proteins bind phosphoserine containing peptide domains (120).

SH2 domains have been shown to mediate both intra and intermolecular interactions. Intermolecular binding to phosphotyrosine allows for the creation of multi-protein complexes, a feature desirable for transduction of a signal (115). An obvious advantage for an intramolecular interaction is that it can be used as a means for allosteric regulation of a protein.

Phosphorylation at other sites is also important for regulating the activity of many enzymes. Protein kinases have the unique ability to exert this regulation through intra and intermolecular phosphorylation. For example, phosphorylation within the ATP binding fold of cyclin dependent kinases prevents ATP binding within the catalytic domain and thereby inhibits kinase activity (reviewed in (17)). Many kinases also possess an "activation domain", found between subdomains VII and VIII of the catalytic domain, that

when phosphorylated leads to activation of the kinase (121). These examples of inhibitory and stimulatory (auto)phosphorylations can occur on serine/threonine and/or tyrosine residues and are believed to involve conformational/steric changes to the catalytic domain, thus affecting activity (17,121).

Protein kinases and in particular tyrosine and dual-specificity kinases are, therefore, poised to control signal transduction networks through the regulation of protein-protein complex formation and protein enzymatic activity.

Clk1 exhibits considerable serine/threonine and tyrosine autophosphorylating activity *in vitro* and *in vivo* and is thus a candidate for controlling one or more of the above mentioned regulatory activities/mechanisms. Biochemical characterization of the Clk1 kinase is important for a complete understanding of its physiological role. To this end, we have initiated an analysis of the amino-terminal non-catalytic region of Clk1 to identify potential phosphorylation sites, domains involved in protein complex formation and in the subnuclear localization of the kinase as well as domains necessary for the *in vivo* activity of Clk1 towards SR proteins. We report here that the amino acids encoded by the alternatively spliced exon (exon B) are essential for autophosphorylation of Clk1 *in vitro* and are required for the catalytic activity of Clk1 towards SR proteins *in vivo*. Furthermore, we present evidence that some of these residues can be phosphorylated *in vitro* by Clk1 and that phosphorylation of a tyrosine residue encoded by exon B may regulate the catalytic activity of Clk1 in mammalian cells.

MATERIALS AND METHODS

Plasmid Constructions - Clk1 Deletions

Clk1 deletion mutations were created from the pECE/M-Clk1 plasmid using PCR-mediated mutagenesis. Deletion mutations starting from amino acid 62 used the common oligonucleotide 5'ctacggtcgacctgctttccagataatag3' (271F; containing a *SaII* site at the 5' end) and a unique oligonucleotide as follows:

- 1) pECE/M-Clk1^{Δ62-83} : 5'catgcgtcgaccgagccagggcatcc3' (341R; *SaII* site at 5' end)
- 2) pECE/M-Clk1^{Δ62-93} : 5'catgcgtcgacaagcagataaccagatgca3' (371R; *SaII* site at 5' end)
- 3) pECE/M-Clk1^{Δ62-111} : 5'catgcgtcgacttacaaaagtaaacacagga3' (425R; *SaII* site at 5' end)
- 4) pECE/M-Clk1^{Δ62-129} : 5'catgcgtcgacgaagagtcaccgaagg3' (478R; *SaII* site at 5' end)
- 5) pECE/M-Clk1^{Δ62-160} : 5'atgaaattgttgatactttaggt3' (571R)

50μl reactions were performed containing 20μg pECE/M-Clk1, 25pmol each oligonucleotide, 0.2mM each dNTP, 2mM MgSO₄, 1 unit Vent or 2 units Vent(exo⁻) DNA Polymerase (New England Biolabs) in buffer supplied by the manufacturer. Typical amplification conditions were hot start reactions (5 min pre-heat at 95°C followed by the addition of enzyme) followed by 95°C, 60sec; 50°C, 60sec; 72°C, 5min for 2 cycles followed by 95°C, 60sec; 60°C, 60sec; 72°C, 5min for 25 cycles. The reaction products were digested with *SaII*, gel-purified and recircularized by ligation. pECE/M-Clk1^{Δ62-160} was generated by phosphorylation of the amplification product with T4 polynucleotide kinase, followed by gel-purification of the DNA and subsequent recircularization by ligation.

- Deletion constructs 1-4 and pECE/M-Clk1^{Δ1-60} carrying the K190R substitution (rendering the kinase domain catalytically inactive) were also generated by substitution of the 3.2-kb *BsrEII/PvuI* fragment (Clk1 amino acids 146-483) with that from pECE/M-Clk1^{K190R}.
- 6) pECE/M-Clk1^{Δ112-129} - Amino acids 1-111 were contained within a ~1.5-kb *PvuI/XhoI* fragment generated by digestion of PCR amplified DNA. pECE/M-Clk1 was amplified using oligonucleotides 5'tacggatcttccactcctacc3' (419F/*XhoI*, *XhoI* site at 5' end) and 5'gatgatgaacaccagccgccttggtatctacaaga3' (ST1100AA). This DNA was restricted with *XhoI* (Clk1 amino acid 111) and *PvuI* (which digests within the β-lactamase gene). Amino acids 130-483 were contained in a ~3.4-kb *SalI/PvuI* fragment isolated from pECE/M-Clk1^{Δ62-129} by *SalI/PvuI* digestion. Ligation of the ~1.5-kb *PvuI/XhoI* and ~3.4-kb *SalI/PvuI* fragments created the Clk1 deletion 112-129 and reconstituted the β-lactamase gene.
- 7) pECE/M-Clk1^{Δ112-129,K190R} was generated as above using the ~3.4-kb *SalI/PvuI* fragment derived from pECE/M-Clk1^{Δ62-129,K190R}.
- 8) pECE/M-Clk1^{Δ1-60/Δ112-129} was created by ligation of a ~1.2-kb *PvuI/SalI* fragment (Clk1 amino acids 62-111) from pECE/M-Clk1^{Δ1-60} and a ~3.6-kb *SalI/PvuI* fragment (Clk1 amino acids 130-483) from pECE/M-Clk1^{Δ112-129}.
- 9) pECE/M-Clk1^{Δ1-60/Δ112-129,K190R} was generated as above using the ~3.6-kb *SalI/PvuI* fragment isolated from M-Clk1^{Δ112-129,K190R}.
- 10) pECE/M-Clk1^{Δ1-165} (Clk1 catalytic domain) was created by substitution of the full-length Clk1 cDNA in pECE/M-Clk1 (~1.7-kb *EcoRI* fragment) with the Clk1 catalytic domain (~1150-bp *EcoRI* fragment). The catalytic domain was amplified from

pTZ19R/Clk1 (18) using oligonucleotides 5'cttgaattccttaggtgaaggtgct3' (591R; *EcoRI* site at 5' end) and 5'acgttctcgagcgacggccagtgc3' (UWXI), as above except PCR conditions were 95°C, 50sec; 60°C, 60sec; 72°C, 2min for 25 cycles. The reaction product was digested with *EcoRI*, gel-purified and ligated into pECE/M-Clk1.

11) pECE/M-Clk1(Δ3'UTR) - The 3'UTR of the Clk1 cDNA was removed by PCR amplification of Clk1 or Clk1^{K190R} cDNAs using the oligonucleotides 5'tcgggtacctcgagatgagacattcaaagaga3' (93R) and 5'ggcgactagtcgtatgcttttaagtgg3' (1539F; *SpeI* site at 5' end). Amplification conditions were similar to those described above except with the following cycling temperatures: 94°C, 30sec; 55°C, 30sec; 72°C, 2min for 2 cycles followed by 94°C, 30sec; 65°C, 30sec; 72°C, 2min for 20 cycles. The reaction product was digested with *BstEII* and *SpeI*, gel-purified and was subcloned into pECE/M-Clk1 following removal of the Clk1 fragment by *BstEII/XbaI* digestion. Generation of pECE/M-Clk1(Δ3'UTR) resulted in the loss Clk1's normal translation stop codon. Two new amino acids are added before a stop codon is encountered within the pECE vector (KHTTR.).

Junction Fragments - The nucleotide sequence and predicted amino acid translation product are indicated. Bolded nucleotide sequence indicates junction, often encoding a restriction site. Newly encoded non-Clk1 amino acids are bolded.

Δ62-83	S R S ⁶¹ T ⁸⁴ E P agc agg tcg acc gag cca	Δ62-93	S R S ⁶¹ T ⁹⁴ S R agc agg tcg aca agc aga
Δ62-111	S R S ⁶¹ T ¹¹² Y K agc agg tcg act tac aaa	Δ62-129	S R S ⁶¹ T ¹³⁰ K S agc agg tcg acg aag agt

Δ112-129 R S S¹¹¹ T¹³⁰ K S aga tcc tcg acg aag agt Δ62-160 S R S⁶¹ T V D¹⁶¹ E I agc agg tcg acc gta gat gaa att Δ1-60 (named ΔXhoII) P G S⁶¹ I N E ccc gga tcc ata aat gag	Δ1-165 L N S¹⁶⁶ L G E ttg aat tcc tta ggt gaa
---	--

12) pECE/M-Clk1^{Y160F,D} - Tyrosine 160 was converted to phenylalanine or aspartate using the pAlter site-directed mutagenesis system (Promega). The 1.7-kb Clk1 *EcoRI* cDNA fragment isolated from pECE/M-Clk1 was subcloned into the *EcoRI* site of pAlter-1 plasmid. Orientation was determined by restriction enzyme analysis. Nucleotides 570 and 571 were substituted using oligonucleotide 5'-ctaagtgaaga(t/g)(a/t)tgaaattgtga3' (Y563F/D, which is partially degenerate at two positions, nucleotides 570 and 571). This oligonucleotide will produce four codons: tat - Tyr, ttt - Phe, gat - Asp, gtt - Val. Plasmids carrying the codons **ttt** (pAlter/Clk1^{Y160F}) and **gat** (pAlter/Clk1^{Y160D}) were identified and chosen by sequencing. pECE/M-Clk1^{Y160F} and pECE/M-Clk1^{Y160D} were generated by substitution of the 550-bp *BstEII/NdeI* in pECE/M-Clk1 with that from pAlter/Clk1^{Y160F} and pAlter/Clk1^{Y160D}, respectively.

GST-Clk1 Deletions - pGEX-3X/Clk1^{K190R} (which expresses GST-Clk1^{K190R}) was created by substitution of the ~1.4-kb *SalI/EcoRI* fragment from pECE/M-Clk1^{K190R} with that from pGEX-3X/Clk1.

pGEX-3X/Clk1^{Δ130-158} (GST-Clk1^{Δ130-158}) and pGEX-3X/Clk1^{Δ62-160} (GST-Clk1^{Δ62-160}) were generated by substitution of the ~1.7-kb *Bam*HI/*Eco*RI Clk1 fragment from pECE/M-Clk1^{Δ130-158} and pECE/M-Clk1^{Δ62-160}, respectively, with that from pGEX-3X/Clk1.

COS-1 Cell Expression of Clk1 Polypeptides - Transfection of COS-1 cells, generation of protein lysates, immunoprecipitation, immunoblot and kinase assays were performed as described in chapter 2. In some cases Histone was included as an exogenous substrate during kinase assays.

GST/GST-Clk1 Purification - GST and GST-Clk1 binding reactions were carried out as described in chapter 2. Bacteria harbouring pGEX-3X (Pharmacia) or pGEX-3X/Clk1 plasmids were induced to express GST and GST-Clk1 fusion proteins, respectively, as per the manufacturer's instructions. Bacteria (50ml induced culture) were lysed in immunoprecipitation (IP) buffer (20mM Tris-HCL, pH 7.5, 150mM NaCl, 2mM EDTA, 1% Triton X-100, 2mM NaF, 2mM sodium pyrophosphate, 500μM sodium vanadate, 200μg/ml PMSF, 2μg/ml aprotinin, 5μg/ml leupeptin) by sonication three times for 20sec.. Lysates were cleared by centrifugation (10min., 4°C at maximum speed in a microcentrifuge). Lysates containing GST or GST-Clk1 were incubated with glutathione-Sepharose 4B beads (Pharmacia) for 1hr at 4°C with rocking. Beads were washed three times with 1ml IP buffer. To quantitate bound GST/GST-Clk1 an aliquot was removed, washed twice with PBS and protein concentrations determined by Bradford assay (BioRad).

Approximately 10 μ g GST or GST-Clk1 coupled to beads was used in each binding experiment.

GST/GST-Clk1 Binding Assay - COS-1 cells were transfected with plasmids and protein lysates extracted in IP buffer as described in chapter 2. Cell lysates were mixed with GST or GST-Clk1 beads and incubated for 2hr at 4°C with rocking. The beads were then washed three times with IP buffer and proteins eluted with sample buffer. Bound proteins were analysed by SDS-PAGE, transferred to nitrocellulose and subsequently immunoblotted with anti-Myc mAb. Immunoblots were visualised with HRP-conjugated goat anti-mouse antibody followed by enhanced chemiluminescence and exposure to Reflection film (Dupont NEN).

Peptide Synthesis and Phosphorylation - The Clk1 peptide (amino acids 130-164, KSHRRKRSRSVEDDEEGHLCQSGDVLSARYEIVD) was synthesized on an Applied Biosystems 431A peptide synthesizer by solid phase peptide synthesis, using a modification of Fluorenylmethoxycarbonyl (Fmoc) chemistry with HBTU (2-(1 H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate) used as the activating agent. Following synthesis the peptide was cleaved from the synthesis resin using a trifluoroacetic acid mixture (containing thiol scavengers to avoid side-chain reactions) for 2h at 22°C, precipitated and washed with T-butyl ether and lyophilised. Peptide purity was confirmed by mass spectroscopy analysis. The peptide, although insoluble in kinase buffer (20mM HEPES, pH 7.1, 10mM MgCl₂, 2mM MnCl₂), was soluble in diluted kinase buffer (20mM

HEPES, pH 7.1, 1mM $MgCl_2$, 0.2mM $MnCl$). For peptide phosphorylation immunoprecipitated pECE/M-Clk1 or affinity purified GST-Clk1 polypeptides were incubated with 10 μ g peptide, 10 μ M ATP and 5 μ Ci [γ - 32 P]ATP in 25 μ l diluted kinase buffer for 20 min at room temperature. Reactions were arrested by the addition of sample buffer followed by boiling. Samples were resolved by 15% or 7½-20% SDS-PAGE and exposure of the gel to storage phosphor screen (Kodak) and detection by PhosphorImager (Molecular Dynamics). Phosphoamino acid analysis was performed on phosphoproteins following transfer of proteins to nitrocellulose as described in chapter 2.

RESULTS

Phosphorylation Sites

To identify potential autophosphorylation sites within the putative regulatory domain (amino acids 1-160) of the Clk1 kinase we generated a series of deletions within this region (Fig. 1). Deletions were created by PCR based mutagenesis using pECE/M-Clk1 as a starting template. All deletion mutants retained the Myc-epitope tag at the amino terminus.

To determine if Clk1 deletion mutants expressed in mammalian cells retained catalytic activity we used an immune complex kinase assay. COS-1 cells were transfected with individual pECE plasmids expressing the series of deletion mutants. M-Clk1 polypeptides were isolated by anti-Myc immunoprecipitation followed by *in vitro* kinase assay. Deletion of the first sixty amino acids (Δ 1-60) has previously been shown to have no gross effect on autophosphorylation (Chap. 2, Fig. 5). Based on this result we decided to make deletions carboxyl-terminal to this site. Mutants lacking the next 67 amino acids (Δ 62-83, Δ 62-93, Δ 62-111 and Δ 62-129) all displayed significant autophosphorylation activity as assessed by *in vitro* kinase assay (Fig. 2A, lanes 1-4). A larger deletion encompassing the next thirty amino acids (Δ 62-160), although expressed at high levels (Fig. 2B, lane 2), displayed undetectable levels of autophosphorylation (Fig. 2A, lane 6). Deletion of the entire regulatory domain (Δ 1-165) resulted in a polypeptide which displayed no kinase activity as assessed by *in vitro* kinase assay (Fig. 2A, lane 5). Expression of the polypeptide was confirmed by anti-Myc mAb immunoblot (Fig. 2B, lane

Figure 1

Schematic Representation of Clk1 Regulatory Domain Deletions.

Deletions were created between amino acids 1 and 165 of Clk1 (defined here as the putative regulatory domain). All deletions have the Myc epitope tag fused at the amino terminus of Clk1. The alternatively spliced exon (exon B) is represented in black and constitutes amino acids 129-159. Deletions were generated with a wild-type catalytic domain or a catalytically inactive domain containing the substitution of amino acid lysine 190 for arginine (K190R). MTG - Myc epitope tag; KRKKR - putative nuclear localisation signal.

Figure 1

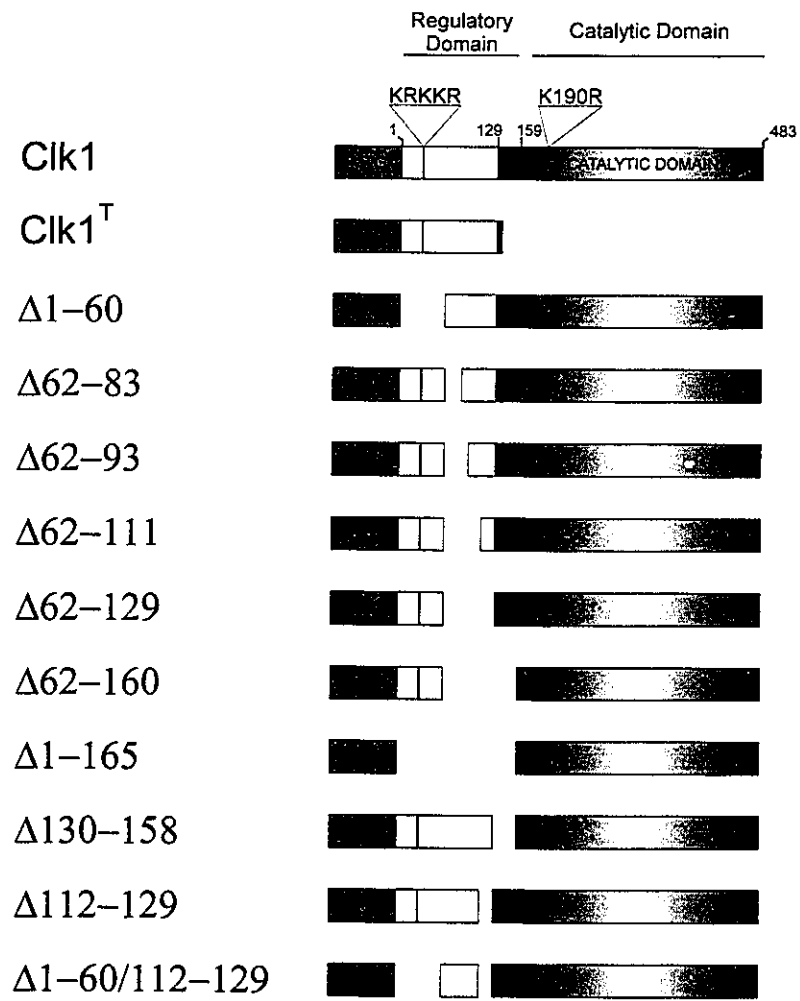


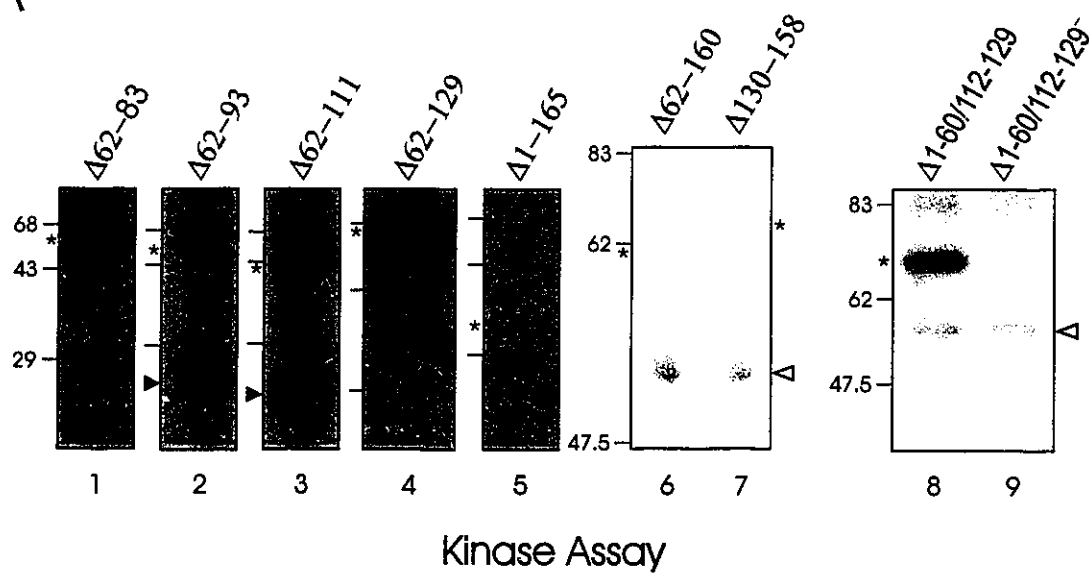
Figure 2

Expression and catalytic activity of Clk1 deletion mutants in COS-1 cells.

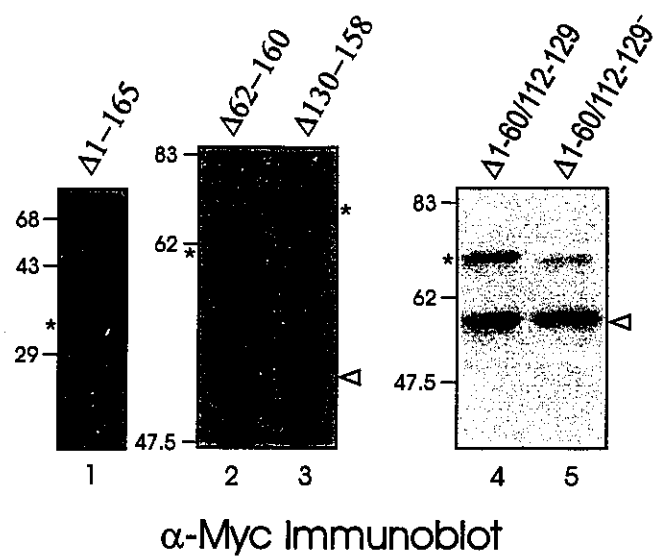
A, Autoradiogram of *in vitro* catalytic activities of Clk1 deletion proteins. **B**, Expression levels were determined for those displaying no catalytic activity by immunoblotting with the anti-Myc antibody followed by enhanced chemiluminescence. COS-1 cells were transfected with individual pECE expression vectors encoding Myc epitope-tagged Clk1 deletion mutants (Δ 62-83, 62-93, 62-111, 62-129, 1-165, 62-160, 130-158, 1-60/112-129 and 1-60/112-129; **A**, lanes 1-9, respectively). Myc-tagged proteins were immunoprecipitated with the anti-Myc monoclonal antibody and subjected to an *in vitro* kinase assay. Proteins were resolved by SDS-PAGE and transfer to nitrocellulose. Positions of molecular mass markers (kDa) are indicated to the left of each lane. The same three markers (68, 43 and 29-kDa) are found in **A**, lanes 1-5. The positions of individual Clk1 polypeptides are indicated with an asterisk. Open arrow head ($\triangleright$) indicates the position of immunoglobulin heavy chain; closed arrow head ($\blacktriangleright$) indicates the position of exogenously added histone.

Figure 2

A



B



1). The precise start of the catalytic domain is ill-defined, thus, a simple explanation for the lack of catalytic activity in this mutant is that some amino acids of the catalytic domain are deleted. The lack of activity in the smaller deletion, $\Delta 62-160$, may also be due to the removal of critical catalytic domain amino acids. To test this idea $\Delta 62-160$ was expressed in bacteria as a GST fusion protein (GST-Clk1 $\Delta 62-160$), affinity purified from bacterial lysates using glutathione-Sepharose and assayed for catalytic activity by *in vitro* kinase assay. GST-Clk1 $\Delta 62-160$ displayed significant catalytic activity indicating that this mutant contains a functional kinase domain (Fig. 3).

The presence of kinase activity in mutant $\Delta 62-129$ and the lack thereof in $\Delta 62-160$ suggested that a domain necessary for autophosphorylation is found between amino acids 130-160. We, therefore, created a mutant which lacked this domain ($\Delta 130-158$) and expressed it in COS-1 cells. Immunoprecipitated M-Clk1 $\Delta 130-158$ displayed no autophosphorylation by *in vitro* kinase assay (Fig. 2A, lane 7) despite high protein expression levels (Fig. 2B, lane 3). To confirm the presence of a functional kinase domain $\Delta 130-158$ was also expressed in bacteria as a GST fusion protein (GST-Clk1 $\Delta 130-158$). Similar to GST-Clk1 $\Delta 62-160$, GST-Clk1 $\Delta 130-158$ possessed catalytic activity indicating that it contains a functional kinase domain (Fig. 4B, lane 2). These results suggest that amino acids 130-158 are required for the autophosphorylating activity of Clk1 expressed in mammalian cells. Interestingly this stretch of amino acids is encoded by the alternatively spliced exon, exon B (see Chap. 2, Fig. 1).

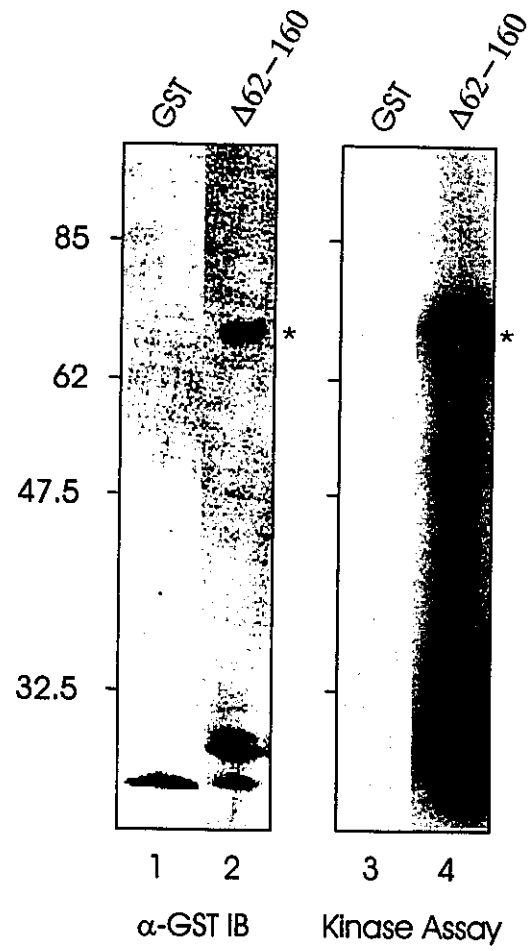
The lack of autophosphorylation of Clk1 $\Delta 130-158$ in mammalian cells may be due to the removal of autophosphorylation sites within that sequence. The primary amino acid

Figure 3

Bacterial expression and catalytic activity of Clk1^{Δ62-160}.

Clk1^{Δ62-160} was expressed in bacteria as a glutathione-S-transferase (GST) fusion protein. GST or GST-Clk1^{Δ62-160} were affinity purified with glutathione-Sepharose 4B and subjected to an *in vitro* kinase assay. Proteins were resolved by SDS-PAGE and transfer to nitrocellulose. **Lanes 1-2**, Immunoblot (IB) of expressed GST and GST-Clk1^{Δ62-160} proteins with the anti-GST monoclonal antibody. **Lanes 3-4**, Phosphorimage of *in vitro* catalytic activities of GST-tagged proteins. GST-Clk1^{Δ62-160} migrates with a molecular mass of approximately 70-kDa (indicated by an asterisk). Positions of molecular mass markers (kDa) are indicated to the left.

Figure 3



sequence of exon B contains 5 serine residues and one tyrosine residue (the codon is split between exons B and C) which may act as sites for autophosphorylation. To determine if the amino acids encoded by exon B could serve as a substrate for the Clk1 catalytic domain we synthesized a peptide encompassing this region (Fig. 4A and see Materials and Methods) and tested for the ability of Clk1 to phosphorylate it *in vitro*. GST-Clk1 expressed in bacteria was affinity purified with glutathione-Sepharose and an *in vitro* kinase assay was performed in the presence of the peptide. As is evident in figure 4B (lanes 4 and 5), GST-Clk1 was able to phosphorylate the peptide. This phosphorylation was specific to catalytically active Clk1 as neither GST alone nor the catalytically inactive point mutant of Clk1 (GST-Clk1^{K190R}) was able to phosphorylate the peptide (Fig. 4B, lanes 1 and 3). The peptide has a predicted molecular mass of approximately 4-kDa but following phosphorylation migrates with a range of mobilities between 4-6-kDa (Fig. 4B). Such mobility differences may be due to phosphorylation of the peptide at multiple sites. Phosphoamino acid analysis of the phosphopeptide (~6-kDa form) revealed that only serine site(s) were phosphorylated (Fig. 4C).

We next tested whether the deletion mutant lacking exon B (GST-Clk1^{Δ130-158}) could phosphorylate the peptide *in trans*. As shown in figure 4B (lane 2), GST-Clk1^{Δ130-158} was able to phosphorylate the peptide indicating that this sequence is not required to be within the primary sequence for its phosphorylation.

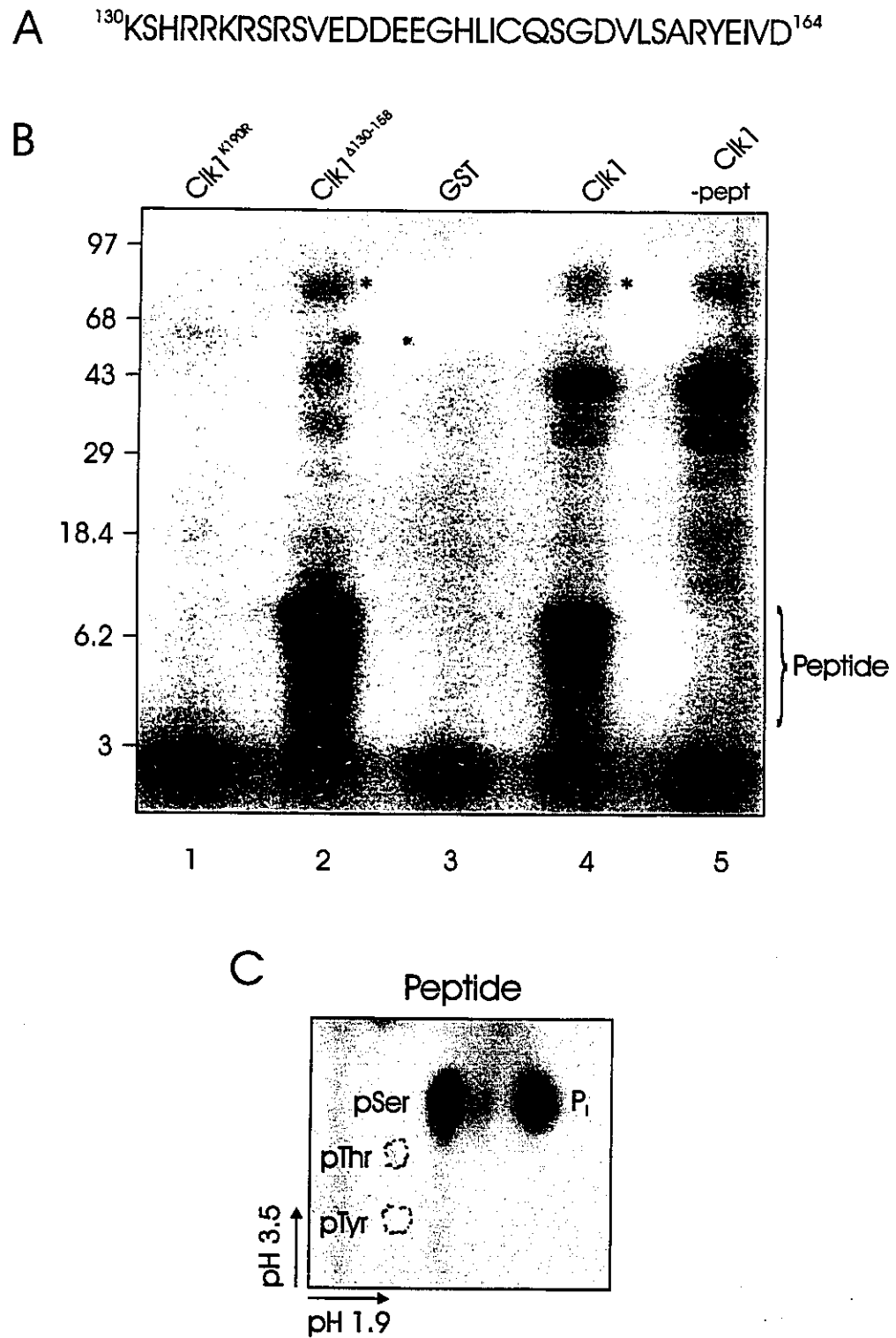
Since Clk1^{Δ130-158} displayed a difference in autophosphorylating activity when expressed either in mammalian cells or bacteria we wished to test the ability of Clk1^{Δ130-158} expressed in COS-1 cells to phosphorylate the peptide. Similar to the results with

Figure 4

Phosphorylation of Clk1 peptide by bacterially expressed Clk1 proteins.

A, Sequence of the synthetic peptide and corresponding amino acid positions in Clk1. **B**, Phosphorimage of *in vitro* catalytic activities of GST-tagged proteins. GST (*lane 3*) and GST-Clk1 fusion proteins (K190R, *lane 1*; Δ 130-158, *lane 2*; and wild-type Clk1, *lanes 4* and *5*) expressed in bacteria were affinity purified using glutathione-Sepharose 4B and subjected to an *in vitro* kinase assay in the presence (*lanes 1-4*) or absence of Clk1 peptide (*lane 5*). Proteins were resolved by 7½-20% SDS-PAGE and transfer to PVDF membrane. The positions of molecular mass markers (kDa) are indicated to the left of the panel. Asterisks mark the positions of full-length GST-Clk1 fusion proteins. **C**, Phosphoamino acid analysis by two-dimensional thin-layer electrophoresis of the peptide phosphorylated by GST-Clk1 (**B**, *lane 4*). Positions of phosphoserine (pSer), phosphothreonine (pThr) and phosphotyrosine (pTyr) standards are indicated. P_i - free inorganic phosphate.

Figure 4



autophosphorylation of the kinases themselves, Clk1^{Δ130-158} expressed in mammalian cells was unable to display significant peptide phosphorylation when compared to the full-length active kinase (Fig. 5A). Peptide phosphorylation by full length Clk1 was on serine residue(s), as seen with GST-Clk1 from bacteria (Fig. 5B).

Tyrosine 160 Phosphorylation

Songyang *et al.* (122) have identified tyrosine 160 in Clk1 as a possible site of phosphorylation. Tyrosine 160 is found within the sequence Y¹⁶⁰EIV, a sequence, which when tyrosine phosphorylated, falls within the consensus sequence for binding the SH2 domain of the fps/fes tyrosine kinases (122). Phosphorylation of tyrosine 160 may be by autophosphorylation or phosphorylation by a different kinase. To determine if tyrosine 160 is a major site of autophosphorylation and/or a regulatory phosphorylation site we created a site-directed mutant of Y160. Two mutants were created, Y160F, to convert tyrosine to the non-phosphorylatable phenylalanine, and Y160D, to convert tyrosine to aspartate, a negatively charged amino acid which may mimic phosphotyrosine. Both mutants were highly expressed when transfected into COS-1 cells (Fig. 6A). Clk1^{Y160F} displayed significant autophosphorylation (Fig. 6B, lane 1) suggesting that phosphorylation at this site is not necessary for activity of the catalytic domain. In contrast, Clk1^{Y160D} exhibited no catalytic activity (Fig. 6B, lane 2) and displayed an increased mobility on SDS-PAGE relative to that of the wild-type kinase. This difference in mobility is similar to that seen with the catalytically inactive mutant Clk1^{K190R} (see Chap. 2). If substitution of tyrosine 160 with aspartate indeed mimics phosphorylation of this residue this result suggests that

Figure 5

Phosphorylation of Clk1 peptide by Clk1 proteins expressed in COS-1 cells.

A, Phosphorimage of *in vitro* catalytic activities of Myc-tagged proteins. COS-1 cells were transfected with pECE vector alone (*lanes 1 and 2*), vector encoding Myc epitope-tagged Clk1, Clk1^{K190R} or Clk1^{Δ130-158} (*lanes 3-4, 5-6, and 7-8, respectively*). Myc-tagged proteins were immunoprecipitated with anti-Myc monoclonal antibody and subjected to an *in vitro* kinase assay in the absence (*lanes 1, 3, 5 and 7*) or presence (*lanes 2, 4, 6 and 8*) of the Clk1 peptide. Proteins were resolved by 15% SDS-PAGE and transfer to PVDF membrane. The positions of molecular mass markers (kDa) are indicated to the left of the panel. **B**, Phosphoamino acid analysis by two-dimensional thin-layer electrophoresis of the peptide phosphorylated by wild-type Clk1. Positions of phosphoserine (pSer), phosphothreonine (pThr) and phosphotyrosine (pTyr) standards are indicated. P_i - free inorganic phosphate.

Figure 5

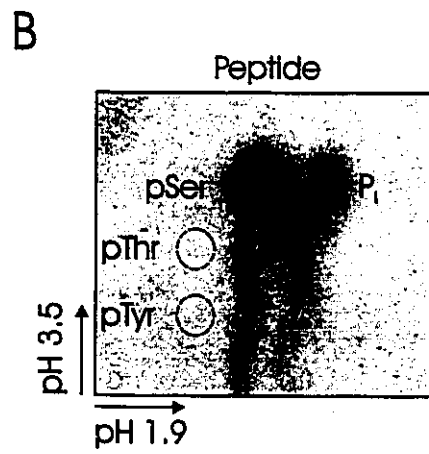
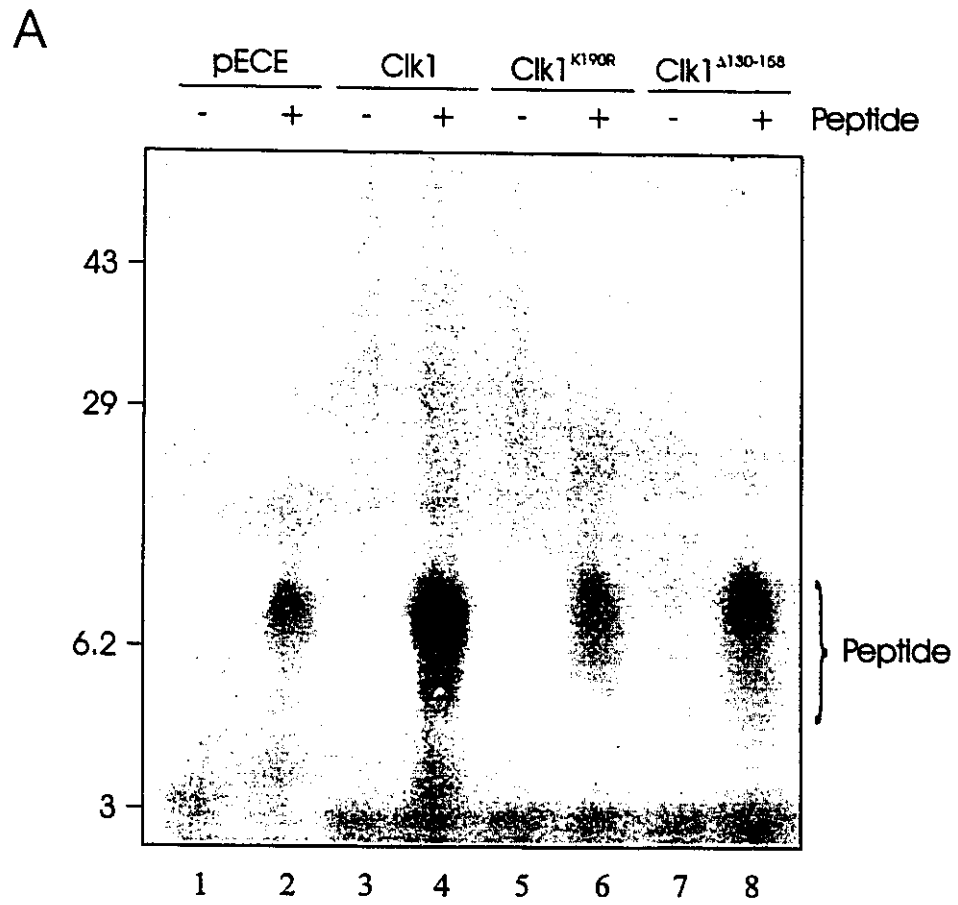
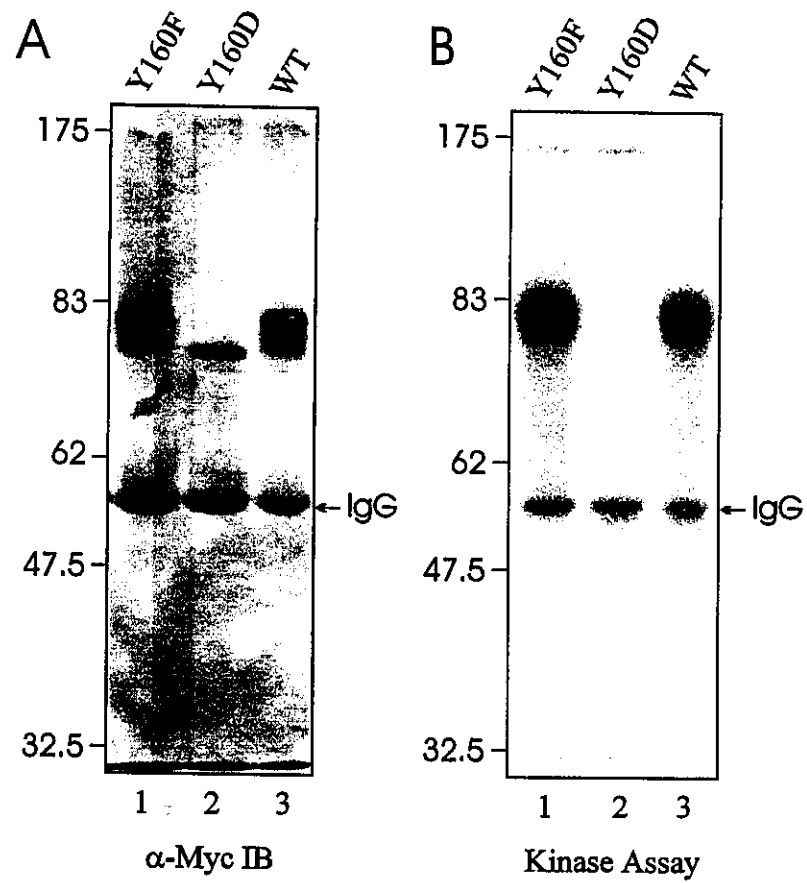


Figure 6

Catalytic activities of Clk1 Y160 mutants in COS-1 cells.

A, Expression of Clk1 Y160 mutants determined by anti-Myc antibody immunoblotting (IB) and enhanced chemiluminescence. **B**, Phosphorimage of *in vitro* catalytic activities of Clk1 Y160 mutants. COS-1 cells were transfected with pECE expression vector encoding Myc epitope-tagged Clk1 Y160F mutant (A,B lane 1), Y160D mutant (A,B lane 2) or wild-type Clk1 (A,B lane 3),. Myc-tagged proteins were immunoprecipitated and subjected to an *in vitro* kinase assay. Proteins were resolved by SDS-PAGE and transfer to nitrocellulose. The positions of molecular mass markers (kDa) are indicated to the left of each panel. IgG - immunoglobulin heavy chain.

Figure 6



phosphorylation of Y160 down-regulates the catalytic activity of Clk1. Sequencing of the catalytic domain of Clk1^{Y160D} indicated that there were no nucleotide substitutions, apart from those at positions 570 and 571, which might explain the lack of catalytic activity. A possible explanation for the lack of activity is that the Y160D causes a conformational change in the kinase, rendering it inactive.

Identification of Clk1:Clk1 Interaction Domains

We have previously shown that Clk1 is able to form a complex with other molecule(s) of Clk1 (Chap. 2, Fig. 8). Truncated Clk1 (Clk1^T) is also able to form a complex with Clk1 indicating that the amino-terminal portion of the molecule contains an interaction domain and that catalytic activity is not required for this interaction.

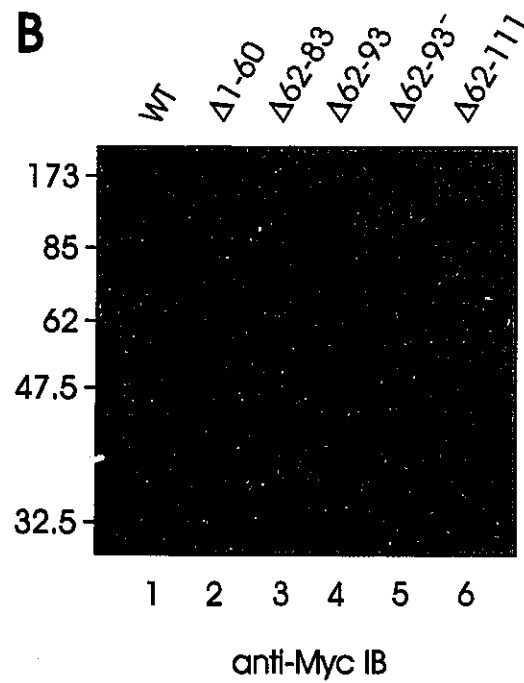
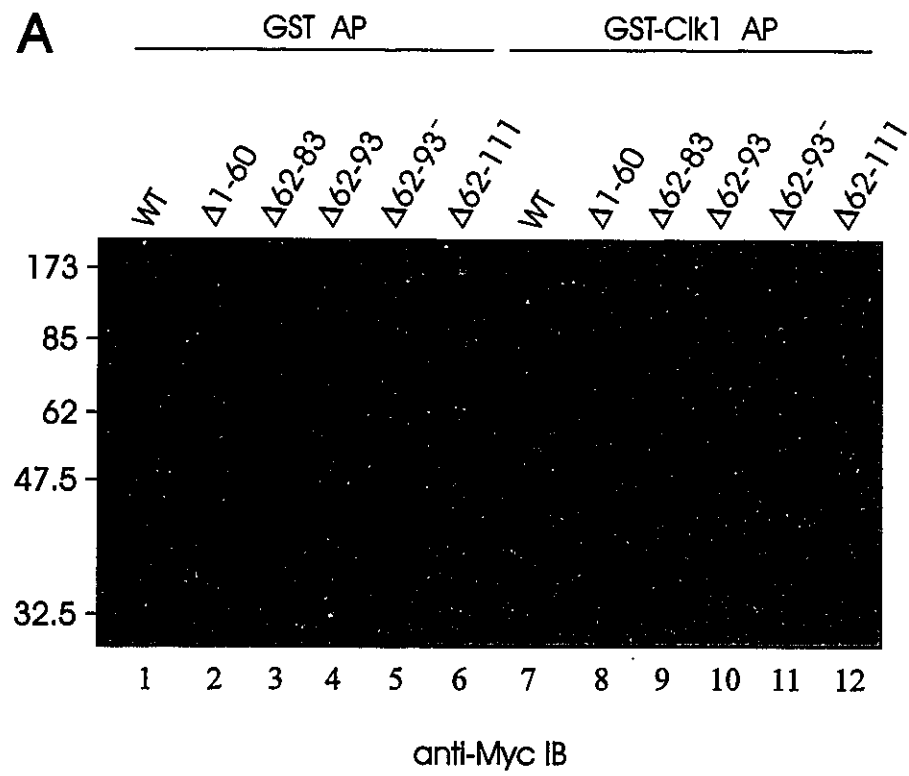
To gain further insight into the biochemical regulation of the Clk1 kinase we wished to determine the domain(s) involved in Clk1:Clk1 binding. We have used the panel of deletion mutants describe above to map these domain(s). Myc-epitope tagged Clk1 deletion mutants expressed in COS-1 cells were incubated *in vitro* with bacterially expressed GST or GST-Clk1 fusion protein coupled to glutathione-Sepharose. Bound proteins were eluted, resolved by SDS-PAGE, transferred to nitrocellulose and immunoblotted with the anti-Myc mAb. None of the Clk1 polypeptides interacted with GST-coupled beads (Fig. 7A, 8A, 9A, lanes 1-6). Full length Clk1 was able to bind immobilized GST-Clk1 as has been previously shown (Fig. 7A, lane 7 and Chap. 2, Fig. 8). To test whether the catalytic domain confers any binding ability for GST-Clk1 the Δ 1-165 mutant was assayed in this system. When expressed in COS-1 cells Δ 1-165 displayed

Figure 7

Binding of Clk1 deletion mutants to GST-Clk1 *in vitro*.

COS-1 cells were transfected with pECE vector encoding Myc epitope-tagged wild-type Clk1 (A, lanes 1 and 7, B, lane 1), or Clk1 deletion mutants $\Delta 1-60$ (A, lanes 2 and 8, B, lane 2), $\Delta 62-83$ (A, lanes 3 and 9, B, lane 3), $\Delta 62-93$ (A, lanes 4 and 10, B, lane 4), $\Delta 62-93$ catalytic mutant (A, lanes 5 and 11, B, lane 5) and $\Delta 62-111$ (A, lanes 6 and 12, B, lane 6). Protein from equivalent fractions of cell lysates was resolved by SDS-PAGE and transfer to nitrocellulose (B) or subjected to affinity purification with bacterially expressed GST coupled to glutathione-Sepharose 4B (GST AP, A, lanes 1-6) or GST-Clk1 coupled to Sepharose (GST-Clk1, A, lanes 7-12) followed by separation using SDS-PAGE and transfer to nitrocellulose. Myc-tagged proteins were visualised by immunoblot analysis (IB) and enhanced chemiluminescence with the anti-Myc antibody. The positions of molecular mass markers (kDa) are indicated to the left of each panel.

Figure 7



undetectable catalytic activity and was unable to form a complex with GST-Clk1 (Fig. 8A, lane 11). In contrast, Clk1^T, although expressed at levels much lower than Δ 1-165, showed considerable binding (compare Fig. 8A, lanes 11,12 and 8B, lanes 5,6). These results indicated that Clk1:Clk1 complex formation is mediated by amino acids found within the amino-terminal regulatory domain.

Deletion of the first sixty amino acids (Δ 1-60) resulted in loss of binding (Fig. 7A, lane 8) despite a reasonable level of expression (Fig. 7B, lane 2). Therefore, a domain within amino acids 1-60 is likely required for binding to GST-Clk1. Deletion of amino acids 62-83 (Δ 62-83) also resulted in a Clk1 polypeptide which was unable to interact with GST-Clk1. All larger carboxyl-terminal deletions also had impaired binding (see Δ 62-93, -111 and -129 in Fig. 7A, lanes 10,12 and Fig. 8A, lane 9, respectively). These polypeptides were all expressed at levels comparable to that of full-length Clk1 (see Δ 62-83, -93, -111 and -129, Fig. 7B and 8B). Thus, Δ 62-83 demarks the smallest deletion with impaired binding. The proximity of this region to Δ 1-60 suggests that there may be a single domain that encompasses the junction of these regions (amino acids 60-62) or that there are at least two domains. This domain(s) is defined here as a positive regulatory domain; one that enhances Clk1:Clk1 binding.

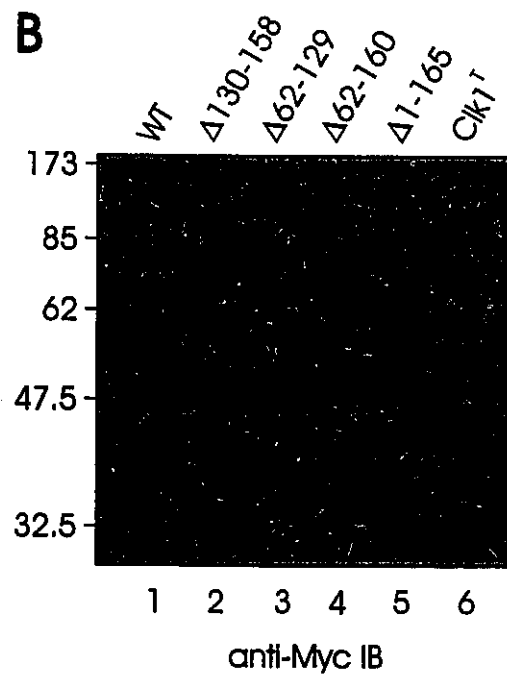
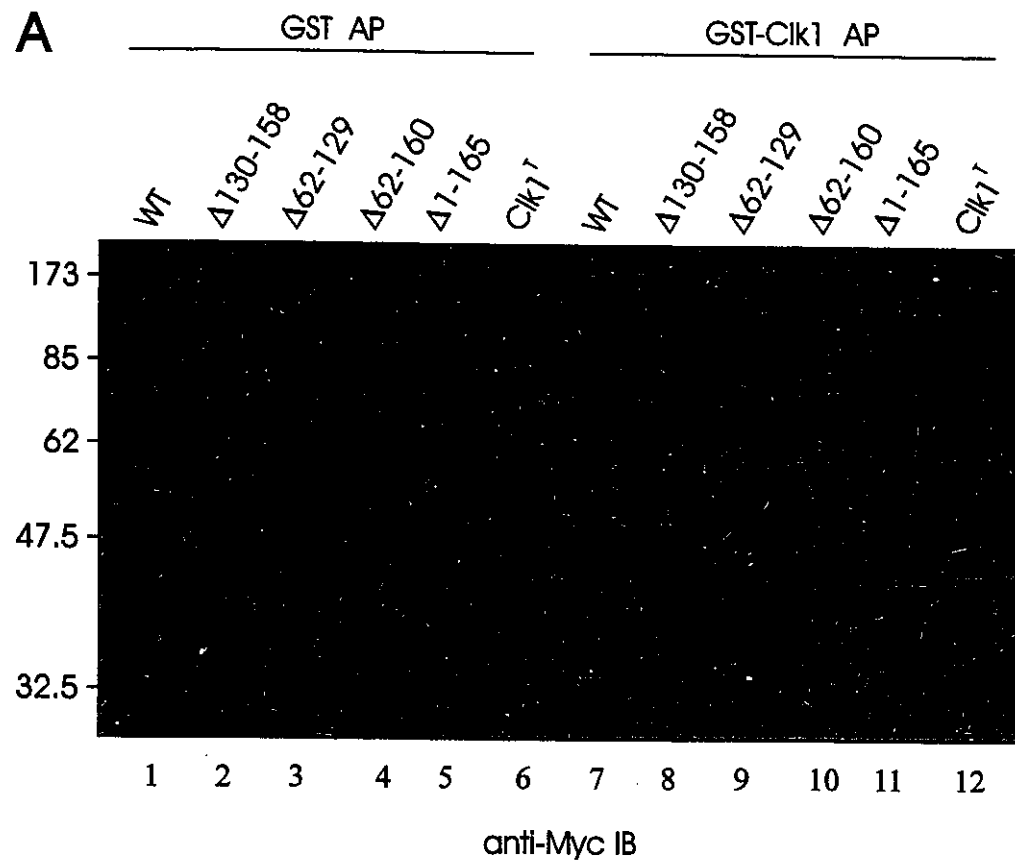
A further deletion, Δ 62-160, which also removed the amino acids encoded by exon B showed considerable binding to GST-Clk1 (Fig. 8A, lane 10). The lack of Δ 62-129 binding and the apparent binding of Δ 62-160 suggested that between amino acids 130-160 resides a domain which inhibits Clk1:Clk1 binding. This domain would be defined as a negative regulatory domain; one that diminishes Clk1:Clk1 binding. An important

Figure 8

Binding of Clk1 deletion mutants to GST-Clk1 *in vitro*.

COS-1 cells were transfected with pECE vector encoding Myc epitope-tagged wild-type Clk1 (A, *lanes 1* and *7*, B, *lane 1*), or Clk1 deletion mutants $\Delta 130-158$ (A, *lanes 2* and *8*, B, *lane 2*), $\Delta 62-129$ (A, *lanes 3* and *9*, B, *lane 3*), $\Delta 62-160$ (A, *lanes 4* and *10*, B, *lane 4*), $\Delta 1-165$ (A, *lanes 5* and *11*, B, *lane 5*) and truncated Clk1 (Clk1^T, A, *lanes 6* and *12*, B, *lane 6*). Protein from equivalent fractions of cell lysates was resolved by SDS-PAGE and transfer to nitrocellulose (B) or subjected to affinity purification with bacterially expressed GST coupled to glutathione-Sepharose 4B (GST AP, A, *lanes 1-6*) or GST-Clk1 coupled to Sepharose (GST-Clk1, A, *lanes 7-12*) followed by separation using SDS-PAGE and transfer to nitrocellulose. Myc-tagged proteins were visualised by immunoblot analysis (IB) and enhanced chemiluminescence with the anti-Myc antibody. The positions of molecular mass markers (kDa) are indicated to the left of each panel.

Figure 8



difference between the deletion mutants $\Delta 62-129$ and $\Delta 62-160$ is that the former displays catalytic activity whereas the latter does not. Thus, the phosphorylation state of the Clk1 deletion polypeptide may contribute to its ability to bind to GST-Clk1 (see below). To test the idea that amino acids 130-160 contain a negative regulatory domain we determined the ability of deletion mutant $\Delta 130-158$ to bind GST-Clk1. Although expressed at a level similar to that of Clk1, Clk1 $\Delta 130-158$ displayed enhanced binding to GST-Clk1 (compare Fig. 8A, lanes 7,8 and 8B, lanes 1,2). Thus, the amino acids encoded by exon B (130-158) contain a domain which, when present in Clk1, inhibit Clk1:Clk1 complex formation. A caveat to this interpretation of the results is that Clk1 $\Delta 130-158$, like Clk1 $\Delta 62-160$ does not display catalytic activity when expressed in mammalian cells. Therefore, as mentioned above the phosphorylation state of the kinase may contribute to its binding ability.

Binding of K190R mutants

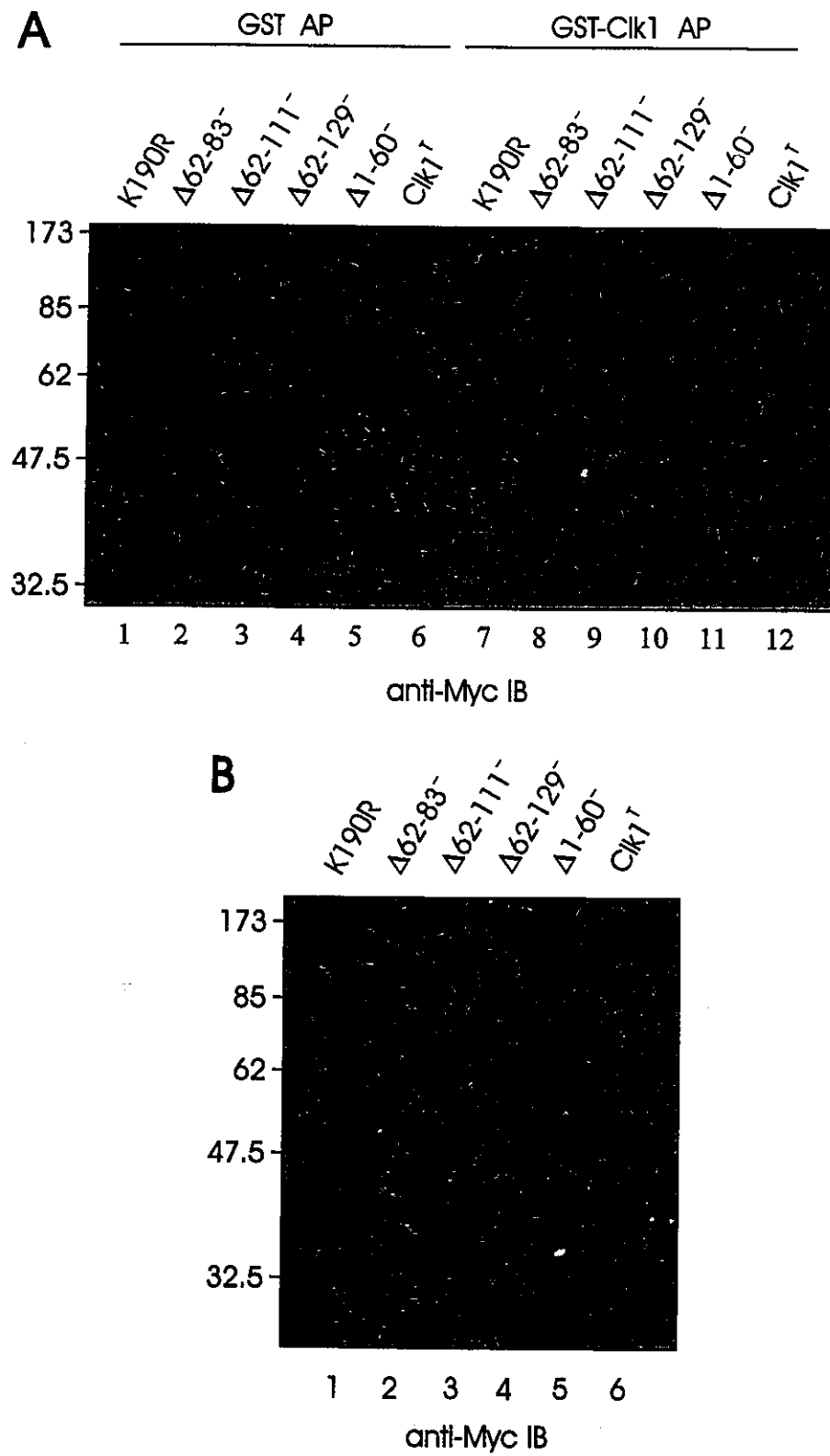
To eliminate the contribution of Clk1 autophosphorylation during these binding studies we have generated a series of deletion mutants containing an inactive catalytic domain. The wild-type catalytic domain was substituted with a catalytic domain carrying the K190R point mutation. In contrast to their catalytically active counterparts, $\Delta 1-60^-$ (superscripted minus sign denotes a K190R catalytic domain), $\Delta 62-83^-$, $\Delta 62-93^-$ and $\Delta 62-111^-$ were able to bind GST-Clk1 (Fig. 9A, lanes 8, 9, 11 and Fig. 7A, lane 11). These results indicated that the positive regulatory domain(s) identified above (between 1-83) is evident only in the presence of a functional kinase domain.

Figure 9

Binding of Clk1^{K190R} deletion mutants to GST-Clk1 *in vitro*.

COS-1 cells were transfected with pECE vector encoding Myc epitope-tagged catalytic mutants of full-length Clk1 (A, *lanes 1 and 7, B, lane 1*), or Clk1 deletion mutants $\Delta 62-83^-$ (A, *lanes 2 and 8, B, lane 2*), $\Delta 62-111^-$ (A, *lanes 3 and 9, B, lane 3*), $\Delta 62-129^-$ (A, *lanes 4 and 10, B, lane 4*), $\Delta 1-60^-$ (A, *lanes 5 and 11, B, lane 5*) and truncated Clk1 (Clk1^T, A, *lanes 6 and 12, B, lane 6*). Protein from equivalent fractions of cell lysates was resolved by SDS-PAGE and transfer to nitrocellulose (B) or subjected to affinity purification with bacterially expressed GST coupled to glutathione-Sepharose 4B (GST AP, A, *lanes 1-6*) or GST-Clk1 coupled to Sepharose (GST-Clk1, A, *lanes 7-12*) followed by separation using SDS-PAGE and transfer to nitrocellulose. Myc-tagged proteins were visualised by immunoblot analysis (IB) and enhanced chemiluminescence with the anti-Myc antibody. Superscripted minus sign denotes the mutant catalytic domain K190R. The positions of molecular mass markers (kDa) are indicated to the left of each panel.

Figure 9



The larger $\Delta 62-129^-$ deletion showed reduced binding to GST-Clk1 compared to Clk1^{K190R}, when expressed at comparable levels (compare Fig. 9A, lanes 7, 10 and 9B, lanes 1,4). The considerable binding of $\Delta 62-111^-$ but reduced binding of $\Delta 62-129^-$ suggests that a domain necessary for binding is contained between amino acids 112-129. To test this idea we created a mutant lacking amino acids 112-129 ($\Delta 112-129^-$). $\Delta 112-129^-$ retained the ability to form a complex with GST-Clk1 (Fig. 10, lane 3) indicating that removal of this sequence is not sufficient to prevent binding.

We next tested the ability of Clk1 lacking two of the putative regulatory domains to bind GST-Clk1. The double mutant $\Delta 1-60/112-129^-$ exhibited much impaired complex formation with GST-Clk1 in comparison to the single mutants (compare Fig. 10, lanes 3-5). Analysis of the polypeptide levels indicated that the reduced binding was not due to low expression levels of $\Delta 1-60/112-129^-$ (Fig. 10, lanes 8-10). The residual binding of $\Delta 1-60/112-129^-$ may be due to the regulatory domain which maps to 62-83 and which is retained in this mutant. We cannot rule out the possibility that the lack of binding of any mutant is not due to conformational changes introduced by the deletion.

Subnuclear “Specklising” Domain

Catalytically inactive Clk1 polypeptides co-localise with SR proteins in nuclear “speckles”. To identify the region(s) necessary for localisation of Clk1 to speckles we expressed Clk1^{K190R} deletion mutants in COS-1 cells and assessed their nuclear localisation by indirect immunofluorescence. We have previously shown that Clk1^T localises to nuclear speckles (Chap. 4) indicating that amino acids 1-129 are sufficient for this compartmentalization.

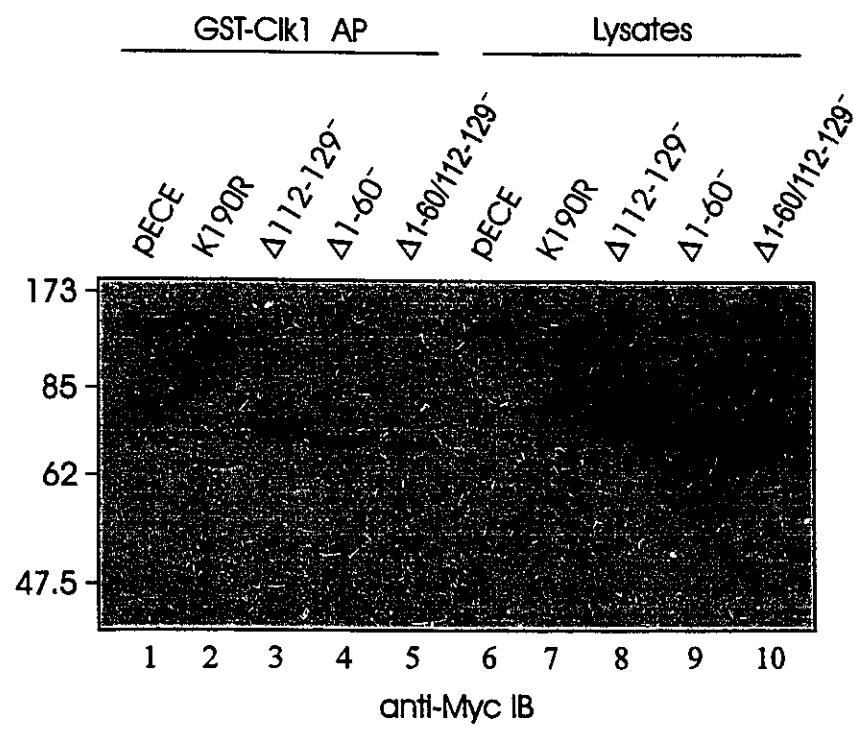


Figure 10

Binding of Clk1^{K190R} deletion mutants to GST-Clk1 *in vitro*.

COS-1 cells were transfected with pECE vector alone (*lanes 1 and 6*) or pECE vector encoding Myc epitope-tagged catalytic mutants of full-length Clk1 (*lanes 2 and 7*), or Clk1 deletion mutants $\Delta 112-129$ (*lanes 3 and 8*), $\Delta 1-60$ (*lanes 4 and 9*) and double mutant $\Delta 1-60/112-129$ (*lanes 5 and 10*). Protein from equivalent fractions of cell lysates was resolved by SDS-PAGE and transfer to nitrocellulose (*Lysates, lanes 6-10*) or subjected to affinity purification with bacterially expressed GST-Clk1 coupled to Sepharose (*GST-Clk1, lanes 1-5*) followed by separation using SDS-PAGE and transfer to nitrocellulose. Myc-tagged proteins were visualised by immunoblot analysis (IB) and enhanced chemiluminescence with the anti-Myc antibody. Superscripted minus sign denotes the mutant catalytic domain K190R. The positions of molecular mass markers (kDa) are indicated to the left of each panel.

Figure 10



Deletion of amino acids between 1-60 or 62-129 had no observable effect on subnuclear localisation (Fig. 12d-f and Fig. 11a-c, respectively). These results suggest that either the deletions were not sufficiently large to remove the domain or that there is more than one domain sufficient to direct Clk1 to speckles and that the deletions did not encompass them.

SR "De-Specklising" Domain

Catalytically active Clk1 causes the redistribution of SR proteins *in vivo* (Chap. 3). To determine the domain(s) required for this activity we have expressed Clk1 deletion mutant in COS-1 cells and determined the subnuclear localization of endogenous SR proteins. Clk1^{Δ1-60}, although widely expressed throughout the cell retained the ability to cause SR protein redistribution (Fig. 12a-c). Deletions from 62-111 had no observable effect on Clk1 activity towards SR proteins whereas Δ62-129 had an intermediary effect with many of the cells expressing this deletion also showing the loss of SR proteins from speckles (Fig. 11d-f). However, there were also cells which displayed a diffuse staining pattern of Clk1 and a speckled pattern of SR proteins (Fig. 11g-i). Deletion of amino acids 130-158 (Clk1^{Δ130-158}), as has been previously shown, resulted in a kinase which could no longer cause the redistribution of SR proteins (Fig. 13a-c and Chap. 4, Fig. 4). Deletion of the entire amino-terminal region of Clk1 (Δ1-165) resulted in a polypeptide which was predominately cytoplasmic and which could no longer cause SR protein redistribution (Fig. 13d,e). These results suggested that amino acids encoded by exon B (amino acids 130-159) are required for the ability of Clk1 to cause the redistribution of SR proteins.

Figure 11

Nuclear localization of Clk1 deletion mutants and their effects on SR protein distribution.

COS-1 cells were transfected with pECE vectors encoding Myc epitope-tagged Clk1^{K190R} deletion mutant $\Delta 62-129^-$ (*panels a-c*), or Clk1 deletions $\Delta 62-129$ (*panels d-i*) or $\Delta 62-160$ (*panels j-l*) and subjected to indirect immunofluorescent staining and confocal microscopy. Confocal images are from the anti-Myc monoclonal antibody (mAb) (α -Myc, *panels a,d,g* and *j*), anti-SR mAb (α -SR, *panels b,e,h* and *k*) or an overlay of the two signals (α -Myc + α -SR, *panels c,f,i* and *l*). Superscripted minus sign denotes the mutant catalytic domain K190R. Alignment of green and red signals appears yellow. Magnification 63X.

Figure 11

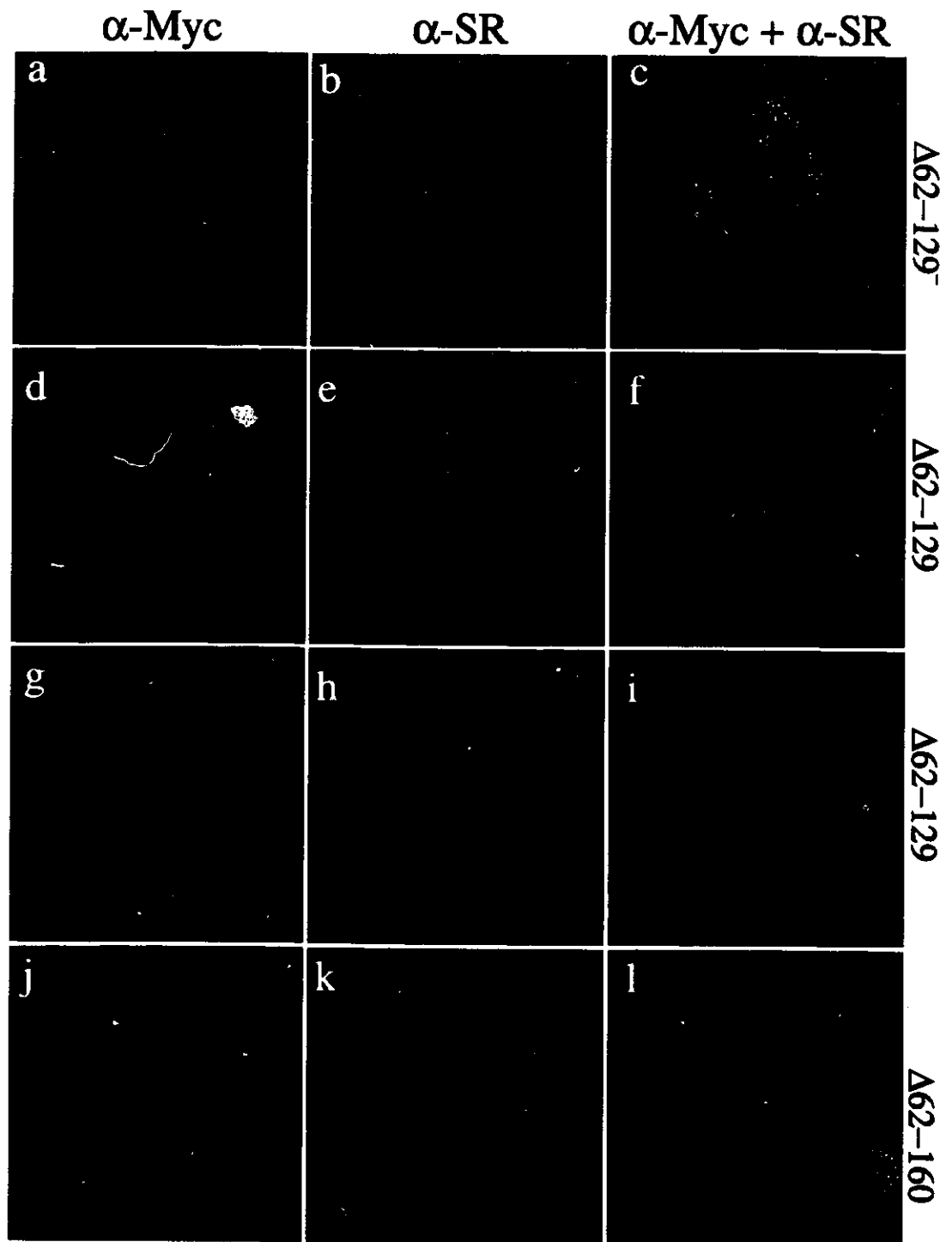


Figure 12

Nuclear localization of Clk1 deletion mutants and their effects on SR protein distribution.

COS-1 cells were transfected with pECE vectors encoding Myc epitope-tagged Clk1 deletion mutants $\Delta 1-60$ (*panels a-c*) and $\Delta 1-60/112-129$ (*panels g-l*) or Clk1^{K190R} deletion mutants $\Delta 1-60^-$ (*panels d-f*) and $\Delta 1-60/112-129^-$ (*panels j-l*) and subjected to indirect immunofluorescent staining and confocal microscopy. Confocal images are from the anti-Myc monoclonal antibody (mAb) (α -Myc, *panels a, d, g* and *j*), anti-SR mAb (α -SR, *panels b, e, h* and *k*) or an overlay of the two signals (α -Myc + α -SR, *panels c, f, i* and *l*). Superscripted minus sign denotes the mutant catalytic domain K190R. Alignment of green and red signals appears yellow. Magnification 40X (*panels j-l*) and 63X (*panels a-l*).

Figure 12

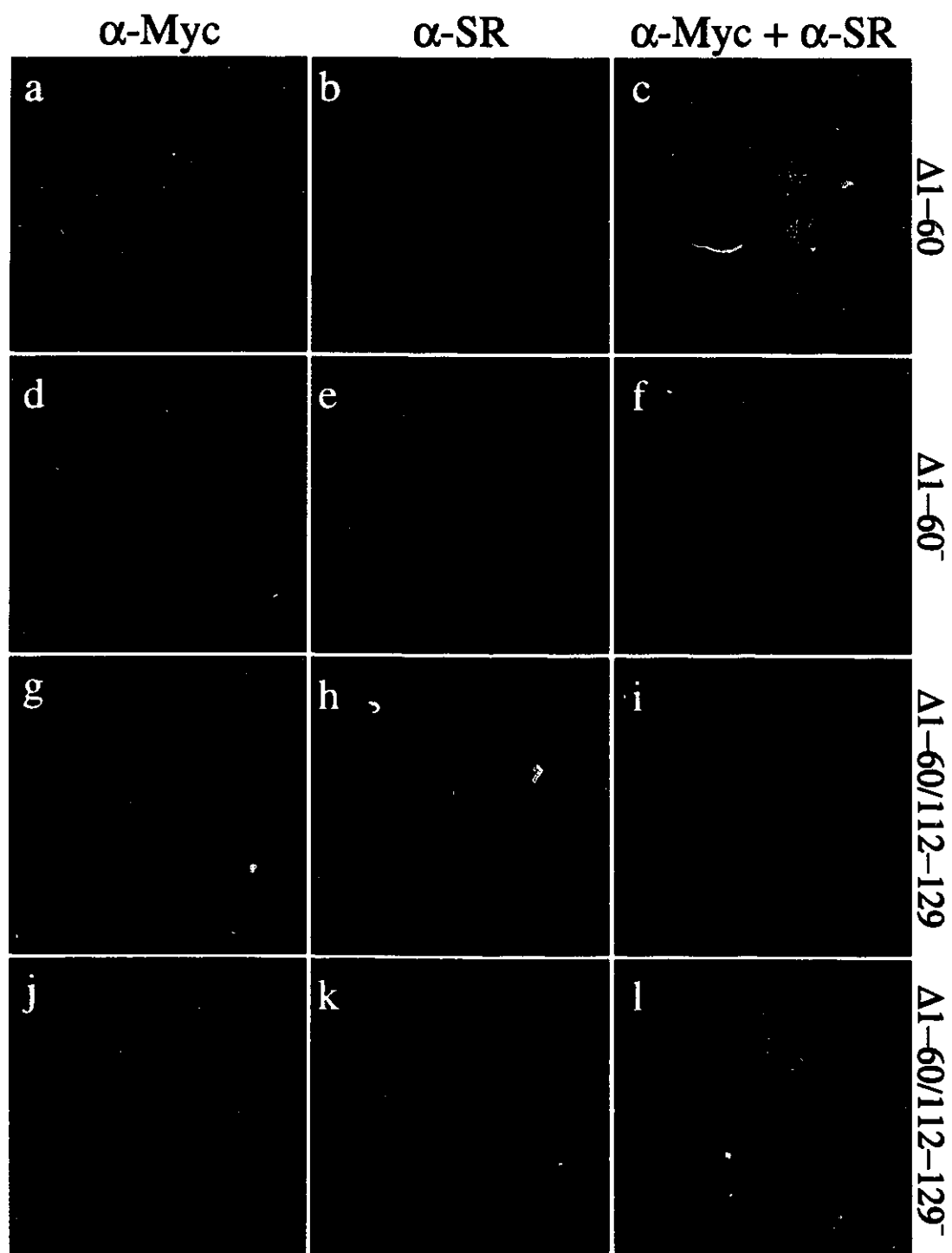
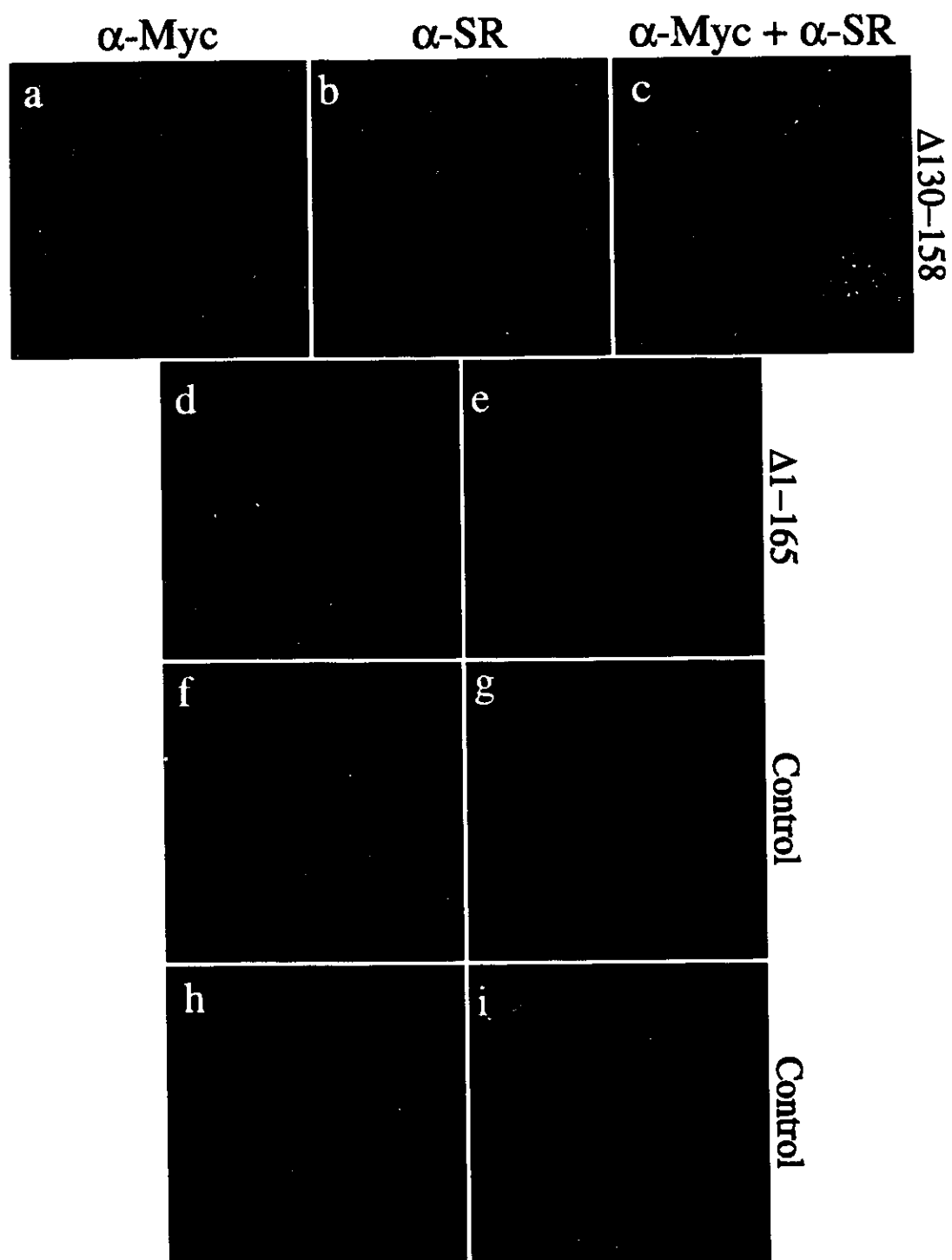


Figure 13

Amino acids encoded by the alternatively spliced exon are necessary for SR protein redistribution.

COS-1 cells were transfected with pECE vector encoding Myc epitope-tagged Clk1 deletion mutant $\Delta 130-158$ (*panels a-c*), $\Delta 1-165$ (*panels d,e*) or wild-type Clk1 (*panels f-I*) and subjected to indirect immunofluorescent staining and confocal microscopy. Cells were double-labelled with anti-Myc monoclonal antibody (mAb) (α -Myc) and anti-SR mAb (α -SR) (*panels a-c*) and visualized by indirect immunofluorescence and confocal microscopy. α -Myc + α -SR (*panel c*) is an overlay of the two signals. Alignment of green and red signals appears yellow. Control samples in *panels f* and *g* lack the Texas Red-coupled anti-IgM secondary antibody specific for the α -SR mAb. Control samples in *panels h* and *i* lack the FITC-coupled anti-IgG secondary antibody specific for the α -Myc mAb. Magnification 63X.

Figure 13



DISCUSSION

Deletions within the amino-terminal non-catalytic region of Clk1 have been used to identify possible regulatory domains. We have identified the amino acids encoded by exon B in Clk1 as important for the regulation of Clk1's catalytic activity (autophosphorylation and SR protein redistribution). Intriguingly, $\Delta 130-158$ can autophosphorylate when expressed in bacteria, suggesting that in COS-1 cells an inhibitor of Clk1 catalytic activity may exist. Although Clk1 is able to phosphorylate serine residue(s) within this region the outcome of this modification is unknown. It remains possible that the phosphorylation state of this domain is inconsequential to Clk1's activity.

Tyrosine 160 (Y160) within the exon B domain was not phosphorylated by Clk1 *in trans*, however, it may be an autophosphorylation site when present *in cis*. Within the peptide, Y160 was located near the carboxyl end, thus the necessary amino acid recognition sequence for phosphorylation by Clk1 at that site may extend beyond the sequence present in the peptide. It is intriguing to speculate that Y160 is recognized by a nuclear tyrosine kinase and when phosphorylated inactivates the Clk kinase. When phosphorylated, Y160 falls within the consensus for binding of the fps/fes tyrosine kinase SH2 domain (122). Little is known about the proteins which interact with these kinases, however, Clk1 and a fps/fes-like kinase may form part of a signalling pathway from the cytoplasm to the nucleus. For instance, the FER kinase has a similar tissue distribution to the Clk kinase, is known to shuttle between the cytoplasm and nucleus and is related to the fps/fes kinases which recognize the Y160 motif (123,124). Y160 is conserved among the

Clk1 family members hClk1, 2 and 3 (21) suggesting that this may be an important residue in Clk signalling.

Receptor tyrosine kinases typically require dimerisation for activation of their catalytic domains (125), however, it is still unclear whether dimerisation plays any role in the activation and signal transduction capabilities of non-receptor kinases. We have identified three domains in Clk1 which regulate its ability to form Clk1:Clk1 complexes *in vitro* (summarized in Fig. 14). Two of these domains are identified in the double deletion mutant $\Delta 1-60/112-129$. This mutant has residual binding when compared to single deletions, indicating that sequences within either deleted region are important for binding. Further deletions within this region will be required to identify the exact sequence(s) involved. $\Delta 1-60/112-129$ was still able to cause the redistribution of SR proteins (Fig. 12g-I) suggesting that dimerisation is not required for this activity.

The third domain maps to the sequence encoded by exon B. Removal of this sequence enhanced Clk1:Clk1 binding. It is important to note that this mutant did not display catalytic activity in COS-1 cells. Thus, the enhanced binding of this mutant may simply be due to the increased charge of the protein relative to that of full-length Clk1.

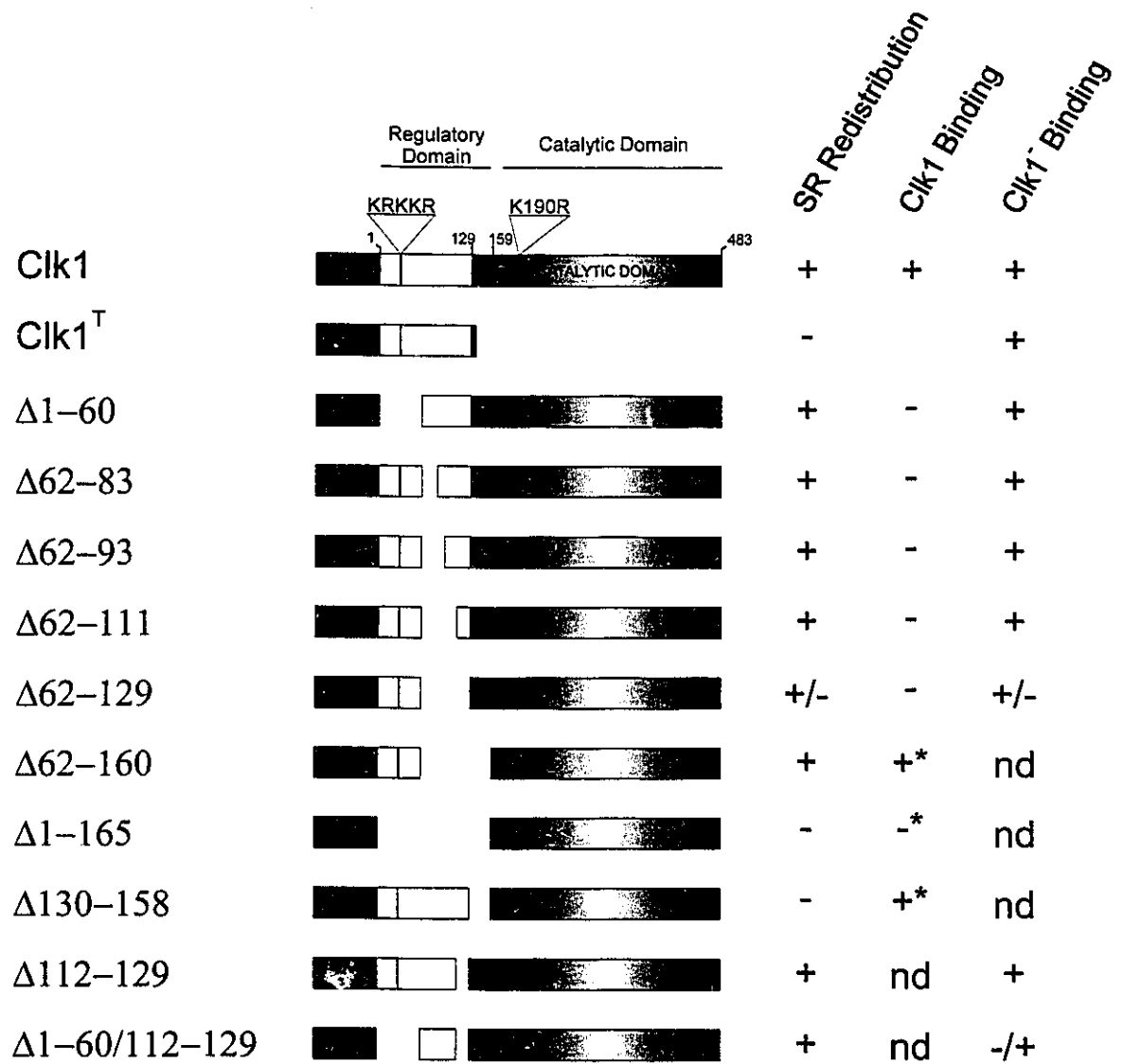
A suggested function of the RS domain in SR proteins is in subnuclear localisation to speckles (38,39). Clk1 itself possesses an amino-terminal non-catalytic domain rich in serine and arginine residues (Chap. 3). However, it contains only 10 SR or RS di-peptides (the longest stretch is RSRS) and, therefore, would not be classified as containing a typical RS domain (126). We were unable to identify a discrete domain which resulted in loss of targetting of Clk1 to nuclear speckles. It is likely, therefore, that several regions

Figure 14

Summary of effects of Clk1 deletions on Clk1 binding and SR protein redistribution.

A summary of the effects of : 1) Clk1 on SR protein redistribution in COS-1 cells, 2) Clk1 deletion mutants on binding to GST-Clk1 (Clk1 Binding), and 3) Clk1^{K190R} deletion mutants on binding to GST-Clk1 (Clk1 Binding). Deletions were created between amino acids 1 and 165 of Clk1 (defined here as the putative regulatory domain). All deletions have the Myc epitope tag fused at the amino terminus of Clk1. The alternatively spliced exon (exon B) is represented in black and constitutes amino acids 130-159. Deletions were generated with a wild-type catalytic domain or a catalytically inactive domain containing the substitution of amino acid lysine 190 for arginine (K190R). MTG - Myc epitope tag; KRKKR - putative nuclear localisation signal. + indicates SR redistribution or Clk1 binding; - indicates no SR redistribution or Clk1 binding; nd - not determined; * - indicates catalytic activity in mammalian cells not demonstrated.

Figure 14



of the polypeptide are involved in the subnuclear targetting of Clk1 and that sequences other than RS domains can also perform this function. The $\Delta 130-158$ mutant was targetted to speckles yet had a dramatically reduced ability to cause SR protein redistribution. This result is consistent with our inability to show catalytic activity of $\Delta 130-158$ in COS-1 cells although it possesses a functional catalytic domain.

hClk2 and hClk3 display a common mode of alternative splicing with that of Clk1 (21) and can cause the redistribution of SR proteins *in vivo* (see Chap. 6). Outside the catalytic domain these kinases show limited sequence identity to Clk1 (21) and no apparent conserved regions. The analogous deletions in Clk2 and Clk3, as performed here, may help to identify functionally similar domains.

It is too early to conclude that Clk1:Clk1 interactions are not an important mode of regulation. Further analysis of the consequences of Clk1 deletions on RNA splicing will aid in elucidating the roles of these regions on Clk1 function.

CHAPTER 6

**The Clk2 and Clk3 Dual-Specificity Protein Kinases Regulate the
Intranuclear Distribution of SR proteins and Influence Pre-mRNA Splicing**

SUMMARY

Three members of the Clk family of kinases (Clk1, 2 and 3) undergo conserved alternative splicing to generate catalytically active (Clk) and inactive (Clk^T) isoforms. The prototype, murine Clk1 (mClk1), is a nuclear dual-specificity kinase that can interact with, and cause the nuclear redistribution of, SR proteins. Human Clk2 and Clk3 (hClk2 and 3) are also found within the nucleus and display dual-specificity kinase activity. The truncated isoforms, hClk2^T and hClk3^T, colocalise with SR proteins in nuclear speckles. Catalytically active hClk2 and hClk3 cause the redistribution of SR proteins and can regulate the alternative splicing of a model precursor mRNA *in vivo*.

INTRODUCTION

Signal transduction in eukaryotes is mediated by networks of proteins, transmitting extracellular stimuli across the plasma membrane, through the cytoplasm and finally into the nucleus. Signalling pathways initiated by such diverse stimuli as mitogens, hormones and stress make use of distinct cytoplasmic kinase cascades to transmit and amplify these signals.

Two distinct MAPK (mitogen-activated protein kinase) pathways have been identified and there is genetic evidence for the existence of others (reviewed in (8,9)). These pathways have as their central activating components a trio of kinases arranged in a cascade. For instance, growth factor stimulation, leading to the activation of Ras and/or PKC, is followed by the sequential phosphorylation and activation of the cytoplasmic MAPK cascade. Thus, Ras and/or PKC can activate the first kinase in the cascade, Raf1,

a MEKK (MAPK or ERK (extracellular signal-regulated kinase) kinase kinase)). Activated Raf1 phosphorylates and activates the second member of the cascade MEK (MAPK/ERK kinase 1 or 2) which in turn phosphorylates and activates the third kinase MAPK (ERK1 or ERK2). Activation of MAPK leads to its nuclear translocation where it can phosphorylate and activate several nuclear proteins, many of which control events resulting in cell proliferation (8,9).

A second pathway includes cytokines and stress as stimuli resulting in the sequential activation of MEKK1, Sek/Jnkk/Mkk4 (the MEK), and Sapk/Jnk (the MAPK) which in turn phosphorylates nuclear targets including the transcription factors c-Jun and TCF-Elk1 (reviewed in (10)). The conserved hierarchy of these cascades is remarkable. The MAPKs, MEKs and MEKKs constitute three families of evolutionarily conserved kinases (8,9). There is genetic evidence for the existence of other distinct pathways, however, the players (kinase cascade members and otherwise) remain to be identified (127,128).

Controlled gene transcription through the regulated activation of transcription factors is generally accepted as critical for normal cellular growth and differentiation and can be mediated in part through MAPK pathways. Equally important control of gene expression occurs post-transcriptionally at the level of RNA processing (25). The RNAs of most transcribed genes undergo capping, splicing, polyadenylation nuclear/cytoplasmic transport, translation and degradation. All of these processes are potential sites for regulation of gene expression and indeed may be regulated by the known MAPK pathways or more likely by distinct kinase cascade signalling pathways.

We and others have identified the Sty/Clk kinase (18,19) which we believe is involved in the regulation of pre-messenger RNA (pre-mRNA) splicing (62) and that we propose is an integral part of a cellular kinase signalling pathway regulating gene expression post-transcriptionally.

Murine Sty/Clk (herein called mClk1) is the prototype of a family of protein kinases which have been conserved throughout evolution (13-24,95). We have shown that the *mClk1* primary transcript is alternatively spliced to generate mRNAs encoding a full-length catalytically active dual-specificity kinase (mClk1) and a truncated, kinase-deficient polypeptide (mClk1^T) (96).

The biological activity of the *mClk1* gene products are largely unknown although possible roles in PC12 differentiation (76) and in the regulation of RNA splicing have been suggested (62). In the latter case, we have shown that mClk1 kinase can interact with and phosphorylate members of the serine/arginine-rich (SR) family of splicing factors (62). Furthermore, overexpression of mClk1 kinase in mammalian cells, initiates the redistribution of nuclear SR proteins suggesting that mClk1 directly regulates the compartmentalization and activity of SR splicing factors.

Additional members of the *Clk* family have been cloned from human cDNA libraries (21,95) and like *mClk1* all three human *Clk* genes express alternatively spliced mRNAs predicted to encoded full-length catalytically active and truncated inactive polypeptides (21).

The identification of related but distinct *Clk* gene products suggests that multiple protein kinases may be involved in the regulation of mRNA splicing. To test this idea we

compared the biochemical and biological properties of human hClk2 and hClk3 gene products. Like mClk1, the full length hClk2 and hClk3 proteins exhibit dual-specificity protein kinase activity *in vitro* and *in vivo* and can effect changes in the subnuclear compartmentalization of SR proteins. Truncated isoforms encoded by all three *Clk* genes co-localise with SR proteins in nuclear speckles. Furthermore, all three Clk kinases regulate the splicing of a mClk1 pre-mRNA. These results suggest the possibility that *Clk* gene products may be involved in the control of constitutive and/or alternative splicing.

MATERIALS AND METHODS

Construction of Plasmids

Human Clk kinase clones (hClk; Dr. K.H. Scheit) were subcloned into the pECE/*M-Sty* expression vector (96) by substitution of the *Sty* fragment for *hClk* fragments. To facilitate subcloning the *Bam*HI restriction site found within the 3' SV40 polyadenylation sequence was destroyed. This was achieved by a partial *Bam*HI digestion of pECE/*M-Sty* followed by blunt ending with T4 DNA polymerase. The linear ~4.7-kb pECE/*M-Sty* fragment was gel-purified and religated to generate pECE/*M-Sty* (Δ *Bam*HI).

The *hClk* fragments (entire coding sequence) were generated by PCR amplification using the following oligonucleotides: *hClk2* and *hClk2^T* : 5'cgggatccatgcccgcctcgaagg^{3'} (including a *Bam*HI site; initiating methionine in bold) and 5'tgtctagatcaccgactgatatcccg^{3'} (including a *Xba*I site; termination codon in bold). *hClk3* and *hClk3^T* : 5'cgggatccatgcatcactgtaagcga^{3'} (including a *Bam*HI site; initiating methionine in bold) and 5'tgtctagatcatctgcttggttgcg^{3'} (including a *Xba*I site; termination codon in bold). 50 μ l amplification reactions were performed containing 100ng plasmid, 25pmol each oligonucleotide, 0.2mM each dNTP, 1 unit Vent DNA Polymerase (New England Biolabs) in buffer supplied by the manufacturer. Amplification conditions were 94°C, 30sec; 60°C, 30sec; 72°C, 90sec for 20 cycles. The reaction products were digested with *Bam*HI and *Xba*I, gel-purified and substituted for the *Sty Bam*HI/*Xba*I fragment in pECE/*M-Sty* (Δ *Bam*HI).

The Myc epitope was confirmed to be in frame by sequencing and is immunoreactive with the mouse monoclonal antibody (mAb) 9E10 in immunoprecipitation/immunoblot analysis (82).

Cell Culture and Transfection

Monkey COS-1 (110) cells were maintained in α -minimal essential medium (α -MEM) supplemented with 10% calf serum and were transiently transfected by electroporation. Trypsinized cells were counted and resuspended at $2-3 \times 10^6$ cells in 0.5ml serum-free α -MEM. 10 μ g of plasmid DNA was added and the cells chilled on ice for 10min. Cells were electroporated using a Gene Pulser (Bio-Rad) at 220V and 960 μ F following which the cells were left at room temperature for 15 min. Transfectants were plated in α -MEM containing 10% calf serum. In experiments to detect the presence of phosphotyrosine transfected cells were grown in the absence or presence of 50 μ M sodium vanadate. Cells were harvested approximately 24h post-transfection.

For splicing analysis COS-1 cells were transfected using Lipofectamine (Gibco BRL) according to the manufacturer's protocol. Cells in 60mm dishes (60-75% confluent) were transfected with 2.5 μ g CMV/*CR-1* reporter gene DNA (129) and 0.5, 2.5 or 5 μ g pECE/*hClk* DNA. pECE DNA was used as filler to keep the total amount of DNA constant. Cells were harvested after 24h in PBS. Protein extracts were generated from one quarter of the sample by direct lysing of the cells in sample buffer (130). Total RNA was extracted from the remaining sample using RNeasy purification (Qiagen) as per the manufacturer's instructions.

Immunoprecipitation/Immunoblot Analysis

Transfected COS-1 cells were lysed in immunoprecipitation buffer (20mM Tris-HCl, pH 7.5, 150mM NaCl, 2mM EDTA, 1% Triton X-100, 2mM NaF, 2mM sodium pyrophosphate, 500 μ M sodium vanadate, 200 μ g/ml phenylmethylsulfonyl fluoride, 2 μ g/ml aprotinin, 5 μ g/ml leupeptin) for 30min on ice. Lysates were cleared by centrifugation. Supernatants were either loaded directly on gels (lysates) or immunoprecipitated with either mAb 9E10 or anti-phosphotyrosine mAb 4G10 (Upstate Biotechnology Inc.). Immunoprecipitates were assayed for kinase activity in 20mM HEPES, pH 7.1, 10mM MgCl₂, 2mM MnCl₂ containing 5 μ Ci [γ ³²P]ATP for 30min at room temperature. Reactions were arrested by the addition of sample buffer followed by boiling. Samples were resolved by SDS-polyacrylamide gel electrophoresis (PAGE) followed by exposure of the gel to phosphor screen and phosphorimager analysis (Molecular Dynamics) or transfer to nitrocellulose membrane for immunoblot analysis.

For immunoblots, membranes were blocked in 5% skim-milk in TBST (150mM NaCl, 10mM Tris-HCl, pH 7.5, 0.05% Tween 20). Anti-Myc immunblotting was performed using 9E10 hybridoma culture supernatant at 1:100. Immunoblots were visualized with horseradish peroxidase-conjugated goat anti-mouse antibody followed by enhanced chemiluminescence and exposure to Reflection film (Dupont NEN).

Phosphoamino acid analysis

Samples were resolved by SDS-PAGE and transfer to PVDF membrane. Following autoradiographic exposure ³²P-labelled bands were excised, acid hydrolysed and subjected

to two-dimensional electrophoresis in the presence of non-radioactive phosphoamino acid standards (4).

Immunofluorescence Assay

Transfected COS-1 cells were grown on gelatin coated (0.1 % gelatin in PBS) coverslips. Cells were fixed and probed as previously described (96). For double staining by indirect immunofluorescence, fixed cells were incubated with mouse mAb 9E10 (α -Myc (IgG), 1:100 hybridoma supernatant) and mouse mAb 104 (α -SR (IgM), 1:2 hybridoma supernatant (32)) for 1h at 37°C, followed by secondary fluorescein-conjugated goat anti-mouse IgG (1:100, Jackson ImmunoResearch Laboratories) specific for mouse IgG and biotin-labelled goat anti-mouse IgM (1:100, Amersham) for 1h at 37°C. Subsequently cells were incubated with streptavidin coupled to Texas Red (1:100, Amersham) for 1h at 37°C. Cells were viewed by confocal microscopy using an upright Leica Confocal Laser Scanning Microscope equipped with a 55 mW krypton/argon air-cooled laser and 40X or 63X (Plan Apo) oil immersion lens.

RT-PCR Analysis

100ng total RNA was used RT-PCR analysis with the following oligonucleotides to amplify *mClk1*-specific products: 5'tggtaggagtgaagaag^{3'} (corresponding to positions 397-414 in the *mClk1* cDNA (18)) and 5'gatggctggcaactagaa^{3'}, directed against the Bovine Growth Hormone polyadenylation signal in the CMV/*CR-1* vector. PCR conditions, including number of cycles and template concentrations, were optimized to maintain

linearity during amplification. Linearity was monitored by serial dilution of the template cDNAs, demonstrating a corresponding decrease in signal, as well as consistent ratios between bands of interest (EB^+/EB^-). PCR products were separated on 1.5% agarose gels, post-stained with SYBR Green 1 (Molecular Probes) and visualised and quantitated with Storm fluorescence imager and ImageQuaNT software (Molecular Dynamics).

The corresponding immunoblot was probed with anti-Myc 9E10 (82) monoclonal antibody and visualised with horseradish peroxidase-conjugated anti-mouse antibody and chemiluminescence.

RESULTS

Expression of hClk2 and hClk3 Proteins in Mammalian Cells

The hClk2 cDNA (also known as PSK-G1 (95)) is predicted to encode a 60.1 kDa protein, whereas the hClk2^T cDNA (also known as hClk2₁₃₉ (21)) would encode a 17.6 kDa truncated polypeptide lacking the catalytic domain (21). Similarly, the hClk3 and hClk3^T (also known as hClk3₁₅₂ (21)) cDNAs are predicted to encode catalytically active and catalytically inactive polypeptides of 58.6 and 19.0 kDa, respectively (21). To detect the expression of these protein products we have added a six-repeat human Myc epitope (60) to each of the hClk cDNAs and expressed them in COS-1 cells. The epitope tag adds 89 amino acids to the amino-terminal end of each polypeptide, increasing the molecular mass by approximately 10-kDa. COS-1 cells transfected with the pECE expression vector alone expressed no proteins detectable by anti-Myc immunoblotting (Fig. 1A, lane 1). COS-1 cells transfected with the pECE vector harbouring Myc-tagged hClk2 or hClk3 cDNAs expressed anti-Myc immunoreactive proteins of approximately 80 and 75-kDa, respectively (Fig. 1A, lanes 2 and 3). Cells transfected with Myc-tagged hClk2^T or hClk3^T expression vectors had anti-Myc immunoreactive proteins of approximately 34 and 35-kDa, respectively (Fig. 1A, lanes 4 and 5).

hClk2 and hClk3 Proteins Exhibit Dual-Specificity Kinase

Activity in Vitro and in Vivo

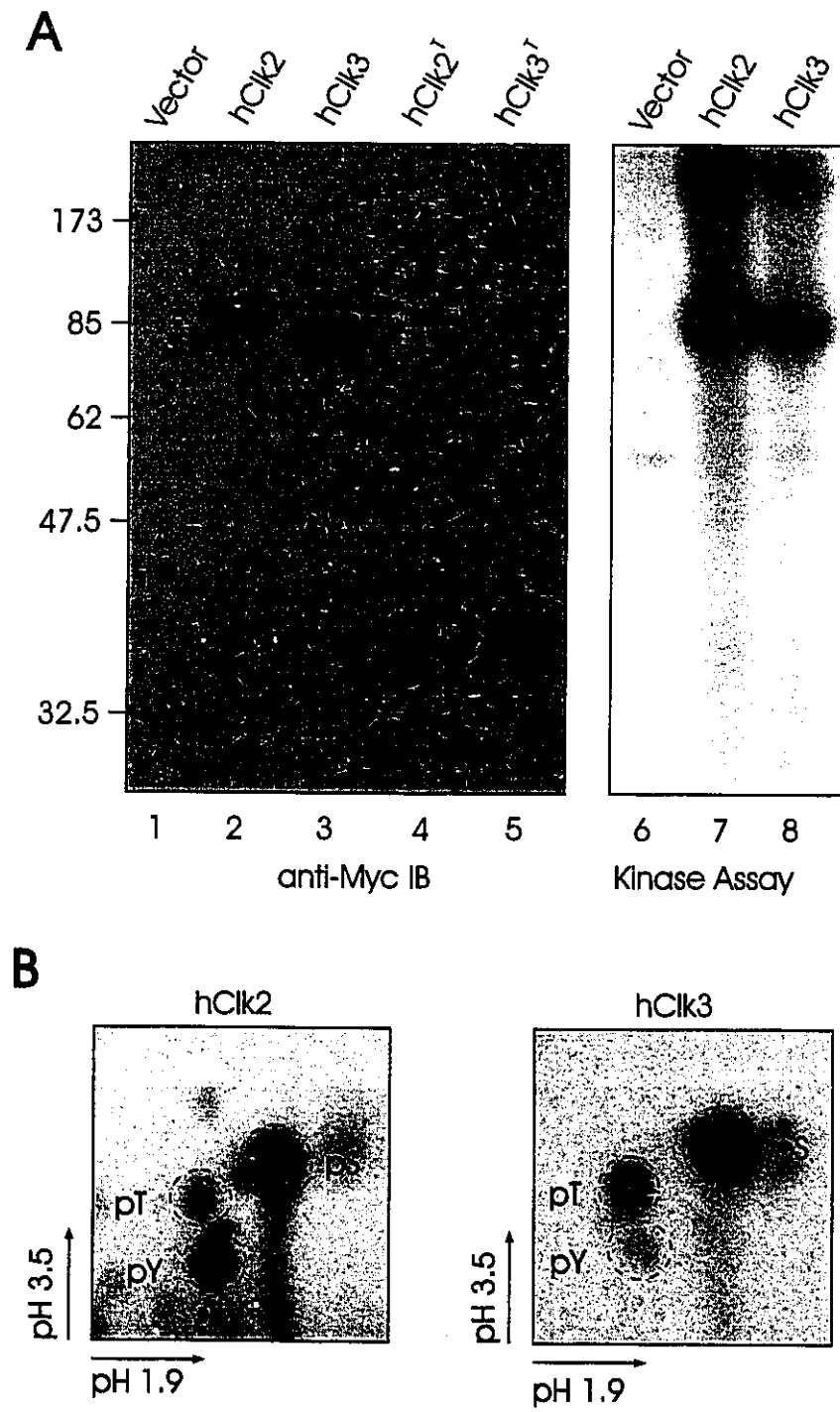
mClk1 expressed in mammalian cells has dual-specificity kinase activity (96). The high sequence similarity within the catalytic domains of mClk1, hClk2 and hClk3

Figure 1

Expression and catalytic activity of hClk2 and hClk3 proteins in COS-1 cells

COS-1 cells were transfected with pECE expression vector alone (*lanes 1,6*) or vector encoding Myc epitope-tagged hClk2 (*lanes 2,7*), hClk3 (*lanes 3,8*), hClk2T (*lane 4*) or hClk3T (*lane 5*). Cells were lysed in immunoprecipitation buffer (see Materials and Methods) and Myc-tagged proteins immunoprecipitated with the anti-Myc monoclonal antibody and subjected to an *in vitro* kinase assay. Proteins were resolved by SDS-PAGE and transfer to nitrocellulose. Protein expression levels were determined by immunoblotting (IB) aliquots of cell lysates with the anti-Myc antibody and enhanced chemiluminescence (A, *lanes 1-5*). Phosphorimage of *in vitro* catalytic activities of hClk2 and hClk3 proteins (A, *lanes 6-8*). Phosphoamino acid analysis by two-dimensional thin layer electrophoresis of hClk2 and hClk3 labelled *in vitro* (B). Positions of molecular mass markers (kDa) are indicated to the left of the panel. The positions of phosphoserine (pS), phosphothreonine (pT) and phosphotyrosine (pY) standards are also indicated.

Figure 1



prompted us to determine whether these two new kinases displayed similar dual-specificity activity. hClk2 and hClk3 proteins expressed in COS-1 cells were immunoprecipitated and subjected to an immune complex kinase assay as described in Materials and Methods. As can be seen in figure 1A, lanes 7 and 8, phosphorylated bands of the expected sizes (80 and 75 kd respectively) are detected and under these assay conditions, hClk2 displays higher *in vitro* kinase activity than hClk3 despite similar levels of expression (compare Fig. 1A, lanes 2 and 3 with lanes 7 and 8). Phosphoamino acid analysis of the phosphorylated bands in figure 1A revealed that both hClk2 and hClk3 had dual specificity protein kinase activity *in vitro* (Fig. 1B).

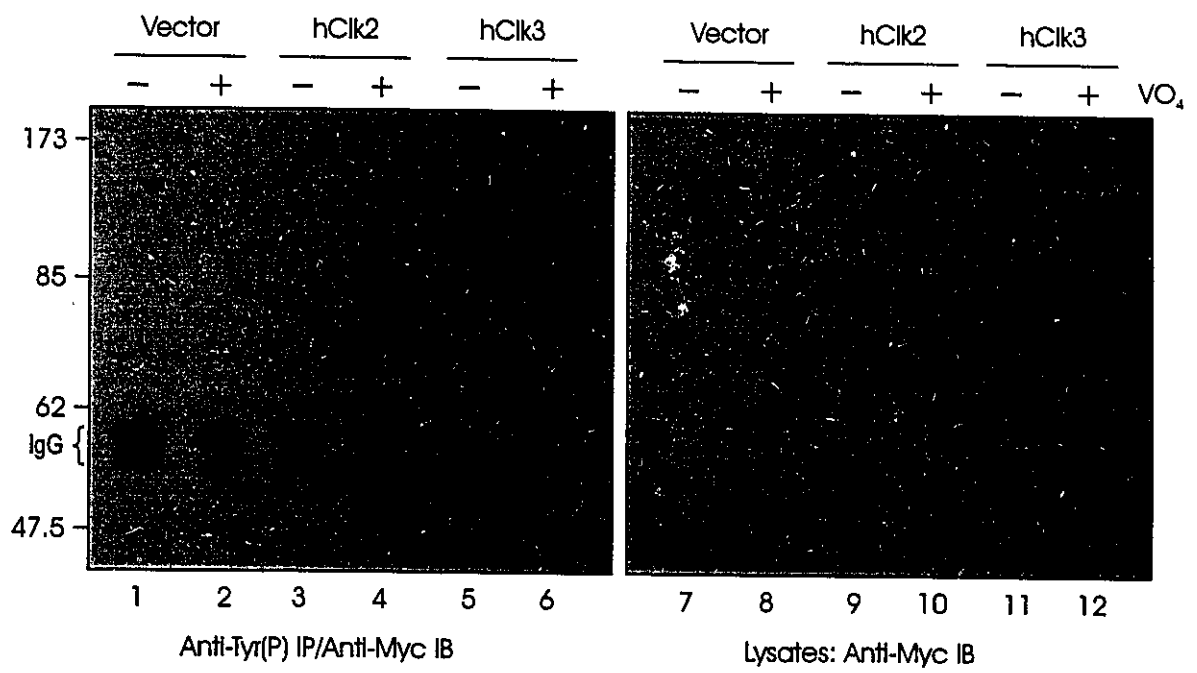
We next determined whether the Clk kinases were phosphorylated on tyrosine residues *in vivo*. Extracts prepared from transfected COS-1 cells were immunoprecipitated with antibodies directed against phosphotyrosine as described in Materials and Methods. Immunoprecipitates were resolved on SDS polyacrylamide gels and immunoblotted with an anti-Myc antibody (82). When cells were grown in the absence or presence of vanadate (a phosphotyrosine phosphatase inhibitor) similar levels of hClk2 (Fig. 2, lanes 9 and 10) or hClk3 (Fig. 2, lanes 11 and 12) protein were detected, however, only in the presence of vanadate were the two kinases significantly phosphorylated on tyrosine residues (Fig. 2, lanes 3-6). As in the *in vitro* kinase assay described above, hClk2 seems to have significantly more tyrosine phosphorylating activity *in vivo* than hClk3.

Figure 2

hClk2 and hClk3 are phosphorylated on tyrosine in vivo.

COS-1 cells were transfected with pECE expression vector alone (*lanes 1,2,7,8*), vector encoding hClk2 (*lanes 3,4,9,10*) or hClk3 (*lanes 5,6,11,12*) and grown in the absence (-, *lanes 1,3,5,7,9,11*) or presence (+, *lanes 2,4,6,8,10,12*) of 50 μ M sodium vanadate (VO_4). Cells were lysed in immunoprecipitation buffer (see Materials and Methods) and Myc-tagged proteins immunoprecipitated (IP) with the anti-phosphotyrosine monoclonal antibody (*Tyr(P)*, *lanes 1-6*). Proteins were resolved by SDS-PAGE and transfer to nitrocellulose. Expression levels were determined from aliquots of cell lysates (*lanes 7-12*). Proteins were visualized by immunoblot analysis (IB) and enhanced chemiluminescence with the anti-Myc antibody. Positions of molecular mass markers (kDa) are indicated to the left of the panel.

Figure 2



Truncated hClk proteins co-localise with splicing factors

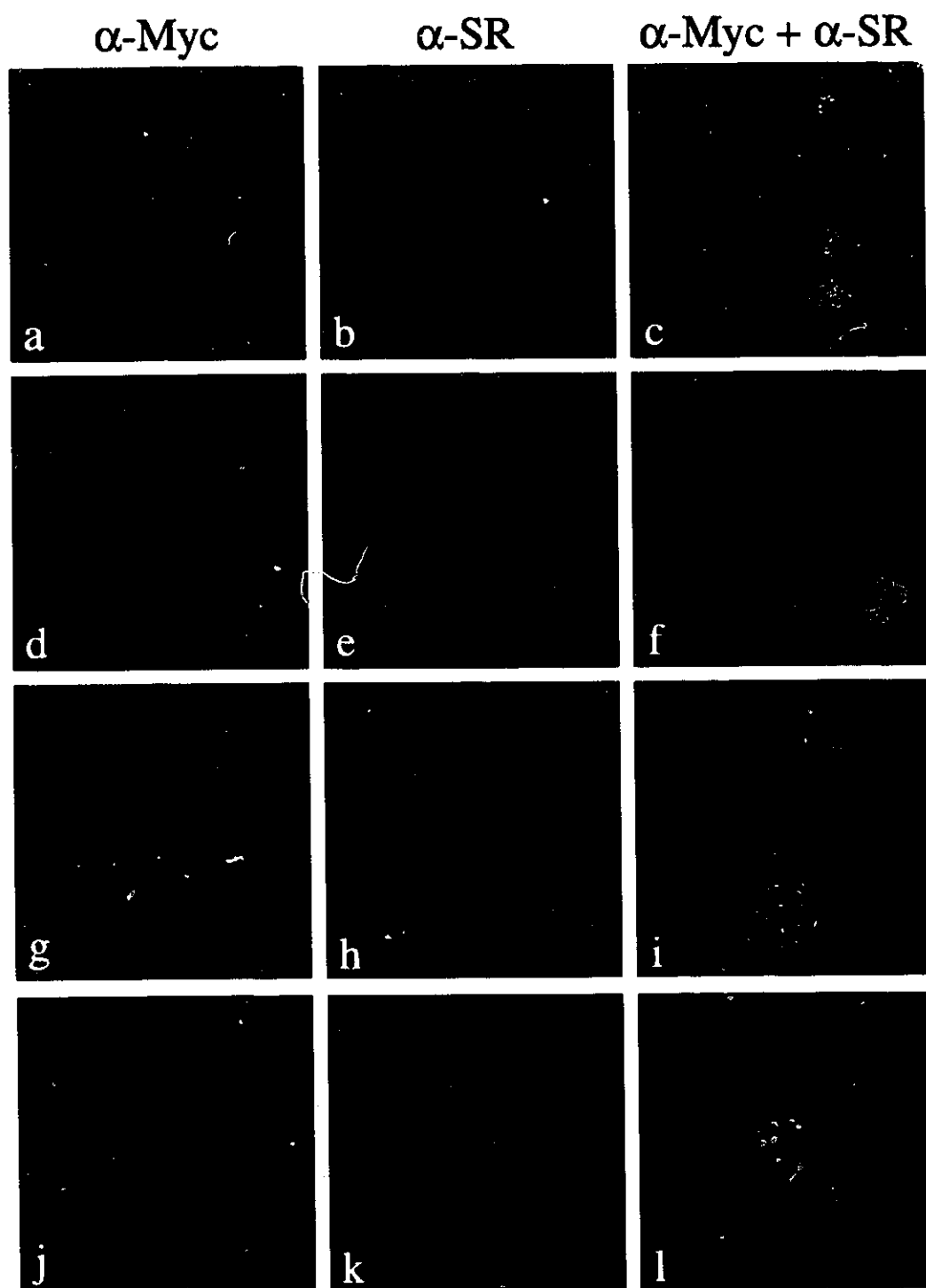
The subcellular localisation of *hClk2* and *hClk3* gene products in transfected COS-1 cells was determined by indirect immunofluorescence. Potential nuclear localization signals have been predicted from the primary amino acid sequence of the hClk proteins (21) and indeed all four hClk isoforms are found in the nuclear compartment (Fig. 3 and 4). The truncated isoforms (*hClk2^T* and *hClk3^T*) had a punctate staining pattern reminiscent of nuclear “speckle” structures that are believed to be stores of mRNA splicing factors in interphase nuclei (Fig. 3 and 4, *panel g*, respectively) (70,71). We have shown previously that the mouse Clk1 kinase can associate with members of the SR protein family (33) and can phosphorylate one of them, ASF/SF2, *in vitro* (62). We therefore tested whether the speckle structures which contain the *hClk2^T* and *hClk3^T* polypeptides also contained SR proteins. To this end, transfected cells were doubly stained with an antibody directed against the Myc tag (to identify *hClk2^T* or *hClk3^T*) and a second monoclonal antibody (mAb 104) directed towards an epitope shared by the SR protein family (32). As can be seen in figures 3 and 4 both *hClk2^T* and *hClk3^T* proteins co-localise with SR proteins in nuclear speckles (Fig. 3 and 4, *panels g-i*, respectively). In a subset of cells *hClk2^T* is found in large nuclear inclusions distinct from speckles (Fig. 3, *panels j-l*). The SR proteins within such cells are no longer found in discrete speckles but are co-localised with *hClk2^T*. The nature and consequence of these nuclear structures is unknown at present.

Figure 3

Nuclear Localization of hClk2 and hClk2^T and their effects on the distribution of SR proteins.

Indirect immunofluorescent staining and confocal microscopy of transfected COS-1 cells using the anti-Myc 9E10 monoclonal antibody (mAb) (α -Myc, *panels a,d,g,j*), anti-SR 104 mAb (α -SR, *panels b,e,h,k*) or an overlay of the two signals (α -Myc + α -SR, *panels c,f,i,l*). Alignment of green and red signals appears yellow. Cells were transfected pECE expression vector encoding hClk2 (*panels a-f*) or hClk2^T (*panels g-l*). Magnification 40X (*panels d-f*); all others at 63X.

Figure 3



The hClk kinases initiate the redistribution of SR proteins

The full-length hClk2 and hClk3 kinases had a diffuse nucleoplasmic staining pattern reminiscent of that seen in cells transfected with the related mouse kinase mClk1 (Fig. 3 and 4, *panels a* and *d*, respectively) (62). We have shown that the mClk1 kinase, by virtue of its catalytic activity causes the re-organization of SR proteins in transfected cells (62). Similarly, hClk2 and hClk3 kinases when overexpressed initiated the redistribution of SR proteins from speckles to a diffuse nucleoplasmic staining (Fig. 3 and 4, *panels a-f*, respectively). hClk3 appeared to have a reduced ability to cause the redistribution of endogenous SR proteins when compared to hClk2 again paralleling the differences we observed in the kinase activity of the two enzymes (compare cells expressing low levels of hClk2 and hClk3 in Fig. 3 and 4, *panels d-f*, respectively).

Clk2 and Clk3 regulate splicing *in vivo*

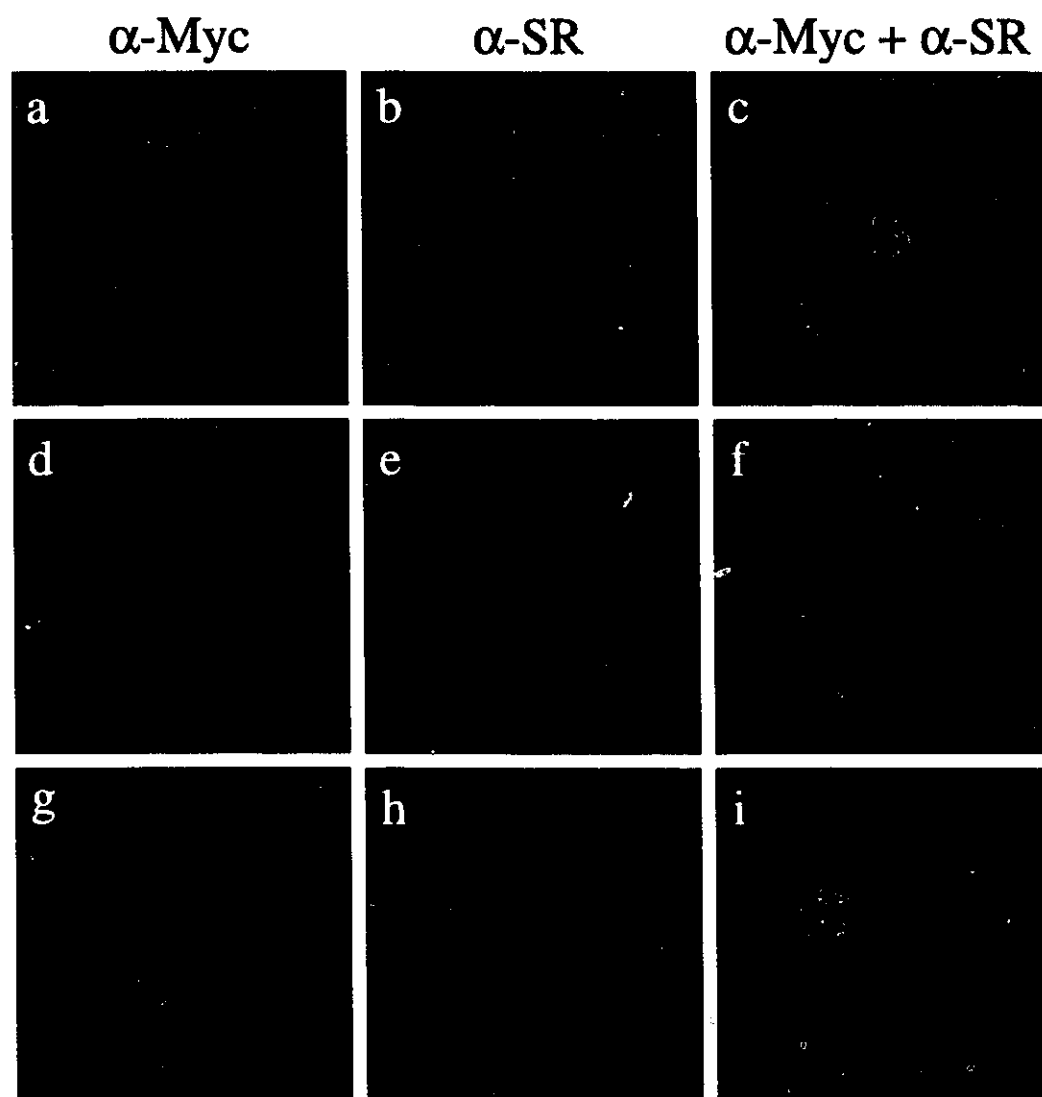
The SR proteins comprise a family of related polypeptides which can regulate constitutive and alternative splicing of pre-mRNAs (34,35). We have shown above that hClk2 and hClk3 overexpression may regulate the distribution of the SR splicing factors. To determine the *in vivo* effects of hClk2 and hClk3 on pre-mRNA splicing we transiently transfected of hClk2 or hClk3 expression vectors along with the mClk1 splicing mini-gene, CMV/*CR-1* (see Chap. 4). CMV/*CR-1* can produce two transcripts, EB⁺ and EB⁻, by alternative splicing of a unique exon (EB). *CR-1* spliced mRNA products were then assessed by RT-PCR analysis as described in Materials and Methods. When expressed in COS-1 cells the *CR-1* mini-gene leads to the production of predominately EB⁻ mRNA (Fig.

Figure 4

Nuclear Localization of hClk3 and hClk3^T and their effects on the distribution of SR proteins.

Indirect immunofluorescent staining and confocal microscopy of transfected COS-1 cells using the anti-Myc 9E10 monoclonal antibody (mAb) (α -Myc, *panels a,d,g*), anti-SR 104 mAb (α -SR, *panels b,e,h*) or an overlay of the two signals (α -Myc + α -SR, *panels c,f,i*). Alignment of green and red signals appears yellow. Cells were transfected pECE expression vector encoding hClk3 (*panels a-f*) or hClk3^T (*panels g-i*). Magnification 40X (*panels a-c*) and 63X (*panels d-i*).

Figure 4



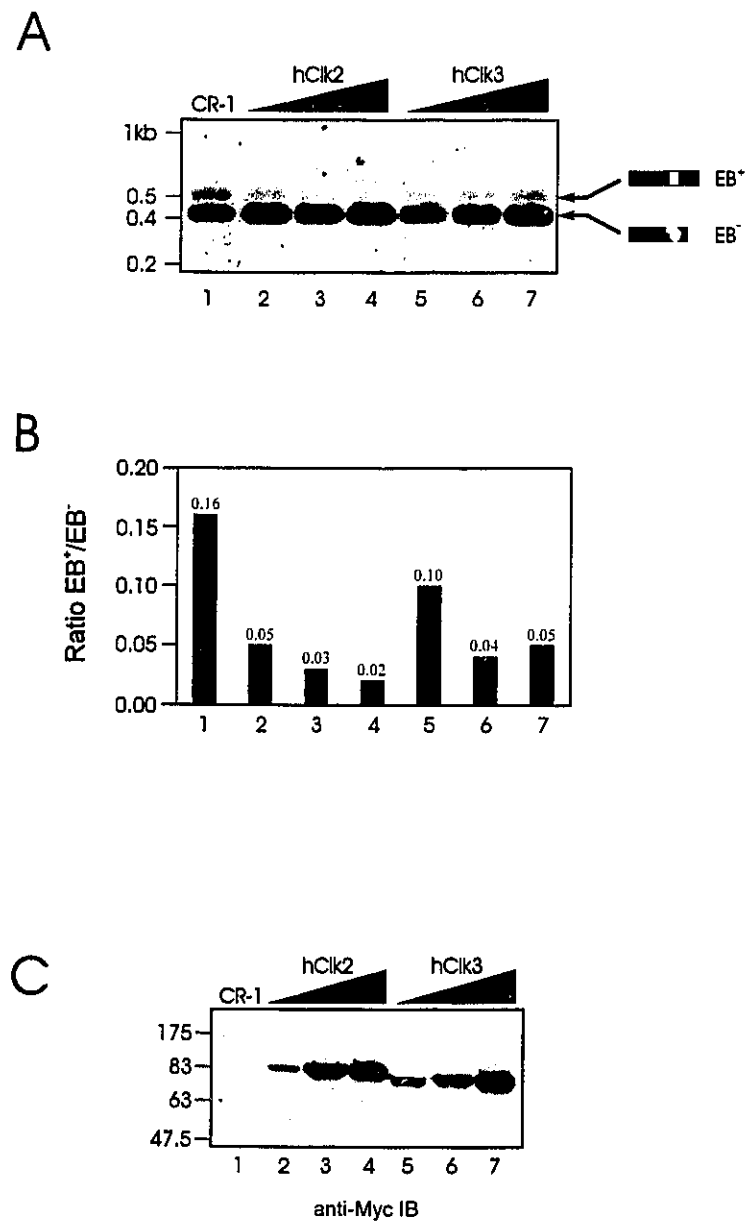
5A,B, lane 1). Cotransfection of hClk2 stimulated exon skipping. The ratio of exon EB inclusion to exon EB skipping decreased as the hClk2 protein level increased, with a maximal increase of approximately eight-fold over *CR-1* alone (Fig. 5A,B,C, lanes 1-4). Similarly, cotransfection of hClk3 promoted exon skipping (Fig. 5A,B,C, lanes 5-7). The maximal increase in the ratio of exon included:skipped mRNA stimulated by hClk3 was only four-fold, again paralleling the differences we observed in the kinase activity of the two enzymes. These results are consistent with the idea that the Clk2 and Clk3 kinases regulate pre-mRNA splicing.

Figure 5

hClk2 and hClk3 regulate alternative splicing *in vivo*.

COS-1 cells were transfected with CMV/*CR-1* mini-gene reporter alone (*lane 1*), or cotransfected with increasing amounts of plasmid (0.5, 2.5 and 5 μ g DNA) encoding hClk2 (*lanes 2,3,4*) or hClk3 (*lanes 5,6,7*), respectively. Cells were harvested in PBS (see Materials and Methods) and total RNA and proteins isolated. *CR-1*-specific RNAs were amplified by RT-PCR resolved by gel electrophoresis, stained with SYBR Green I and visualised by fluorimagery (A). Positions of molecular size standards (kb) are indicated to the left of the panel. Ratio of *CR-1* mRNAs shown in A (B). hClk expression levels were determined from aliquots of cell lysates (C). Proteins were resolved by SDS-PAGE and transfer to nitrocellulose. Proteins were visualized by immunoblot analysis (IB) and enhanced chemiluminescence with the anti-Myc antibody. Positions of molecular mass markers (kDa) are indicated to the left of the panel.

Figure 5



DISCUSSION

Three lines of evidence support the idea that the Clk proteins comprise a unique subfamily of protein kinases. Firstly, Clk1, Clk2 and Clk3 share considerable sequence identity in their catalytic domains (62-73%) although they have significantly divergent amino-terminal domains (21). Secondly, the three Clk genes display a conserved splicing pattern which has not been found in other members of the protein kinase superfamily (21). Thirdly, we show here that Clk2 and Clk3 share biochemical and biological properties with Clk1. All three enzymes, have dual-specificity protein kinase activity, are found in the nucleus and co-exist with naturally truncated isoforms. Like Clk1 (62), Clk2 and Clk3 upon overexpression are able to effect changes in the subnuclear localization of SR proteins and can affect pre-mRNA splicing *in vivo*.

Alternative splicing is critical to the control of cellular and viral gene expression (25). Many factors are involved in regulating splicing, however, SR proteins have been implicated as essential components of both constitutive and alternative splicing programs (recently reviewed in (34,35)). One of the important questions that still remains to be answered is how is the splicing machinery regulated. It has been shown that the level of expression of SR proteins affects splicing decisions such as exon exclusion/inclusion (114,131). Tissue specific variations in SR protein expression could account for developmental regulation of mRNA splicing (45) however, it could be argued that post-translational modification of SR proteins would be a more dynamic regulatory mechanism. SR proteins are known to be phosphorylated (32,33) and previous studies have demonstrated that the SRPK1 and mClk1 kinases can phosphorylate and initiate the re-

distribution of SR proteins *in vivo* (62,106). In the case of SRPK1, it is clear that this kinase is able to influence pre-mRNA splicing *in vitro* (106). The results presented here indicate that the two related but distinct kinases, hClk2 and hClk3, are able to affect SR protein distribution and pre-mRNA splicing *in vivo*. Taken together these results support the idea that protein kinases are able to influence pre-mRNA splicing and thus provide a possible connection between signal transduction pathways and splicing.

What is the relationship between SR protein re-localization and pre-mRNA splicing? It has been suggested that the speckles represent stores of inactive splicing factors and that splicing occurs in the nucleoplasm at or near sites of transcription (67,70,71). As mentioned above the concentration of SR proteins has been shown to affect alternative splicing patterns (33). We suggest that the Clk kinases regulate SR protein release from speckles thus altering their effective nucleoplasmic concentration. Alternatively phosphorylation of SR proteins could directly affect their activity in the splicing reaction. Possibly each of the Clk kinases phosphorylates a discrete subset of SR proteins thereby regulating the splicing of a distinct cohort of pre-mRNAs. Alternatively the three kinases could phosphorylate the same SR proteins but interface with distinct signal transduction pathways.

One of the unique properties of the Clk kinase family is the common mode of differential splicing (21,96) which results in the co-expression of catalytically inactive and catalytically active isoforms. Our demonstration here that the hClk2^T and hClk3^T isoforms associate with SR proteins but do not affect their localization suggests that the natural activity of the truncated proteins may be to antagonize the full length active kinase.

Recently, we have found that overexpression of the mouse Clk1^T protein promotes exon inclusion while the Clk1 kinase promotes exon exclusion (see Chap. 4). These observations coupled with the demonstration that Clk1 can form heterodimers with Clk1^T (96) suggests that the alternatively spliced Clk isoforms may be self regulatory.

CONCLUSION

The murine Clk1 kinase was identified as a novel enzyme displaying the unusual property of dual-specificity phosphorylation when expressed in bacteria. This inspired the question: **what is the role(s) of Clk1 in mammalian cells?** We have used molecular biological techniques in an attempt to answer this question. The work presented in this thesis points to a function for mClk1, and the related hClk2 and 3 kinases, in pre-mRNA splicing.

The dissection of the non-catalytic domain of mClk1 has identified regions which may impart levels of regulation on mClk1 function. These include possible dimerisation domains and regulatory phosphorylation sites (Fig. 1). An important issue this raises is whether the related kinases hClk1, 2 and 3 have similar dimerisation and phosphorylation motifs as mClk1.

Based on the results presented in this thesis and previous work in the field a possible model of mClk1 function is presented in Figures 2 and 3.

Nuclear speckles contain a cohort of splicing factors, including SR proteins (69-71). Active mClk1 liberates SR proteins from these storage sites, thus, increasing their effective nucleoplasmic concentration. Indeed, changes in SR protein concentration have been shown to affect pre-mRNA splicing (33,45,50,114) and may be a mode of splicing regulation.

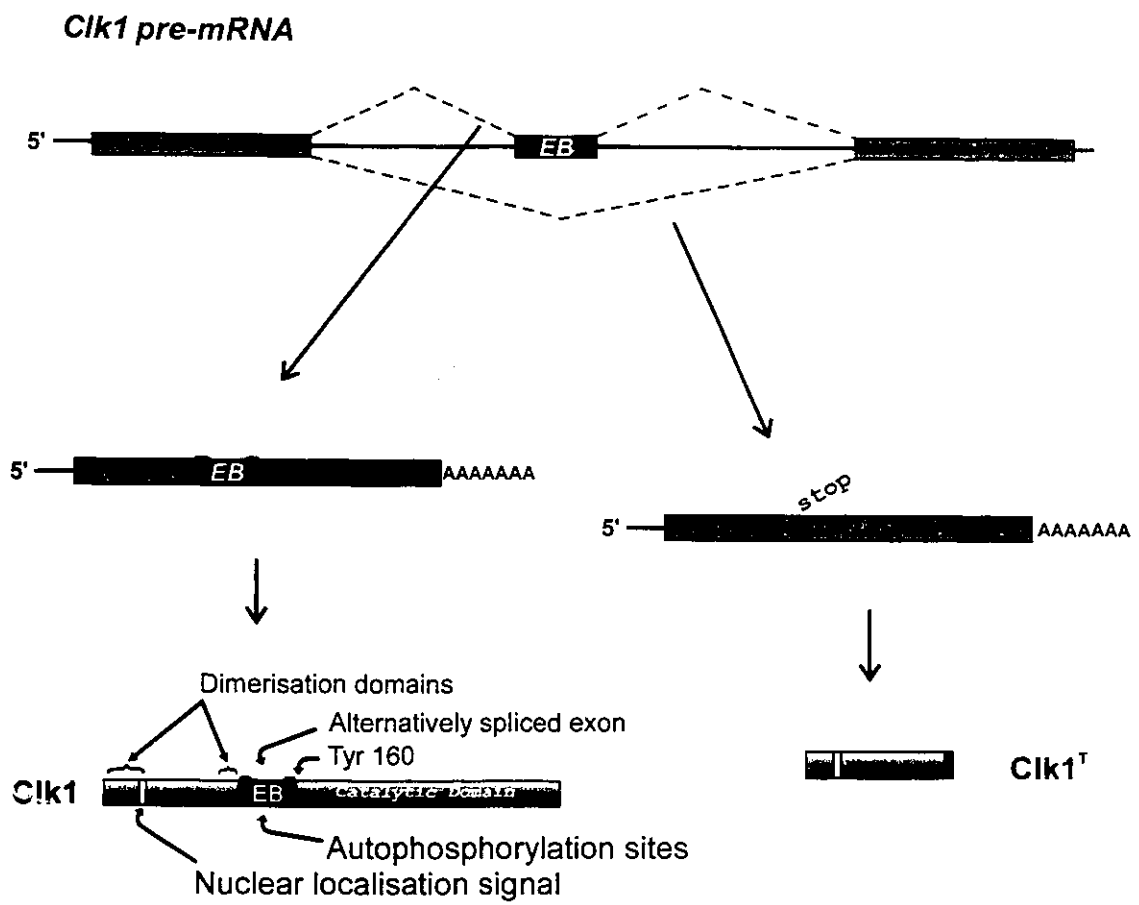
How does mClk1 regulate its own splicing? The active mClk1 kinase promotes skipping of exon EB and therefore down regulates production of its own full-length mRNA. In contrast, the truncated Clk1^T protein promotes exon EB inclusion thus

Figure 1

Schematic representation of Clk pre-mRNA splicing and translated protein products.

The mClk1 precursor mRNA is alternatively spliced through the inclusion or exclusion of exon EB. EB inclusion leads to production of full-length mClk1 polypeptide containing a nuclear localisation site at the amino-terminus and autophosphorylation sites and a potential upstream kinase phosphorylation site (Tyr 160) encoded by EB. EB exclusion leads to production of truncated mClk1T through the introduction of a translation stop codon by frame shifting. This polypeptide lacks the catalytic domain and EB-encoded phosphorylation sites but retains the nuclear localisation signal.

Figure 1



increasing the production of the full-length *mClk1* kinase mRNA. In the proposed model, there exists in the *mClk1* pre-mRNA a splicing element in or near the alternatively spliced exon, EB, which can be bound by an SR protein(s). This *cis* splicing motif may be either a repressor sequence (Fig. 2) or an enhancer sequence (Fig. 3) of exon inclusion.

In the presence of a splicing repressor element just upstream of exon EB the model proposes that: active mClk1 homodimers phosphorylate SR proteins causing their release from the storage speckles, promoting their binding to the repressor element and, thus, favouring skipping of exon EB. As a consequence, the level of Clk^T protein is increased which in turn forms inactive Clk:Clk^T heterodimers. These heterodimers are unable to phosphorylate SR proteins and thus exon EB inclusion is favoured (Fig. 2).

In the presence of an exonic splicing enhancer (Fig. 3) **hypophosphorylated** SR proteins bind the enhancer, stimulate spliceosomal recognition of the adjacent upstream 3' splice site/branch point, thus promoting exon inclusion. Active mClk1 homodimers would cause the **hyperphosphorylation** of these SR proteins, preventing their further association with the enhancer element, thus, promoting exon skipping at the expense of exon inclusion. Hyperphosphorylation of SR proteins would also cause their premature release from speckles, sites of SR protein recycling or spliceosomal complex assembly.

The two possibilities mentioned here may not be mutually exclusive, as phosphorylation of a particular SR protein by mClk1 may increase binding to certain splicing elements (ie. intron repressors) while inhibiting binding to others (ie. exon enhancers). Alternatively, phosphorylation of a particular SR protein by mClk1 may

Figure 2

A model for mClk1-mediated splicing regulation of the *mClk1* pre-mRNA through a splicing repressor element.

In this model the mClk1 pre-mRNA contains a splicing repressor element upstream of the alternatively spliced exon EB. mClk1 phosphorylated SR proteins can bind this element and prevent inclusion of EB. See text for details.

Figure 2

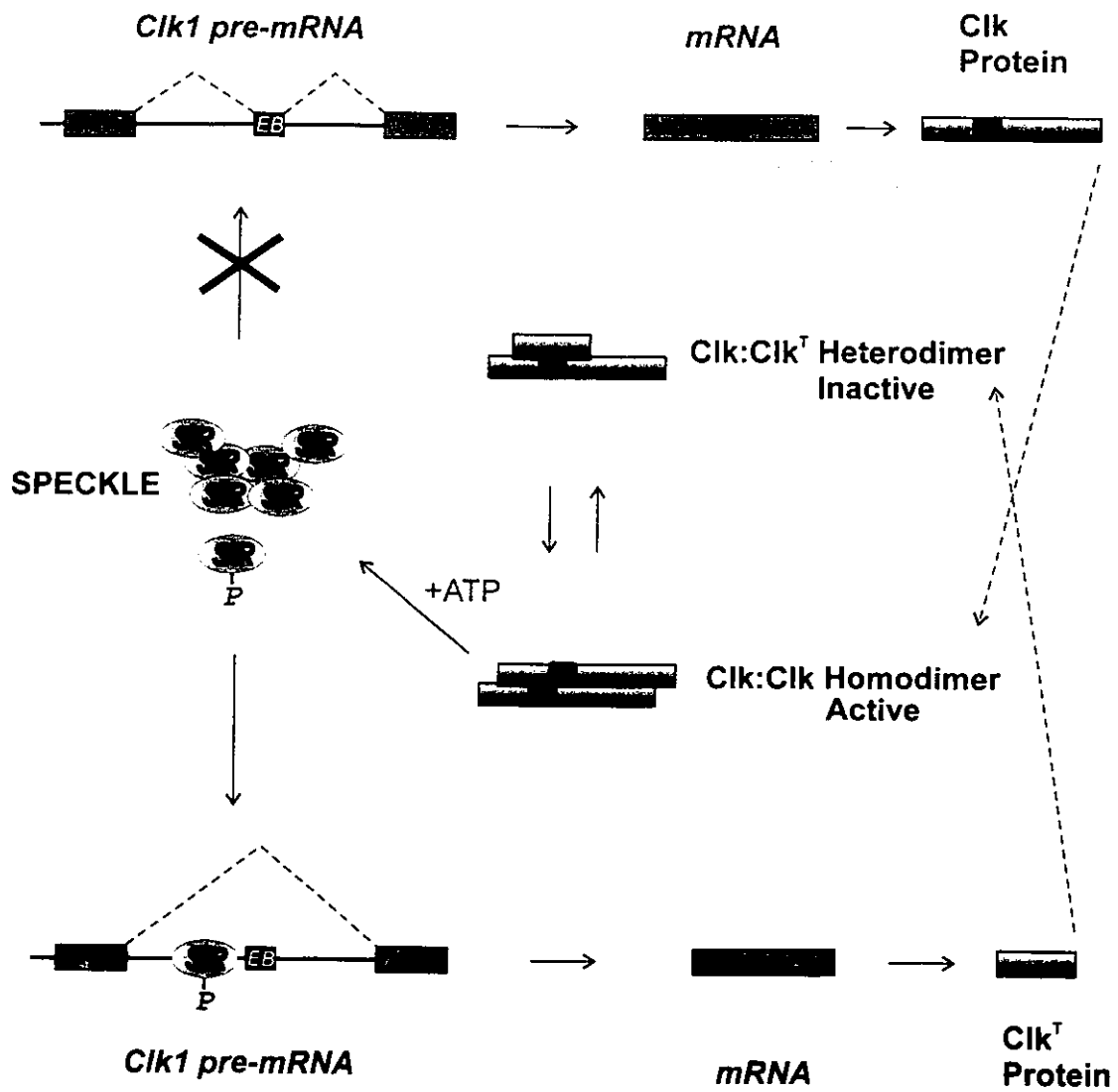
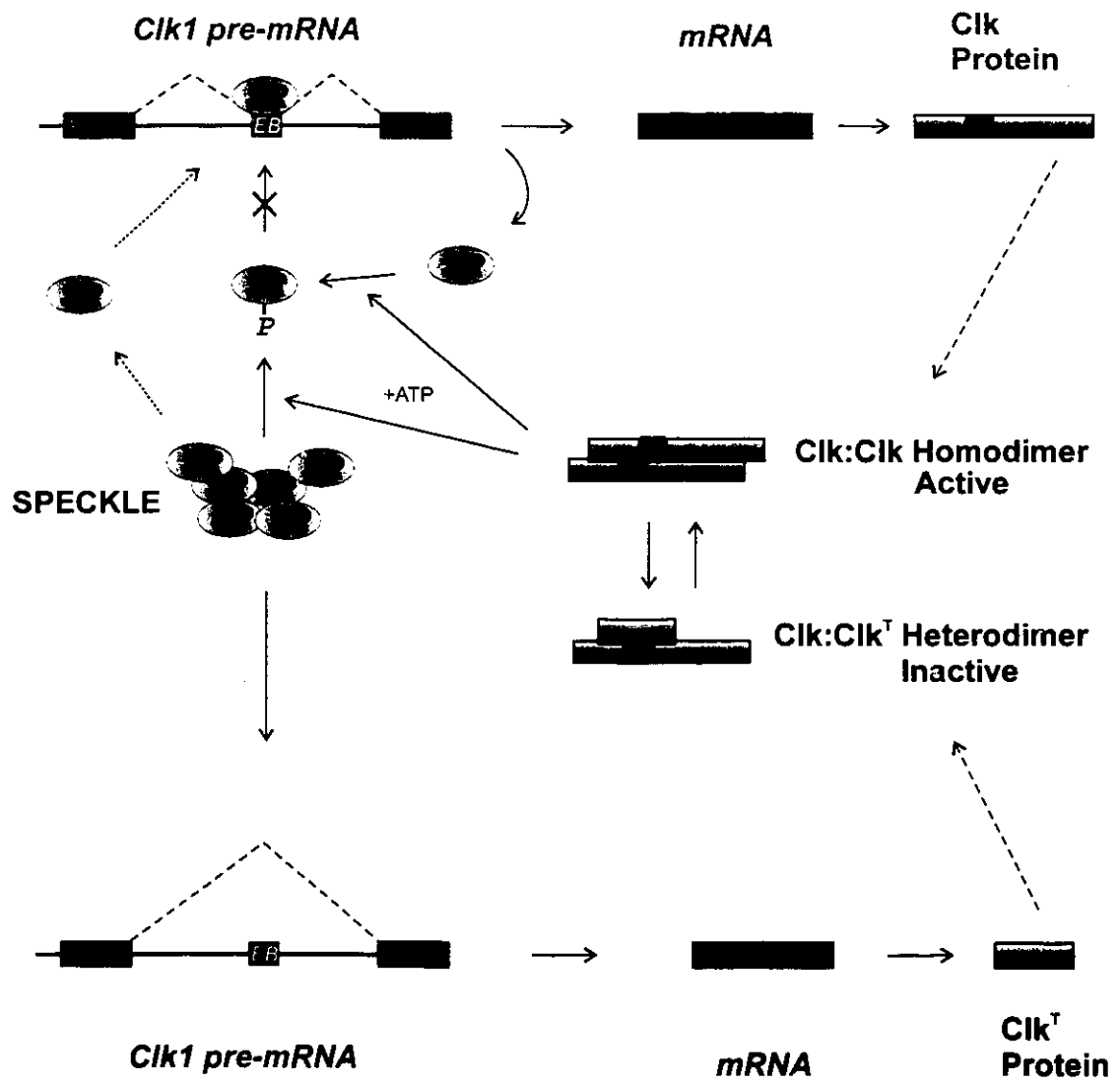


Figure 3

A model for mClk1-mediated splicing regulation of the *mClk1* pre-mRNA through an exon splicing enhancer.

In this model the mClk1 pre-mRNA contains a splicing enhancer element within the alternatively spliced exon EB. Binding of SR proteins to this element stimulates recognition of EB and exon inclusion. mClk1 phosphorylated SR proteins are unable to bind this element with the resultant enhancement of exon skipping. See text for details.

Figure 3



promote binding to splicing elements (Fig. 2), whereas phosphorylation of a different SR protein may actually inhibit its binding.

The above model has been outlined for the regulation of the *mClk1* pre-mRNA, however, mClk1 likely regulates the alternative splicing of numerous pre-mRNAs (containing splicing elements) through the phosphorylation of different SR protein family members.

What role do the Clk family members play? We propose that the individual Clks (Clk1, 2 and 3) each regulate a unique subset of SR proteins through their phosphorylation. Such SR protein regulation may impart control of splicing of a unique cohort of pre-mRNAs.

Additionally, the individual Clks may interface between intracellular signalling networks and the post-transcriptional mechanism of mRNA splicing. It is intriguing to speculate that Clk1 is regulated through stress activation of the SAPK pathways, whereas Clk2 or 3 activity is regulated in response to mitogen stimulation of the MAPK pathway. The intrinsic dual-specificity catalytic activity of the Clks may allow these kinases to play key parts in such signal transduction pathways. As mentioned previously, proteins requiring dual phosphorylation on serine/threonine and tyrosine residues may be the substrates of independent PSKs and PYKs, or alternatively, the substrates of a single DSK poised to perform both modifications. Along with PYKs, DSKs and their substrates have the potential to create signalling complexes mediated through dynamic phosphorylation-dependent protein binding sites. Autophosphorylation of DSKs on tyrosine may create the binding sites for secondary signalling proteins containing SH2 and/or PTB domains (116).

It also remains to be seen if components of the RNA splicing machinery are regulated by direct tyrosine phosphorylation.

Controlled gene expression is critical for normal cellular growth and differentiation. The evolutionary conservation of the Clk family of dual-specificity kinases suggests that they play key roles in regulating this expression through pre-mRNA splicing. One might, therefore, predict that de-regulation of splicing would lead to developmental defects and/or disease. Indeed, it has been suggested that at least 15% of all inherited human diseases are a consequence of splicing defects (132). Future work in elucidating the mechanism of regulation of splicing is essential for a complete understanding of such events.

REFERENCES

1. **Hanks, S.K., Quinn, A.M. and Hunter, T.** (1988) The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science* **241**, 42-52.
2. **The Protein Kinase *FactsBook*: Protein-Serine Kinases.** (1995) Eds.: G. Hardie and S. Hanks (Academic Press Limited, London)
3. **The Protein Kinase *FactsBook*: Protein Tyrosine Kinases.** (1995) Eds.: G. Hardie and S. Hanks (Academic Press Limited, London)
4. **Cooper, J.A., Sefton, B.M. and Hunter, T.** (1983) Detection and quantification of phosphotyrosine in proteins. *Methods. Enzymol.* **99**, 387-402.
5. **Lindberg, R.A., Quinn, A.M. and Hunter, T.** (1992) Dual-specificity protein kinases: will any hydroxyl do? *Trends. Biochem. Sci.* **17**, 114-119.
6. **Douville, E.M.J., Duncan, P.I., Abraham, N. and Bell, J.C.** (1994) Dual specificity kinases - a new family of signal transducers. *Cancer Metast. Rev.* **13**, 1-7.
7. **Howell, B.W., Gray, D.A. and Bell, J.C.** (1991) Evolution of complex intracellular signalling networks: diversification and conserved features of protein kinases. *seminars in DEVELOPMENTAL BIOLOGY* **2**, 345-351.
8. **Seger, R. and Krebs, E.G.** (1995) The MAPK signaling cascade. *FASEB J.* **9**, 726-735.
9. **Pelech, S.L.** (1996) Signalling pathways: Kinase connections on the cellular intranet. *Curr. Biol.* **6**, 551-554.
10. **Kyriakis, J.M. and Avruch, J.** (1996) Protein kinase cascades activated by stress and inflammatory cytokines. *BioEssays* **18**, 567-577.
11. **Seger, R., Ahn, N.G., Boulton, T.G., Yancopoulos, G.D., Panayotatos, N., Radziejewska, E., Ericsson, L., Bratlien, R.L., Cobb, M.H. and Krebs, E.G.** (1991) Microtubule-associated protein 2 kinases, ERK1 and ERK2, undergo autophosphorylation on both tyrosine and threonine residues: Implications for their mechanism of activation. *Proc. Natl. Acad. Sci. USA* **88**, 6142-6146.
12. **Kohtz, J.D., Jamison, S.F., Will, C.L., Zuo, P., Lührmann, R., Garcia-Blanco, M.A. and Manley, J.L.** (1994) Protein-protein interactions and 5'-splice-site recognition in mammalian mRNA precursors. *Nature* **368**, 119-124.

13. **Kosako, H., Gotoh, Y., Matsuda, S., Ishikawa, M. and Nishida, E.** (1992) *Xenopus* MAP kinase activator is a serine/threonine/tyrosine kinase activated by threonine phosphorylation. *EMBO J.* **11**, 2903-2908.
14. **Gonzalez, F.A., Raden, D.L. and Davis, R.J.** (1991) Identification of substrate recognition determinants for human ERK1 and ERK2 protein kinases. *J. Biol. Chem.* **266**, 22159-22163.
15. **Kyriakis, J.M., App, H., Zhang, X.-F., Banerjee, P., Brautigan, D.L., Rapp, U.R. and Avruch, J.** (1992) Raf-1 activates MAP kinase-kinase. *Nature* **358**, 417-421.
16. **Kyriakis, J.M., Brautigan, D.L., Ingebritsen, T.S. and Avruch, J.** (1991) pp54 Microtubule-associated protein-2 kinase requires both tyrosine and serine/threonine phosphorylation for activity. *J. Biol. Chem.* **266**, 10043-10046.
17. **Nigg, E.A.** (1995) Cyclin-dependent protein kinases: key regulators of the eukaryotic cell cycle. *BioEssays* **17**, 471-480.
18. **Howell, B.W., Afar, D.E.H., Lew, J., Douville, E.M.J., Icely, P.L.E., Gray, D.A. and Bell, J.C.** (1991) STY, a tyrosine-phosphorylating enzyme with sequence homology to serine/threonine kinases. *Mol. Cell. Biol.* **11**, 568-572.
19. **Ben-David, Y., Letwin, K., Tannock, L., Bernstein, A. and Pawson, T.** (1991) A mammalian protein kinase with potential for serine/threonine and tyrosine phosphorylation is related to cell cycle regulators. *EMBO J.* **10**, 317-325.
20. **Johnson, K.W. and Smith, K.A.** (1991) Molecular cloning of a novel human cdc2/CDC28-like protein kinase. *J. Biol. Chem.* **266**, 3402-3407.
21. **Hanes, J., von der Kammer, H., Klaudiny, J. and Scheit, K.H.** (1994) Characterization by cDNA cloning of two new human protein kinases. *J. Mol. Biol.* **244**, 665-672.
22. **Yun, B., Farkas, R., Lee, K. and Rabinow, L.** (1994) The *Doa* locus encodes a member of a new protein kinase family and is essential for eye and embryonic development in *Drosophila melanogaster*. *Genes Dev.* **8**, 1160-1173.
23. **Bender, J. and Fink, G.R.** (1994) AFC1, a LAMMER kinase from *Arabidopsis thaliana*, activates STE12-dependent processes in yeast. *Proc. Natl. Acad. Sci. USA* **91**, 12105-12109.

24. **Padmanabha, R., Gehrung, S. and Snyder, M.** (1991) The KNS1 gene of *Saccharomyces cerevisiae* encodes a non-essential protein kinase homologue that is distantly related to members of the CDC28/cdc2 gene family. *Mol. Gen. Genet.* **229**, 1-9.
25. **Smith, C.W.J., Patton, J.G. and Nadal-Ginard, B.** (1989) Alternative splicing in the control of gene expression. *Ann. Rev. Genet.* **23**, 527-577.
26. **Leff, S.E., Rosenfeld, M.G. and Evans, R.M.** (1986) Complex transcriptional units: Diversity in gene expression by alternative RNA processing. *Annu. Rev. Biochem.* **55**, 1091-1117.
27. **Guthrie, C. and Patterson, B.** (1988) Spliceosomal snRNAs. *Ann. Rev. Genet.* **22**, 387-419.
28. **Lamm, G.M. and Lamond, A.I.** (1993) Non-snRNP protein splicing factors. *Biochem. Biophys. Acta* **1173**, 247-265.
29. **Moore, M.J., Query, C.C. and Sharp, P.A.** (1993) In: *The RNA World*, Eds.: R.F. Gesteland and J.F. Atkins (Cold Spring Harbor Laboratory Press, New York) pp 303-357.
30. **Zamore, P.D., Patton, J.G. and Green, M.R.** (1992) Cloning and domain structure of the mammalian splicing factor U2AF. *Nature* **355**, 609-614.
31. **Valcárcel, J., Gaur, R.K., Singh, R. and Green, M.R.** (1996) Interaction of U2AF⁶⁵ RS region with pre-mRNA of branch point and promotion base pairing with U2 snRNA. *Science* **273**, 1706-1709.
32. **Roth, M.B., Murphy, C. and Gall, J.G.** (1990) A monoclonal antibody that recognizes a phosphorylated epitope stains lampbrush chromosome loops and small granules in the amphibian germinal vesicle. *J. Cell Biol.* **111**, 2217-2223.
33. **Zahler, A.M., Lane, W.S., Stolk, J.A. and Roth, M.B.** (1992) SR proteins: a conserved family of pre-mRNA splicing factors. *Genes Dev.* **6**, 837-847.
34. **Valcárcel, J. and Green, M.R.** (1996) The SR protein family: pleiotropic functions in pre-mRNA splicing. *Trends. Biochem. Sci.* **21**, 296-301.
35. **Manley, J.L. and Tacke, R.** (1996) SR proteins and splicing control. *Genes Dev.* **10**, 1569-1579.

36. **Wu, J.Y. and Maniatis, T.** (1993) Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. *Cell* **75**, 1061-1070.
37. **Cáceres, J.F. and Krainer, A.R.** (1993) Functional analysis of pre-mRNA splicing factor SF2/ASF structural domains. *EMBO J.* **12**, 4715-4726.
38. **Li, H. and Bingham, P.M.** (1991) Arginine/Serine-rich domains of the su(w^a) and tra RNA processing regulators target proteins to a subnuclear compartment implicated in splicing. *Cell* **67**, 335-342.
39. **Hedley, M.L., Amrein, H. and Maniatis, T.** (1995) An amino acid sequence motif sufficient for subnuclear localization of an arginine/serine-rich splicing factor. *Proc. Natl. Acad. Sci. USA* **92**, 11524-11528.
40. **Krainer, A.R., Conway, G.C. and Kozak, D.** (1990) Purification and characterization of pre-mRNA splicing factor SF2 from HeLa cells. *Genes Dev.* **4**, 1158-1171.
41. **Krainer, A.R., Mayeda, A., Kozak, D. and Binns, G.** (1991) Functional expression of cloned human splicing factor SF2: homology to RNA-binding proteins, U1 70K, and Drosophila splicing regulators. *Cell* **66**, 383-394.
42. **Ge, H. and Manley, J.L.** (1990) A protein factor, ASF, controls cell-specific alternative splicing of SV40 early pre-mRNA in vitro. *Cell* **62**, 25-34.
43. **Ge, H., Zuo, P. and Manley, J.L.** (1991) Primary structure of the human splicing factor ASF reveals similarities with drosophila regulators. *Cell* **66**, 373-382.
44. **Fu, X.-D., Mayeda, A., Maniatis, T. and Krainer, A.R.** (1992) General splicing factors SF2 and SC35 have equivalent activities *in vitro*, and both affect alternative 5' and 3' splice site selection. *Proc. Natl. Acad. Sci. USA* **89**, 11224-11228.
45. **Zahler, A.M., Neugebauer, K.M., Lane, W.S. and Roth, M.B.** (1993) Distinct functions of SR proteins in alternative pre-mRNA splicing. *Science* **260**, 219-222.
46. **Fu, X.-D.** (1993) Specific commitment of different pre-mRNAs to splicing by single SR proteins. *Nature* **365**, 82-85.
47. **Zuo, P. and Manley, J.L.** (1994) The human splicing factor ASF/SF2 can specifically recognize pre-mRNA 5' splice sites. *Proc. Natl. Acad. Sci. USA* **91**, 3363-3367.

48. **Fu, X.-D. and Maniatis, T.** (1990) Factor required for mammalian spliceosome assembly is localized to discrete regions in the nucleus. *Nature* **343**, 437-441.
49. **Tacke, R. and Manley, J.L.** (1995) The human splicing factors ASF/SF2 and SC35 possess distinct, functionally significant RNA binding specificities. *EMBO J.* **14**, 3540-3551.
50. **Mayeda, A. and Krainer, A.R.** (1992) Regulation of alternative pre-mRNA splicing by hnRNP A1 and splicing factor SF2. *Cell* **68**, 365-375.
51. **Watakabe, A., Tanaka, K. and Shimura, Y.** (1993) The role of exon sequences in splice site selection. *Genes Dev.* **7**, 407-418.
52. **Lavigne, A., La Branche, H., Kornblihtt, A.R. and Chabot, B.** (1993) A splicing enhancer in the human fibronectin alternate ED1 exon interacts with SR proteins and stimulates U2 snRNP binding. *Genes Dev.* **7**, 2405-2417.
53. **Sun, Q., Mayeda, A., Hampson, R.K., Krainer, A.R. and Rottman, F.M.** (1993) General splicing factor SF2/ASF promotes alternative splicing by binding to an exonic splicing enhancer. *Genes Dev.* **7**, 2598-2608.
54. **Ramchatesingh, J., Zahler, A.M., Neugebauer, K.M., Roth, M.B. and Cooper, T.A.** (1995) A subset of SR proteins activates splicing of the cardiac troponin T alternative exon by direct interactions with an exonic enhancer. *Mol. Cell. Biol.* **15**, 4898-4907.
55. **Kanopka, A., Mühlemann, O. and Akusjärvi, G.** (1996) Inhibition by SR proteins of splicing of a regulated adenovirus pre-mRNA. *Nature* **381**, 535-538.
56. **Fu, X.-D. and Maniatis, T.** (1992) The 35-kDa mammalian splicing factor SC35 mediates specific interactions between U1 and U2 small nuclear ribonucleoprotein particles at the 3' splice site. *Proc. Natl. Acad. Sci. USA* **89**, 1725-1729.
57. **Mermoud, J.E., Cohen, P. and Lamond, A.I.** (1992) Ser/Thr-specific protein phosphatases are required for both catalytic steps of pre-mRNA splicing. *Nucleic Acids Res.* **20**, 5263-5269.
58. **Mermoud, J.E., Cohen, P.T.W. and Lamond, A.I.** (1994) Regulation of mammalian spliceosome assembly by a protein phosphorylation mechanism. *EMBO J.* **13**, 5679-5688.

59. Tazi, J., Kornstädt, U., Rossi, F., Jeanteur, P., Cathala, G., Brunel, C. and Lührmann, R. (1993) Thiophosphorylation of U1-70K protein inhibits pre-mRNA splicing. *Nature* **363**, 283-286.
60. Roth, M.B., Zahler, A.M. and Stolk, J.A. (1991) A conserved family of nuclear phosphoproteins localized to sites of polymerase II transcription. *J. Cell Biol.* **115**, 587-596.
61. Woppmann, A., Will, C.L., Kornstädt, U., Zuo, P., Manley, J.L. and Lührmann, R. (1993) Identification of an snRNP-associated kinase activity that phosphorylates arginine/serine rich domains typical of splicing factors. *Nucleic Acids Res.* **21**, 2815-2822.
62. Colwill, K., Pawson, T., Andrews, B., Prasad, J., Manley, J.L., Bell, J.C. and Duncan, P.I. (1996) The Clk/Sty protein kinase phosphorylates SR splicing factors and regulates their intranuclear distribution. *EMBO J.* **15**, 265-275.
63. Nikolakaki, E., Simos, G., Georgatos, S.D. and Giannakouros, T. (1996) A nuclear envelope-associated kinase phosphorylates arginine-serine motifs and modulates interactions between the lamin B receptor and other nuclear proteins. *J. Biol. Chem.* **271**, 8365-8372.
64. Rossi, F., Labourier, E., Forné, T., Divita, G., Derancourt, J., Riou, J.F., Antoine, E., Cathala, G., Brunel, C. and Tazi, J. (1996) Specific phosphorylation of SR proteins by mammalian DNA topoisomerase I. *Nature* **381**, 80-82.
65. LeMaire, M.F. and Thummel, C.S. (1990) Splicing precedes polyadenylation during *Drosophila* E74A transcription. *Mol. Cell. Biol.* **10**, 6059-6063.
66. Baurén, G. and Wieslander, L. (1994) Splicing of Balbiani ring gene pre-mRNA occurs simultaneously with transcription. *Cell* **76**, 183-192.
67. Zhang, G., Taneja, K.L., Singer, R.H. and Green, M.R. (1994) Localization of pre-mRNA splicing in mammalian nuclei. *Nature* **372**, 809-812.
68. Spector, D.L., Fu, X.-D. and Maniatis, T. (1991) Associations between distinct pre-mRNA splicing components and the cell nucleus. *EMBO J.* **10**, 3467-3481.
69. Spector, D.L., Schrier, W.H. and Busch, H. (1983) Immunoelectron microscopic localization of snRNPs. *Biol. Cell* **49**, 1-10.

70. **Fakan, S., Leser, G. and Martin, T.E.** (1984) Ultrastructural distribution of nuclear ribonucleoproteins as visualized by immunocytochemistry on thin sections. *J. Cell Biol.* **98**, 358-363.
71. **Spector, D.L.** (1993) Macromolecular domains within the cell nucleus. *Annu. Rev. Cell Biol.* **9**, 265-315.
72. **Roth, M.B.** (1995) Spheres, colied bodies and nuclear bodies. *Curr. Op. Cell Biol.* **7**, 325-328.
73. **Featherstone, C. and Russell, P.** (1991) Fission yeast p107^{wee1} mitotic inhibitor is a tyrosine/serine kinase. *Nature* **349**, 808-811.
74. **Nakielnny, S., Cohen, P., Wu, J. and Sturgill, T.** (1992) MAP kinase activator from insulin-stimulated skeletal muscle is a protein threonine/tyrosine kinase. *EMBO J.* **11**, 2123-2129.
75. **Rabinow, L. and Birchler, J.A.** (1989) A dosage-sensitive modifier of retrotransposon-induced alleles of the *Drosophila white* locus. *EMBO J.* **8**, 879-889.
76. **Myers, M.P., Murphy, M.B. and Landreth, G.** (1994) The dual-specificity CLK kinase induces neuronal differentiation of PC12 cells. *Mol. Cell. Biol.* **14**, 6954-6961.
77. **Sambrook, J., Fritsch, E.F. and Maniatis, T.** **Molecular cloning: a laboratory manual**, (1989) (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.)
78. **Sanger, F., Nicklen, S. and Coulson, A.R.** (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467.
79. **Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J.** (1979) Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* **18**, 5294-5299.
80. **Jacobson, A.** (1987) Purification and fractionation of Poly(A)⁺ RNA. *Methods. Enzymol.* **152**, 254-261.
81. **Ellis, L., Clauser, E., Morgan, D.O., Edery, M., Roth, R.A. and Rutter, W.J.** (1986) Replacement of insulin receptor tyrosine residues 1162 and 1163 compromises insulin-stimulated kinase activity and uptake of 2-deoxyglucose. *Cell* **45**, 721-732.
82. **Evan, G.I., Lewis, G.K., Ramsay, G. and Bishop, J.M.** (1985) Isolation of monoclonal antibodies specific for human c-myc proto-oncogene product. *Mol. Cell. Biol.* **5**, 3610-3616.

83. **Muller, A.J., Pendergast, A., Havlik, M.H., Puil, L., Pawson, T. and Witte, O.N.** (1992) A limited set of SH2 domains binds BCR through a high-affinity phosphotyrosine-independent interaction. *Mol. Cell. Biol.* **12**, 5087-5093.
84. **Lipman, D.J. and Pearson, (1985)** Rapid and sensitive protein similarity searches. *Science* **227**, 1435-1441.
85. **Klein, R., Conway, D., Parada, L.F. and Barbacid, M.** (1990) The trkB tyrosine protein kinase gene codes for a second neurogenic receptor that lacks the catalytic kinase domain. *Cell* **61**, 647-656.
86. **Werner, S., Duan, D.R., de Vries, C., Peters, K.G., Johnson, D.E. and Williams, L.T.** (1992) Differential splicing in the extracellular region of fibroblast growth factor receptor 1 generates receptor variants with different ligand-binding specificities. *Mol. Cell. Biol.* **12**, 82-88.
87. **Dorai, T. and Wang, L.** (1990) An alternative non-tyrosine protein kinase product of the c-src gene in chicken skeletal muscle. *Mol. Cell. Biol.* **10**, 4068-4079.
88. **Lewin, B. Genes,** (1987) (John Wiley & Sons, Inc. New York)
89. **Sosnowski, B.A., Belote, J.M. and McKeown, M.** (1989) Sex-specific alternative splicing of RNA from the transformer gene results from sequence-dependent splice site blockage. *Cell* **58**, 449-459.
90. **Scott, R.E.M., Autelitano, D.J., Lugo, D.L., Blum, M., Roberts, J.L. and Pintar, J.E.** (1990) Developmental changes in levels of proopiomelanocortin intron A-containing heterogeneous nuclear RNA and mature messenger RNA in the anterior and neurointermediate lobes of the rat pituitary. *Mol. Endocrinol.* **4**, 812-82.
91. **Qian, L., Theodor, L., Carter, M., Vu, M.N., Sasaki, A.W. and Wilkinson, M.F.** (1993) T Cell receptor- β mRNA splicing: regulation of unusual splicing intermediates. *Mol. Cell. Biol.* **13**, 1686-1696.
92. **Ullrich, A. and Schlessinger, J.** (1990) Signal transduction by receptors with tyrosine kinase activity. *Cell* **61**, 203-212.
93. **Thomis, D.C. and Samuel, C.E.** (1993) Mechanism of interferon action: evidence for intermolecular autophosphorylation and autoactivation of the interferon-induced, RNA-dependent protein kinase PKR. *J. Virol.* **67**, 7695-7700.

94. **Reith, A.D., Ellis, C., Lyman, S.D., Anderson, D.M., Williams, D.E., Bernstein, A. and Pawson, T.** (1991) Signal transduction by normal isoforms and *W* mutant variants of the *Kit* receptor tyrosine kinase. *EMBO J.* **10**, 2451-2459.
95. **Hanks, S.K. and Quinn, A.M.** (1991) Protein kinase catalytic domain sequence database: identification of conserved features of primary structure and classification of family members. *Methods. Enzymol.* **200**, 38-62.
96. **Duncan, P.I., Howell, B.W., Marius, R.M., Drmanic, S., Douville, E.M.J. and Bell, J.C.** (1995) Alternative splicing of STY, a nuclear dual specificity kinase. *J. Biol. Chem.* **270**, 21524-21531.
97. **Fields, S. and Song, O.** (1989) A novel genetic system to detect protein-protein interactions. *Nature* **340**, 245-246.
98. **Soulard, M., Della Valle, V., Siomi, M.C., Piñol-Roma, S., Codogno, P., Bauvy, C., Bellini, M., Lacroix, J.-C., Monod, G., Dreyfuss, G. and Larsen, C.-J.** (1993) hnRNP G: sequence and characterization of a glycosylated RNA-binding protein. *Nucleic Acids Res.* **21**, 4210-4217.
99. **Schmidt, G. and Werner, D.** (1993) Sequence of a complete murine cDNA reflecting an S phase-prevalent transcript encoding a protein with two types of nucleic acid binding motifs. *Biochem. Biophys. Acta* **1216**, 317-320.
100. **Ayane, M., Preuss, U., Köhler, G. and Nielsen, P.J.** (1991) A differentially expressed murine RNA encoding a protein with similarities to two types of nucleic acid binding motifs. *Nucleic Acids Res.* **19**, 1273-1278.
101. **Verheijen, R., Kuijpers, H., Vooijs, P., van Venrooij, W. and Ramaekers, F.** (1986) Distribution of the 70k U1 RNA-associated protein during interphase and mitosis. Correlation with other U RNP particles and proteins of the nuclear matrix. *J. Cell Sci.* **86**, 173-190.
102. **Gotoh, Y., Moriyama, K., Matsuda, S., Okumura, E., Kishimoto, T., Kawasaki, H., Suzuki, K. and Nishida, E.** (1991) *Xenopus* M phase MAP kinase: isolation of its cDNA and activation by MPF. *EMBO J.* **10**, 2661-2668.
103. **Tang, Z., Coleman, T.R. and Dunphy, W.G.** (1993) Two distinct mechanisms for negative regulation of the Wee1 protein kinase. *EMBO J.* **12**, 3427-3436.
104. **Amrein, H., Hedley, M.L. and Maniatis, T.** (1994) The role of specific protein-RNA and protein-protein interactions in positive and negative control of pre-mRNA splicing by *Transformer 2*. *Cell* **76**, 735-746.

105. **Tazi, J., Daugeron, M.-C., Cathala, G., Brunel, C. and Jeanteur, P.** (1992) Adenosine phosphorothioates (ATP α S and ATP γ S) differentially affect the two steps of mammalian pre-mRNA splicing. *J. Biol. Chem.* **267**, 4322-4326.
106. **Gui, J., Lane, W.S. and Fu, X.-D.** (1994) A serine kinase regulates intracellular localization of splicing factors in the cell cycle. *Nature* **369**, 678-682.
107. **Huang, S., Deerinck, T.J., Ellisman, M.H. and Spector, D.L.** (1994) *In vivo* analysis of the stability and transport of nuclear poly(A)⁺ RNA. *J. Cell Biol.* **126**, 877-899.
108. **Krainer, A.R., Conway, G.C. and Kozak, D.** (1990) The essential pre-mRNA splicing factor SF2 influences 5' splice site selection by activating proximal sites. *Cell* **62**, 35-42.
109. **Lundgren, K., Walworth, N., Booher, R., Dembski, M., Kirschner, M. and Beach, D.** (1991) *Cell* **64**, 1111-1122.
110. **Gluzman, Y.** (1981) SV40-transformed simian cells support the replication of early SV40 mutants. *Cell* **23**, 175-182.
111. **Pear, W.S., Nolan, G.P., Scott, M.L. and Baltimore, D.** (1993) Production of high-titer helper-free retroviruses by transient transfection. *Proc. Natl. Acad. Sci. USA* **90**, 8392-8396.
112. **Klymkowsky, M.W.** (1987) *Development* **8**, 140-142.
113. **DuBridge, R.B., Tang, P. and Hsia, H.C.** (1987) Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system. *Mol. Cell. Biol.* **7**, 379-387.
114. **Cáceres, J.F., Stamm, S., Helfman, D.M. and Krainer, A.R.** (1994) Regulation of alternative splicing in vivo by overexpression of antagonistic splicing factors. *Science* **265**, 1706-1709.
115. **Signal Transduction: a special issue**, (1991) *Trends Genet.* **7**, 343-418.
116. **Pawson, T.** (1995) Protein modules and signalling networks. *Nature* **373**, 573-580.
117. **van der Geer, P. and Pawson, T.** (1995) The PTB domain: a new protein module implicated in signal transduction. *Trends. Biochem. Sci.* **20**, 277-280.
118. **Bork, P. and Sudol, M.** (1994) The WW domain: a signalling site in dystrophin? *Trends. Biochem. Sci.* **19**, 531-533.

119. **Chan, D.C., Bedford, M.T. and Leder, P.** (1996) Formin binding proteins bear WWP/WW domains that bind proline-rich peptides and functionally resemble SH3 domains. *EMBO J.* **15**, 1045-1054.
120. **Muslin, A.J., Tanner, J.W., Allen, P.M. and Shaw, A.S.** (1996) Interaction of 14-3-3 with signaling proteins is mediated by the recognition of phosphoserine. *Cell* **84**, 889-897.
121. **Johnson, L.N., Noble, M.E.M. and Owen, D.J.** (1996) Active and inactive protein kinases: structural basis for regulation. *Cell* **85**, 149-158.
122. **Songyang, Z., Shoelson, S.E., McGlade, J., Olivier, P., Pawson, T., Bustelo, X.R., Barbacid, M., Sabe, H., Hanafusa, H., Yi, T., Ren, R., Baltimore, D., Ratnofsky, S., Feldman, R.A. and Cantley, L.C.** (1994) Specific motifs recognized by the SH2 domains of Csk, 3BP2, fps/fes, GRB-2, HCP, SHC, Syk, and Vav. *Mol. Cell. Biol.* **14**, 2777-2785.
123. **Pawson, T., Letwin, K., Lee, T., Hao, Q.-L., Heisterkamp, N. and Groffen, J.** (1989) The *FER* gene is evolutionarily conserved and encodes a widely expressed member of the *FPS/FES* protein-tyrosine kinase family. *Mol. Cell. Biol.* **9**, 5722-5725.
124. **Hao, Q., Ferris, D.K., White, G., Heisterkamp, N. and Groffen, J.** (1991) Nuclear and cytoplasmic location of the *FER* tyrosine kinase. *Mol. Cell. Biol.* **11**, 1180-1183.
125. **Heldin, C.-H.** (1995) Dimerization of cell surface receptors in signal transduction. *Cell* **80**, 213-223.
126. **Birney, E., Kumar, S. and Krainer, A.R.** (1993) Analysis of the RNA-recognition motif and RS and RGG domains: conservation in metazoan pre-mRNA splicing factors. *Nucleic Acids Res.* **21**, 5803-5816.
127. **English, J.M., Vanderbilt, C.A., Xu, S., Marcus, S. and Cobb, M.H.** (1995) Isolation of MEK5 and differential expression of alternatively spliced forms. *J. Biol. Chem.* **270**, 28897-28902.
128. **Lin, A., Minden, A., Martinetto, H., Claret, F., Lange-Carter, C., Mercurio, F., Johnson, G.L. and Karin, M.** (1995) Identification of a dual specificity kinase that activates the Jun kinases and p38-Mpk2. *Science* **268**, 286-290.

129. **Yang, X., Bani, M.-R., Lu, S.-J., Rowan, S., Ben-David, Y. and Chabot, B.** (1994) The A1 and A1^B proteins of heterogeneous nuclear ribonucleoparticles modulate 5' splice site selection *in vivo*. *Proc. Natl. Acad. Sci. USA* **91**, 6924-6928.
130. **Laemmli, U.K.** (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.
131. **Mayeda, A., Helfman, D.M. and Krainer, A.R.** (1993) Modulation of exon skipping and inclusion by heterogeneous nuclear ribonucleoprotein A1 and pre-mRNA splicing factor SF2/ASF. *Mol. Cell. Biol.* **13**, 2993-3001.
132. **Krawczak, M., Reiss, J. and Cooper, D.N.** (1992) The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: causes and consequences. *Hum. Genet.* **90**, 41-54.