uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

John Mark

AUTEUR DE LA THÈSE / AUTHOR OF THESIS

Ph.D. (Biochemistry)

GRADE / DEGREE

Department of Biochemistry, Microbiology and Immunology

FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Protein characterization by biochemical analysis:
MKP3 binding and auto-inhibitory functions unexpected lability in the SARS nucleocapsid protein

TITRE DE LA THÈSE / TITLE OF THESIS

Mary Alice Hefford

DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

Vasek Mezl

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

John Baenziger

Anthony Clarke

Robert Ben

Martin Young

Gary W. Slater

Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

Protein characterization by biochemical analysis:
MKP3 binding and auto-inhibitory functions
Unexpected lability in the SARS nucleocapsid protein

By
John K. Mark

Thesis submitted to the Faculty of Graduate and Postdoctoral Studies
in partial fulfilment of the requirements for the Doctorate of Philosophy (Ph.D.) degree in

Biochemistry, Microbiology, & Immunology
Faculty of Medicine
University of Ottawa

Your file Votre référence
ISBN: 978-0-494-46516-5
Our file Notre référence
ISBN: 978-0-494-46516-5

Abstract

Two proteins were tested using biochemical assay techniques to determine the reasons for their loss of function or loss of primary structure. The first protein that was analysed, MAP kinase phosphatase 3 (MKP3), dephosphorylates the extracellular signal-regulated kinases (ERKs). Dephosphorylation occurs following ERK binding to MKP3's N-terminal domain, which activates MKP3's allosteric C-terminal phosphatase domain.

To determine the mechanism for MKP3's loss of function in pancreatic adenocarcinoma cells, experiments were conducted to test the effect of potential protein-protein interactions on MKP3 activity. Surface plasmon resonance binding studies using the full-length MKP3 and N- and C-terminal truncation mutants showed that the N- and C-terminal MKP3 domains could bind with a dissociation constant (K_D) of 1.43 +/- 0.41 μM. The low K_D value between the two domains is consistent with the presence of an interdomain binding site. A novel MKP3 activity assay was then developed to test if the binding affected MKP3's phosphatase activity. Chemically-activated MKP3 C-terminal domain was capable of hydrolysing *p*-nitrophenol phosphate with a catalytic efficiency (k_{cat}/K_M) ~33% greater than the full-length MKP3 protein. However, when MKP3 C-terminal domain activity was tested in the presence of the N-terminal domain, the addition led to a dose-dependant inhibition with a K_I (inhibition constant) of 0.68 +/- 0.37 μM. The similarity between the K_D and K_I constants, for MKP3's N- and C-terminal binding and inhibition, implies that MKP3 interdomain binding results in inhibition of MKP3 activity. This finding was compared to experiments showing that pancreatic adenocarcinoma cells contain a non-functional, wild-type MKP3 that forms high molecular weight oligomers. Biochemical analysis and the detection of the oligomers indicates that MKP3 overexpression in adenocarcinomas may lead

to a loss of enzyme function through intercalation of N- and C-terminal domains via a "domain swap" interaction that results in MKP3 autoinhibition.

The second study was performed on the SARS nucleocapsid (N) protein. N protein displayed an unexpected lability and multiple lower molecular weight bands were detected upon N protein purification. Mass spectrometry was used to identify the bands as the N- and C-terminal fragments of SARS N protein, which result from a specific cleavage after each arginine in the N protein's ^{184}*SSRSSSRSRGNSR*^{196} sequence. Fluorescent resonance energy transfer (FRET) based assays that were developed to test for non-specific protease activity showed negative results. This may indicate that the SARS N protein possesses a specific autocatalytic activity.

Acknowledgements

Thanks to everyone who reads this. You know who you are.

Table of Contents

Part II.

List of abbreviations

a.a.	Amino acids
AU	Absorbance unit
DABCYL	4-((4-(dimethylamino)phenyl)azo)benzoic acid
DMF	Dimethyl formamide
DMSO	Dimethyl-sulfoxide
DTT	Dithiothreitol
DUSP	Dual specificity phosphatase
EDC	1-Ethyl-3-[3-dimethylaminopropyl] carbodiimide hydrochloride
EDTA	Elthylenediamine tretraacetic acid
EDANS	5-[(2-aminoethyl)amino]naphthalene-1-sulfonic acid
ERK	Extracellular receptor regulated kinase
FGF	Fibroblast growth factor
FITC	Fluorescein isothiocyanate
FRET	Fluorescent resonance energy transfer
HBS	HEPES buffered saline
HPLC	High performance/pressure liquid chromatography
IC_{50}	50% Inhibition concentration
IPTG	Isopropyl β-D-1-thiogalactopyranoside
JNK	c-Jun N-terminal kinase
k_a	Association rate
K_D	Dissociation constant
K_I	Inhibitory constant
KIM	Kinase interacting motif
LB media	Luria-Bertanii media
LC/MS	Tandem liquid chromatography-mass spectrometry
MALDI-TOF/MS	Matrix assisted laser desorption-time of flight mass spectrometry
MAPK	Mitogen activated protein kinase
MKP3	MAP kinase phosphatase 3
MKP3:CT	C-terminal MKP3 (Δ1-154)
MKP3:NT	N-terminal domain MKP3 (Δ155-381)
MKP3:WT	Wild-type MKP3
MS	Mass spectrometry
MW	Molecular weight
NES	Nuclear export signal
NHS	N-hydroxysuccinimide
PCR	Polymerase chain reaction
PEST	Proline, glutamate, serine, threonine rich sequence
PTP	Protein tyrosine phosphatase
pNPP	para-nitrophenol phosphate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SELDI-TOF/MS	Surface enhanced laser desorption ionization-time of flight mass spectrometry
SPR	Surface plasmon resonance
TBS	Tris buffered saline

List of figures and illustrations

Part I

Part II

List of tables

Part I

Part II

0.0. General introduction

Cells possess a number of mechanisms that regulate biological functions, but post-translational modification, arguably one of the cell's simplest mechanisms, may be the most potent. The example of protein phosphorylation, a focus in this thesis, shows the effects modification can have on overall cell function. Protein phosphorylation leads to the replacement of a hydroxyl group (OH^-) with phosphate (PO_4^{3-}) to serine, threonine, or tyrosine and yet it is the single change responsible for important cell and tissue functions. These functions include: muscle contraction via ATP hydrolysis by myosin (*1*), G-protein subunit recruitment and signalling (*2;3*) and the maintenance of cell homeostasis through active ion channels (*4;5*).

In this thesis, two proteins have been studied. Their activities show that they affect or are affected by forms of post-translational modification. The first protein is MKP3 (MAP kinase phosphatase 3), which is the phosphatase responsible for extracellular-signal regulated kinase (ERK) pathway deactivation, but MKP3 shows an unexplained loss of function when overexpressed under patho-physiological conditions. The second protein is the SARS coronavirus nucleocapsid protein, which has been shown to self assemble and bind viral RNA, but shows a susceptibility to peptide bond hydrolysis.

The unifying theme between the two proteins studied in this thesis was biochemical analysis of protein activity. Biochemical assay methods were developed and used to determine the mechanisms behind MKP3's loss of function and determine the mechanism behind SARS N proteolysis.

1.0. Introduction

1.0.1. Loss of MKP3 function in pancreatic adenocarcinomas

The feedback between cellular kinases and their conjugate phosphatases is necessary for the propagation and termination of most signalling cascades. In MAP kinase signalling, an appropriate stimulus leads to an exponentially growing signal cascade that results from the serial activation of protein kinases (*6-8*). Activation can then be opposed by negative feedback, driven by a specific antagonistic phosphatase (*9-11*). This form of negative regulation is known to occur in ERK1 and ERK2-mediated signalling, which is abrogated by MAP kinase phosphatase 3 (MKP3) upregulation (*12*). In this example, the tight substrate specificity for ERK, that MKP3 possesses, results in specific deactivation of the ERK1 and ERK2 cascades (*11-13*). MKP3 specificity is due to ERK-specific binding at MKP3's N-terminal binding domain and the ERK-induced (allosteric) activation at MKP3's C-terminal catalytic domain (*12-15*). Both binding and catalysis have been documented in cell-free, *in vitro* experiments (*11;12;16;17*).

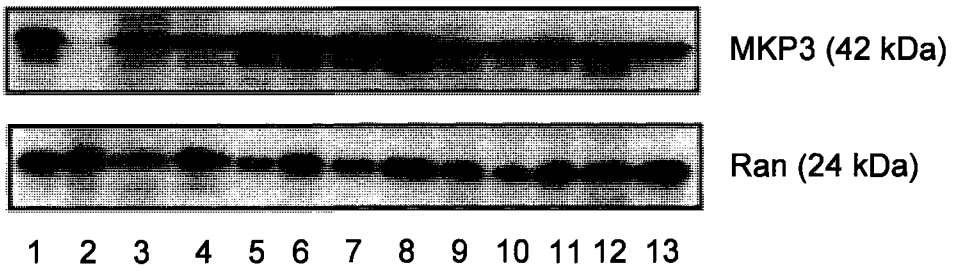
MKP3's binding and catalytic activities imply that it is a negative regulator of the cell's mitotic and developmental signalling pathways. Additionally, because constitutively activating RAS (which activates ERK) and/or ERK mutations are prevalent in pancreatic adenocarcinomas, researchers have suggested that MKP3 could act as a tumour suppressor by deactivating ERK (*18-23*). In fact, previous research initially indicated that pancreatic adenocarcinomas did not express MKP3 protein (*22*). These data correlated with previous studies that detailed a loss of heterozygosity near the *Mkp3* gene locus at 12q22 (*19-21;24*). The studies implied that MKP3 exhibits tumour suppressor function and that the loss of MKP3 resulted in unchecked activation of the ERK signalling pathway. This interpretation

was consistent with data showing anomalous, constitutively active ERK (due to constitutively activating *Ras* mutations) in 90% of pancreatic adenocarcinomas (*18;20*).

However, subsequent work by R. Aubin at Health Canada (Figure 1) showed that many pancreatic adenocarcinoma (PA) cell lines, and even PA's that were previously thought to be MKP3 deficient (*22*), did express wild-type MKP3. The apparent disagreement between these two studies can be attributed to the use of an antibody with a poor affinity for MKP3 in the earlier study (*22*) compared to the anti-MKP3 antibody used by Aubin (mouse anti-MKP3, MAB3576, R&D Systems). In spite of the disagreement over the presence or absence of MKP3 during the later stages of oncogenesis, the two results do show agreement in that constitutively activated RAS or ERK can sustain tumour growth in pancreatic adenocarcinoma. However, finding of high MKP3 levels in pancreatic adenocarcinoma cells introduce a contradiction that cannot be explained by the current model of MKP3 function. This is because MKP3 possesses an excellent affinity and catalytic efficiency in the presence of ERK1 or ERK2 (*15;25;26*). Therefore, by the current view that MKP3 acts as a negative regulator of ERK (*27-31*), one would predict that co-expression of MKP3 would lead to ERK deactivation. The preliminary data showed that MKP3 in its wild-type form was present in pancreatic adenocarcinomas and appeared non-functional.

A potential explanation for MKP3's loss of function is that, in addition to ERK-specific binding and allosteric activation, MKP3 possesses a secondary level of control over its enzymatic activity. Such a regulatory mechanism may involve a third protein that mediates MKP3's ability to bind or dephosphorylate ERK. Alternatively, MKP3 may be exhibiting a substrate or protein level effect, such as post-translational modification or conformational change to mediate its activity. The first hypothesis is unlikely, only p38, which has 50% amino acid identity to ERK2 (*13*) and protein kinase CK2α, which increases

Figure 1. Western detection of MKP3 in pancreatic adenocarcinomas. Western blot analysis to detect the expression of MKP3 was performed by Dr. R Aubin. 20 μg total cellular protein from each of the indicated cell lines were immunodetected. RAN (Ras-associated nuclear protein) was used as the protein loading control. Lane 1. HeLa cells (positive control), Lane2. Normal pancreas, Lane 3. HPDE4, Lane 4. HPDE6, Lane 5-12 pancreatic adenocarcinoma cell lines. Lane5. CRL1420 (MIAPaCa-2), Lane 6. CRL1469, Lane 7. CRL1687 (BxPC-3), Lane 8. CRL1837 (Su86.86), Lane 9. HTB79, Lane 10. HTB80, Lane 11. HTB134, Lane 12. HTB147.


MKP3 (42 kDa)
Ran (24 kDa)
1 2 3 4 5 6 7 8 9 10 11 12 13

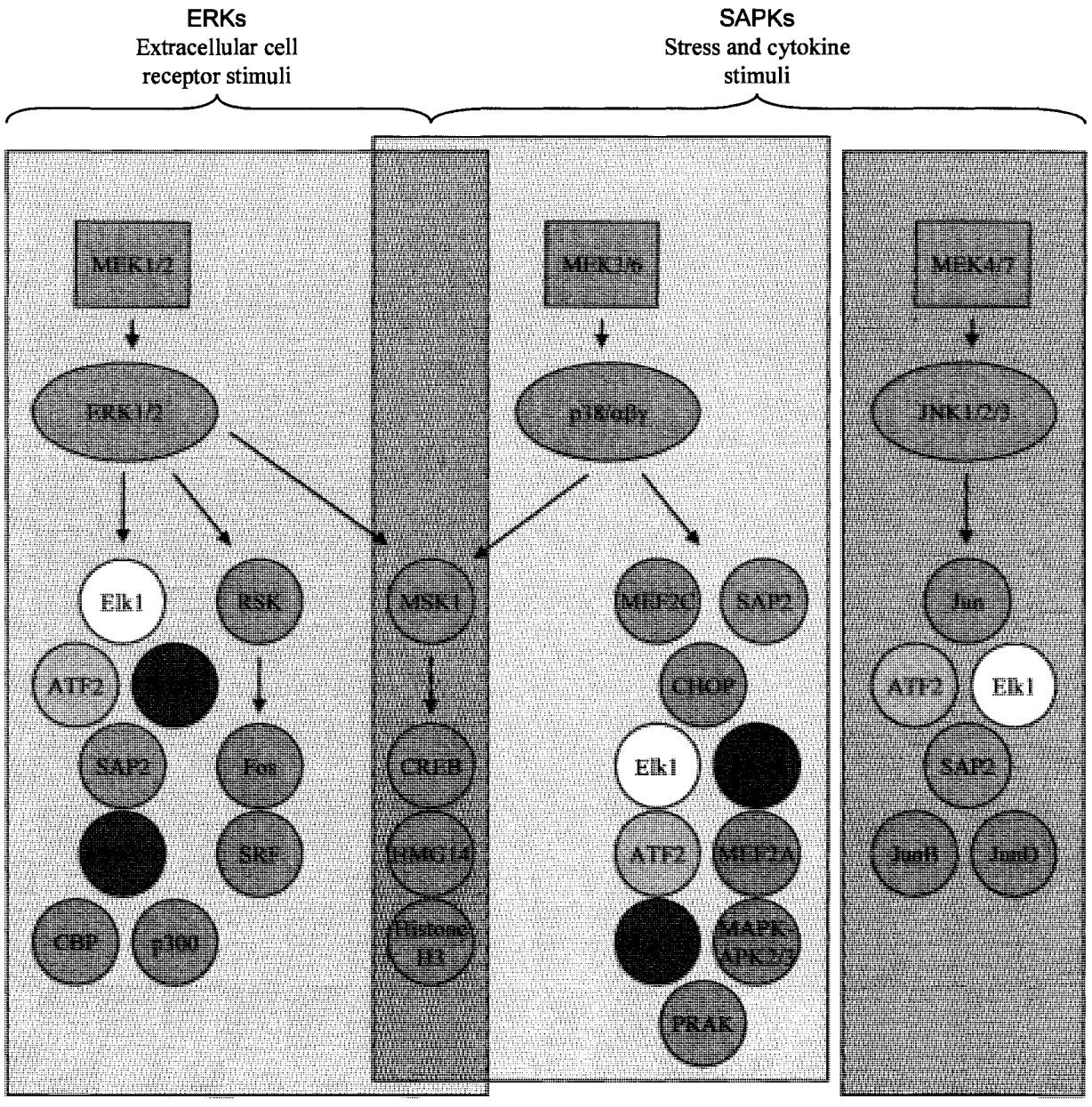
MKP3 phosphatase activity (*32*) have been observed to bind to MKP3. Therefore, it is more likely that MKP3 activity is altered at the protein level rather than through a third party.

1.1. MAP kinases

1.1.1. Overview: MAP kinases

The MAP kinases play major roles in cell regulation by aggregating and coordinating the responses involved in survival, adaptation and development (*6;33*). These critical roles have also become the criteria used to subdivide the kinases into two general categories: the stress activated protein kinases (SAPKs) and the extracellular receptor regulated kinases (ERKs). The SAPK group includes the c-Jun N-terminal kinases (JNKs) and the p38 kinases that respond to cellular stress and oxidative damage. The extracellular receptor regulated kinase group includes the ERKs that are mediated by extracellular receptor stimuli, such as growth factors (*33;34*). While these MAP kinases may be functionally diverse, all members possess a conserved system of activation that results in the addition of phosphate to tyrosine and/or threonine residues at a *-TXY- or -TXT-* "MAPK activation motif" (*6;7;33;35;36*). This activating mechanism is the defining feature of the MAP kinases and it allows them to remain in an inactive state in the absence of phosphorylation. MAPK activity is present only upon phosphorylation of the activation motif, whereupon MAP kinase activity leads to the phosphorylation of intracellular substrates and transcription factors (Figure 2).

Figure 2. MAPK subgroups and signalling pathways. The MAP kinases may be divided into the stress activated protein kinases (SAPKs) or the extracellular receptor regulated kinases (ERKs). The SAPK subgroup consists of the JNK and p38 kinase isoforms that are activated due to cell stress or cytokines. The extracellular receptor regulated kinases, contain the self-named ERKs that respond to growth factors and phorbol esters. The figure shows the general, top-down SAPK and ERK activation pathways. Pathways act independently, but overlap (pathway cross-talk) does occur. The components of the SAPK pathways are highlighted in green (JNK) and yellow (p38), the ERK pathway is highlighted in blue. Potential cross-talk via shared ERK and p38 components is indicated in the overlapped shading. Common transcription factors between all MAPK pathways include Elk1, ATF2, SAP2, and Mnk1/2.


ERKs
Extracellular cell
receptor stimuli
SAPKs
Stress and cytokine
stimuli
MEK1/2
ERK1/2
Elk1
RSK
ATF2
SAP2
Fos
SRF
CBP
p300
MSK1
CREB
DDB2/14
Histone H3
MEK3/6
p38/gIIy
MEF2C
SAP2
CHOP
Elk1
ATF2
MEF2A
MAPK
APK2/3
PRAK
MEK4/7
JNK1/2/3
Jun
ATF2
Elk1
SAP2
JunB
JunD

1.1.2. c-Jun N-terminal kinase

The c-Jun N-terminal kinases (JNK) constitute one of the major groups within the SAPK category. JNKs play a role in responding to various stress-related and proinflammatory stimuli (e.g., cytokines, UV DNA damage, or growth factor deprivation) and are activated by upstream MEK4 and MEK7 kinases. These MEKs are themselves activated by MEKK1-4. Secondary modifiers, such as MLK2 and 3, Tpl-2, DLK, TAO1 and 2, TAK1 and ASK1, and ASK2, may also be involved in the activation pathway and may modulate the response to various stress-linked stimuli. Once activated, JNKs can relocate to the nucleus and phosphorylate the c-JUN proto-oncoprotein (a component of the AP-1 transcription factor). Other transcription factors that may also be phosphorylated, as minor targets, include ATF2, NF-ATc1, HSF1 and STAT3 (*34*).

1.1.3. p38 kinase

The p38s (α, β, γ and δ isoforms), which are also members of the SAPK group, are activated by environmental stress as well as proinflammatory cytokines. Unlike JNK, p38 appears to function by aiding in the induction of TNFα production through macrophages. While the precise mechanism by which p38 is involved is unclear, studies have shown that inhibitors of p38s are able to prevent TNFα production, a result suggesting that p38 may affect TNFα by altering mRNA translation (*34;36*). p38α possesses 50% amino acid identity with ERK2, but unlike ERK2, the p38 proteins do not show appreciable activity in response to mitogenic stimuli (*13*). Instead, activation of the p38 kinases occurs via specific MAPKKKs that are responsive to stressors such as oxidative damage, UV radiation, hypoxia, or ischemia (*6*). These environmental factors lead to the activation of MEKK1-4, MLK2 and 3, DLK, ASK1, Tpl-2 and TAK1, which can then phosphorylate MEKs 3 to 6, leading to

activation of p38. The activated p38 then remains in the cytoplasm or translocates to the nucleus and phosphorylates its known substrates: phospholipase A2, Tau, ATF1 and 2, MEF2A, Sap1, Elk1, NF-κB, Ets1, p53, MSK1, MSK2, MNK1, MNK2, MK2 and MK3 (*34*).

1.1.4. Extracellular receptor regulated kinase

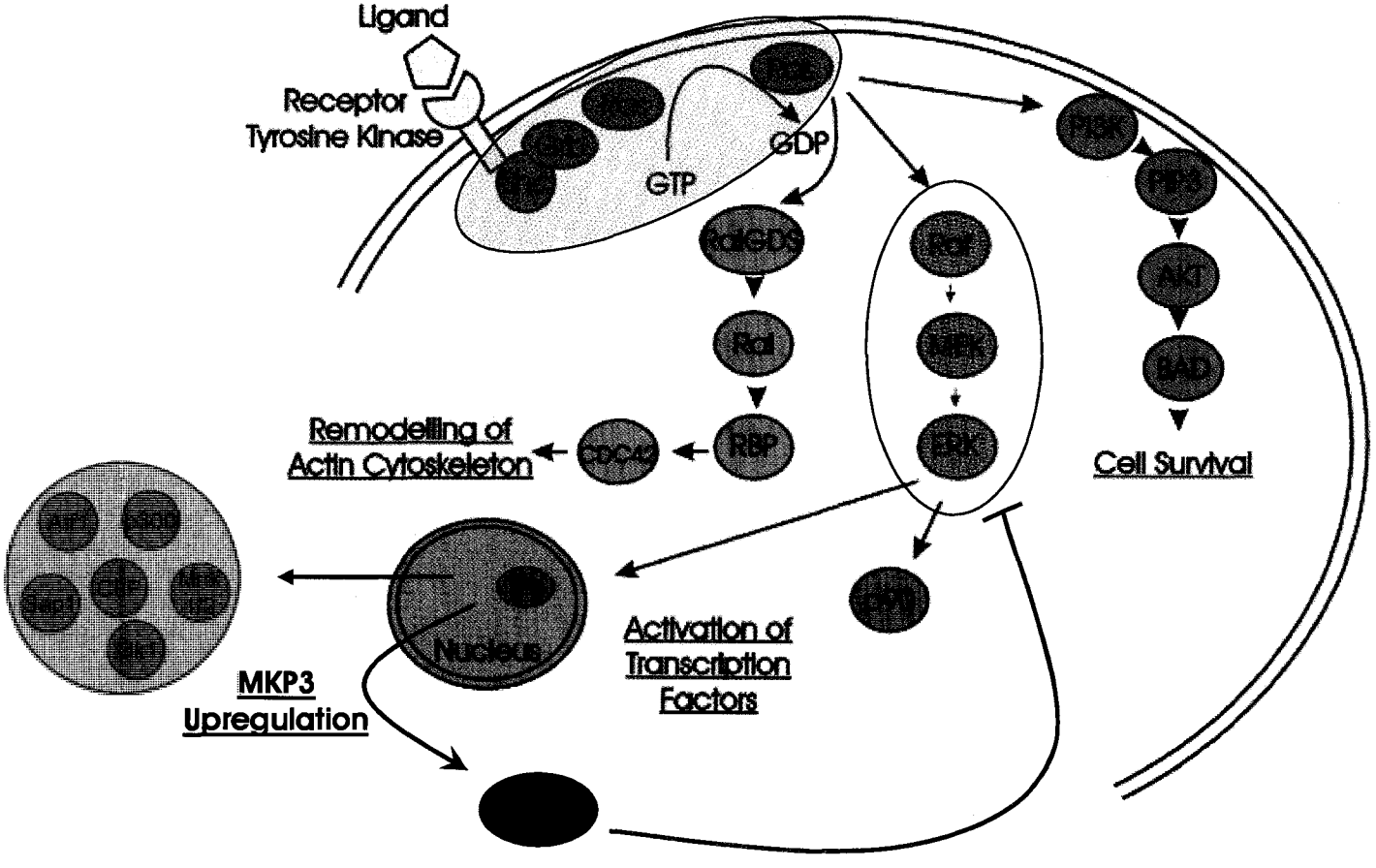
While the SAPKs are activated by cytokines or cellular damage, the ERK family of kinases are primarily influenced by stimulation and activation of surface receptors by growth factors and phorbol esters (*34*). Figure 3 shows ERK activation by the extracellular activation of a transmembrane receptor. This results in an initiating signal that is transduced to the interior of the cell by receptor autophosphorylation and subsequent G-protein recruitment that results in the phosphorylation of ERK effectors, such as RAS (*6;8;37*). The activated RAS protein can then phosphorylate A-RAF, B-RAF and RAF-1 molecules, leading to phosphorylation of MEK1 and MEK2 and ultimately to phosphorylation of the ERK kinases (*38*). When activated, ERK can phosphorylate cytosolic proteins or translocate to the nucleus to further activate transcription factors, an event that ultimately results in a cellular response to the original signalling stimulus (*33*). One of the genes that is upregulated by ERK is *Mkp3* which can act as an inbuilt control that dephosphorylates ERK1 and ERK2, resulting in negative regulation (*11*).

1.2 MAP kinase phosphatases

1.2.1. MAP kinase phosphatases and MAPK pathway regulation

The MAP kinase phosphatases are a specific subgroup of 11 dual specificity phosphatases that share several similar features. Each MKP contains an N-terminal MAP

Figure 3. Regulation of the ERK-MAPK activation pathway. The ERK family of MAP kinases becomes activated following extracellular receptor binding. Ligand/receptor binding and subsequent autophosphorylation leads to G-protein complex formation and results in activation of Ras kinase activity (green highlight). The activated Ras is capable of initiating a phosphorylation cascade (yellow highlight) that ultimately leads to activation of ERK through phosphorylation. Activated ERK can translocate to the nucleus and activate various transcription factors (in grey). The *Mkp3* gene is believed to be one of the many targets of ERK-mediated transcription factors. The production of MKP3 protein negatively regulates the ERK signal and results in self-limited activation of the ERK pathway.


Ligand
Receptor
Tyrosine Kinase
GTP
GDP
PI3K
PIP3
AKT
BAD
Cell Survival
Ral
MEK
ERK
RalGDS
Ral
RBP
CDC42
Remodelling of
Actin Cytoskeleton
Nucleus
Activation of
Transcription
Factors
MKP3
Upregulation

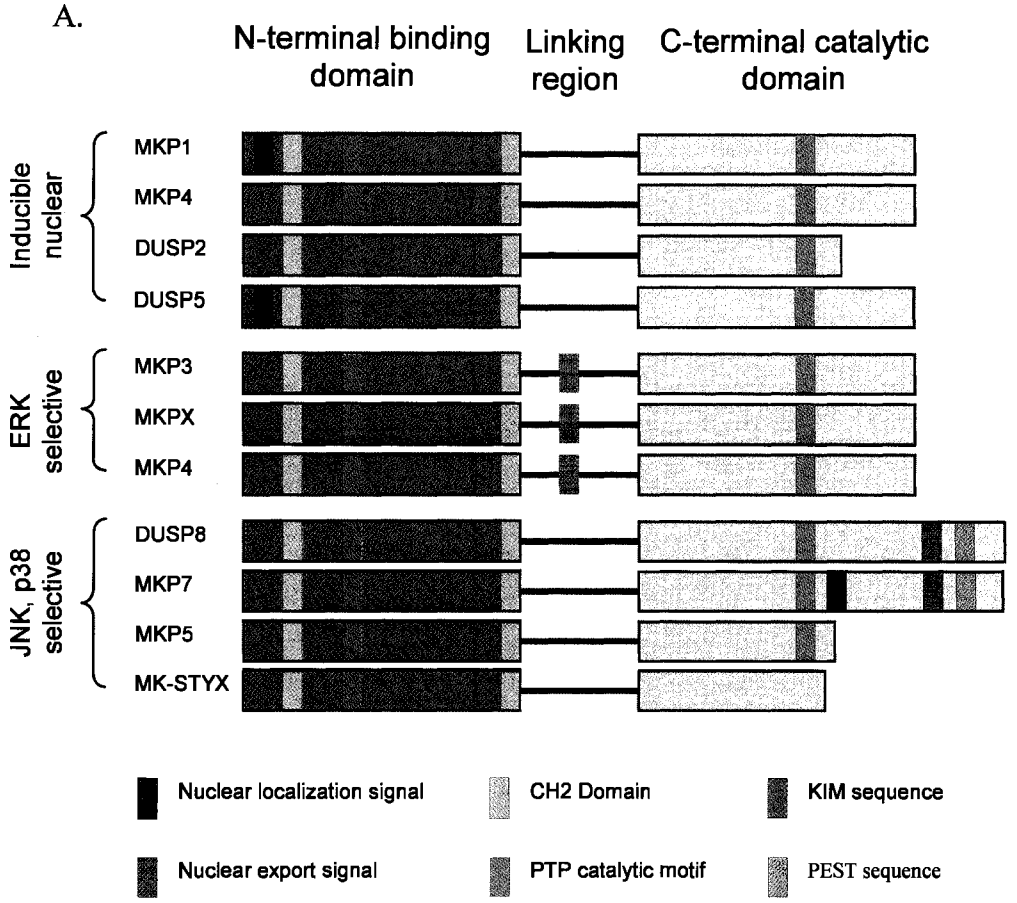
kinase binding domain, which contains two CDC25 constant homology domains (CH2 domains) and a conserved docking region called a kinase interacting motif (KIM) that confers MAPK specificity (*39-42*). Additionally, all MKPs, except for DUSP24/MK-STYX, possess a conserved C-terminal catalytic domain that possesses phosphatase activity (Figure 4, (*43;44*)). The shared catalytic domain permits deactivation of various MAPKs through a direct, acid-catalyzed hydrolysis of the phosphorylated MAPK activation motif. MKP enzyme activity has been characterized in cell-free assays and does not require the recruitment of intermediate docking proteins (*12;15*). This result suggests that the general mechanism for MKP-regulated MAPK deactivation is through intracellular expression leading to direct negative feedback against specific activated MAP kinases (*27;28;30;31*).

1.2.2. MAP kinase phosphatase categories

The MAP kinase phosphatases can be subdivided into three categories based on their intracellular localization and substrate specificity: the inducible nuclear, the JNK/p38 selective, and the cytoplasmic, ERK-selective MKPs (Table I). The first subgroup of MKPs, the inducible nuclear group, consists of the MKP1, DUSP5/PAC1, MKP2, DUSP2 phosphatases. MKP1 may be involved in the regulation of signalling through ERK, JNK and p38 pathways (*45-48*). Other known functions of the groups' members include metabolic homeostasis (MKP1, (*44*)) and positive regulation of MAPK activity in immune cells and maintenance of certain immune responses (PAC1, (*49*)). Next, the JNK/p38 selective group contains DUSP8, MKP5 and MKP7. Members of this MKP subgroup show a cytoplasmic and nuclear distribution (*50*). While the biological functions of most of this group remain largely undetermined, MKP5 has been shown to be involved with JNK regulation and innate or adaptive immunity (*50*). The last group of MKPs is the ERK-selective group that contains MKP3, MKPX and MKP4 enzymes. It is believed that the cytoplasmic, ERK-selective

Figure 4. General structure of the MAP kinase phosphatases and KIM alignment. The MAP kinase phosphatases are distinct subfamily of the protein tyrosine phosphatases, sharing several common characteristics, which have been shown diagrammatically. **A.** The MAP kinase phosphatases contain similar structures that include a conserved catalytic domain and a MAPK-binding N-terminal domain with a conserved CH2 domain and kinase interacting motif (KIM) sequence. **B.** The binding specificity of MKPs results from a charged amino acid sequence with alternating hydrophobicity and hydrophilicity known as a KIM sequence. The known docking motifs of the MKPs are shown with the KIM sequences of the ERK-selective MKPs highlighted in red. ERK binding depends on a KIM sequence with RRLQK or RRLRR sequence.

A.


N-terminal binding domain
Linking region
C-terminal catalytic domain
Inducible nuclear
MKP1
MKP4
DUSP2
DUSP5
ERK selective
MKP3
MKPX
MKP4
JNK, p38 selective
DUSP8
MKP7
MKP5
MK-STYX
Nuclear localization signal
CH2 Domain
KIM sequence
Nuclear export signal
PTP catalytic motif
PEST sequence

B.

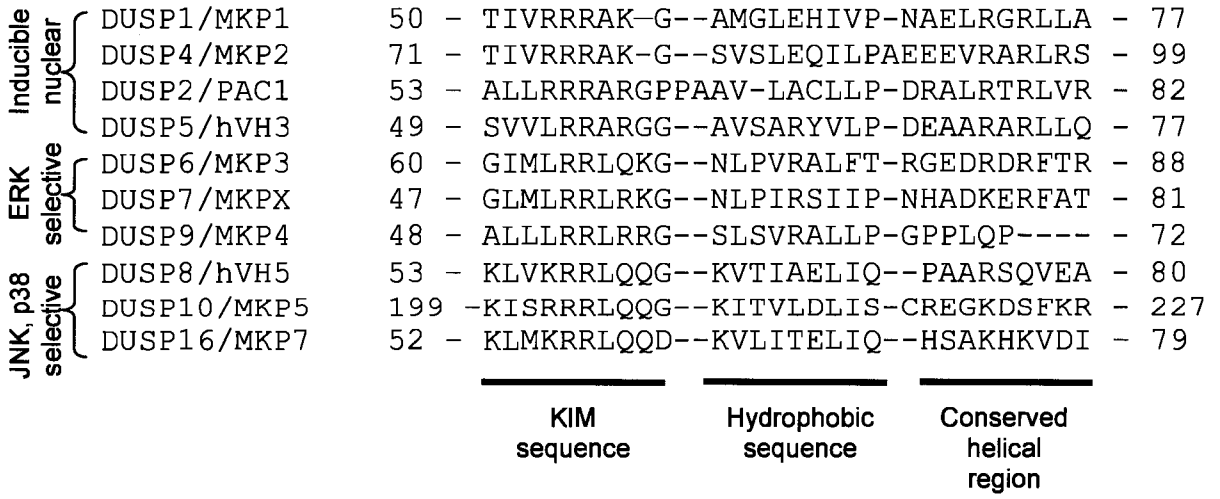

Inducible nuclear
ERK selective
JNK, p38 selective
DUSP1/MKP1 50 - TIVRRRAK-G--AMGLEHIVP-NAELRGRLLA - 77
DUSP4/MKP2 71 - TIVRRRAK-G--SVSLEQILPAEEEVRARLRS - 99
DUSP2/PAC1 53 - ALLRRRARGPPAAV-LACLLP-DRALRTRLVR - 82
DUSP5/hVH3 49 - SVVLRRARGG--AVSARYVLP-DEAARARLLQ - 77
DUSP6/MKP3 60 - GIMLRRLQKG--NLPVRALFT-RGEDRDRFTR - 88
DUSP7/MKPX 47 - GLMLRRLRKG--NLPIRSIIP-NHADKERFAT - 81
DUSP9/MKP4 48 - ALLLRRLRRG--SLSVRALLP-GPPLQP---- - 72
DUSP8/hVH5 53 - KLVKRRLQQG--KVTIAELIQ--PAARSQVEA - 80
DUSP10/MKP5 199 -KISRRRLQQG--KITVLDLIS-CREGKDSFKR - 227
DUSP16/MKP7 52 - KLMKRRLQQD--KVLITELIQ--HSAKHKVDI - 79
KIM sequence
Hydrophobic sequence
Conserved helical region

Table I. MAP kinase phosphatase family members

	Protein	Other names	Chromosomal location	Subcellular localization	Substrate	Physiological function
Inducible nuclear	MKP1	DUSP1,CL100, erp, 3CH134, hVH1	5q34	Nuclear	JNK, p38, ERK	Negative regulator of immune function and metabolic control
	MKP2	DUSP4, Typ1, Sty8, hVH2	8p12-p11	Nuclear	JNK, p38, ERK	Unknown
		DUSP2, PAC1	2q11	Nuclear	p38, ERK	Positive regulator of inflammatory responses
		DUSP5, hVH3, B23	10q25	Nuclear	ERK	Unknown
ERK selective	MKP3	DUSP6, Pyst1, rVH6	12q22-q23	Cytoplasmic	ERK	Negative regulator of ERK signalling
	MKPX	DUSP7, Pyst2, B59	3p21	Cytoplasmic	ERK	Unknown
	MKP4	DUSP9, Pyst3	Xq28	Cytoplasmic	ERK>p38	Essential in placental development and function
JNK, p38 selective		DUSP8, M3/6, hVH5, HB5	11p15.5	Cytoplasmic/Nuclear	JNK, p38	Unknown
	MKP5	DUSP10	1q41	Cytoplasmic/Nuclear	JNK, p38	Functions in innate and adaptive immunity
	MKP7	DUSP16	12p12	Cytoplasmic/Nuclear	JNK, p38	Unknown
	MK-STYX	DUSP24		Cytoplasmic		Unknown, non-catalytic

MKPs play important roles in development; for example, MKP4 has been shown to be essential for placental development and function (*51-53*), while MKP3 has been shown to be extensively involved in developmental regulation and axial patterning during embryogenesis (*27-29;46;54-59*).

1.2.3. MAP kinase phosphatase 3 specificity

Due to its importance in development and as a regulator of ERK signalling, MKP3 is arguably the most studied member of the MKPs (*20;43;44;46;60*). As with other MKPs, MKP3 possesses dual specificity and is able to dephosphorylate phospho-tyrosine and phospho-threonine residues (*13;17*). However, unlike other MKPs, MKP3 is selective for only ERK1 and ERK2 and exhibits no appreciable activity on other MAP kinases (*13*). As Table I shows, most members of the MKP family possess a broad specificity towards the MAP kinases. Therefore, the tight specificity of MKP3 has been the focus of several studies that attempted to determine the mechanisms (both physical and kinetic) that regulate MKP3's ability to dephosphorylate ERK kinases (*15;26;61;62;62;63*). What has been revealed is that to achieve the explicit regulation of ERK signalling pathways, MKP3 relies on several complementary methods that act synergistically, resulting in an unrivalled specificity among the MKP family members (*26;46*). These methods can be broken down and categorized as follows: physical separation/sequestration, temporal separation/kinetic mechanisms and allosteric effects.

1.2.3.a. Physical separation/sequestration

One consequence of maintaining shared components between signalling pathways is that the components in cellular pathways can potentially interact with one another. While there are many occasions when this cross-talk is desirable, binding is scrupulously controlled and non-specific interactions are prevented in the ERK/MKP3 interaction. One mechanism

that prevents cross-talk is sequestration by cell type. This form of sequestration results when the expression of a regulatory protein (i.e., MKP3) is restricted to specific cell types, whereas a general or shared component of signalling pathways is expressed broadly in many tissues (i.e., MAP kinases). The result is that activation or deactivation of the pathway becomes cell type specific. The tissue specific expression of MKP3 has been studied in both chicken and mouse models that indicate that MKP3 is involved in early embryogenesis, but primarily localized to the major developmental loci that are responsible for axial patterning, such as the limb bud and neural tube (*27-30;46;54-59*). Furthermore, in addition to cellular localization, MKP3 displays a subcellular localization in the cell cytoplasm (*17;64-66*). This form of sequestration by MKP3's cytoplasmic localization is made possible by a nuclear export signal (NES), a leucine and glycine rich region (-^{164}*VLGLGGLRISSD*^{175}-) in the MKP3 amino acid sequence (*64*). The resulting localization allows MKP3 to interact specifically with ERK contained within the cytoplasm and it has been suggested that the MKP3 and ERK binding may result in further cytoplasmic compartmentalization of ERK. This effect can prevent ERK's nuclear entry and activation of downstream initiation factors present in the nucleus (*40;46*).

Binding specificity, which is not technically a form of sequestration but an important mechanism of physical separation, allows for selectivity between MAPKs and other signalling components through a system of modular docking motifs. MKP3 possesses several motifs that signal ERK specificity. First, MKP3 possesses two CDC25 constant homology domains that facilitate MAP kinase binding; however, MKP3 also possesses a unique -^{65}RRLQK^{69}- sequence within its kinase interacting motif (Figure 4B). The system of charged residues, which was identified by Tanoue and coworkers (*39*), specifically targets MKP3 to an oppositely charged surface region found on ERK1 and ERK2 and prevents JNK or p38

binding. More recently, Zhou (*67*) has identified additional binding regions on MKP3 at a.a. position 101-107 (-^{101}NSSDWNE^{107}-), which can increase the specificity of MKP3 binding to the ERK1 and ERK2 kinases, as well as a region at 364-367 that appears to function similarly to ERK-binding -FXFP- like sequence (-^{364}FTAP^{367}-). Each of these motifs is selective for ERK and prevents non-specific binding.

1.2.3.b. Kinetic mechanisms

While the physical sequestering mechanisms available to MKP3 allow for the physical separation of signalling components, the specific kinetic properties of an enzyme cascade driven pathway lead to temporal separation that aids MKP3's substrate specificity. It is believed that MKP3 responds to activated ERK1 and ERK2 through a negative feedback loop, which occurs when activated ERK translocates to the nucleus and initiates the transcription of various gene products. The *Mkp3* gene is a target and therefore *Mkp3* mRNA can become upregulated in tandem with ERK activation and lead to a subsequent rise in MKP3 protein levels. The timing of protein translation coincides with the peak periods of ERK activity and can not only generate a temporal gradient to ERK deactivation (which has been shown to be essential in proper chick and xenopus development (*46;68-70*)), but non-specific MKP3 interactions are also limited due to the abundance of cellular ERK at the time MKP3 is expressed.

1.2.3.c. Allosteric and inhibitory effects

The final mechanism involved in MKP3 specificity is allosteric enzyme activation, which is linked to the presence ERK1 and ERK2 (*12;16;53*). This allostery is in contrast to other protein tyrosine phosphatases (PTP) and even other related MKPs within the ERK-selective group that can act as general phosphatases against a range of substrates (Table I; (*60;71-73*)). The specificity MKP3 shows towards ERK1 and ERK2 can be studied using *in*

vitro assays using para-nitrophenol phosphate (pNPP, a small aryl phosphate). While other members of the MKP family can dephosphorylate small molecules, studies show that MKP3 possesses very little basal activity towards pNPP in the absence of ERK (*15*). However, the addition of ERK can lead to a more than 100-fold increase in MKP3 catalytic efficiency, which indicates an allosteric activation of the MKP3 catalytic domain by ERK. The ERK-allostery prevents non-specific phosphatase activity in the absence of the *in vivo* substrate (*61;63;74*).

1.2.4. Structure of MKP3

The structures of both MKP3's N- and C-terminal domains (a.a. 1-154 and 204-347 regions, respectively) have been determined and provide an insight into MKP3's ability to bind and become allosterically activated in the presence of ERK. The structure of MKP3's N-terminal binding domain was solved using high resolution NMR (*75;76*) and showed that the ERK binding domain possesses five alpha-helices surrounding an open, twisted five-strand beta-sheet. This structure is characteristic of a Rossmann fold (a motif common in many enzyme binding domains showing beta-alpha-beta-alpha-beta alternation). Following the structure determination, Farooq *et al. (75;76)* tested the function of the N-terminal domain, but similar to Camps' original work (*12*), the binding domain was tested using experiments that included ERK to activate the full-length MKP3 protein (*75*). Increasing inhibition of the activated full-length MKP3 was observed with increasing concentrations of the MKP3 N-terminal domain. This inhibition of ERK-activated MKP3 was due to competitive binding for ERK between the N-terminal domain and the full-length MKP3. Several deletion mutants (including regions 1-154, 1-178, 1-221 and a full-length C293S non-catalytic mutant) were also tested using the same assay and all mutants showed an ability to inhibit MKP3 phosphatase activity. The results of the study indicated that the

relative ability of each protein to bind ERK (leading to deactivation of MKP3) was 1-154 < 1-179 < 1-221 < C293S. This result implied that although the a.a. 1-154 sequence contained the ERK binding domain, the ability to bind ERK was enhanced by the presence of C-terminal amino acids and regions important for ERK binding likely exist outside the 1-154 N-terminal domain sequence for which the NMR structure was determined.

The C-terminal domain structure was also determined (*74*), but unlike the N-terminal domain's solution structure, the C-terminal domain was determined by x-ray crystallography. Stewart and coworkers crystallized the MKP3 catalytic domain in the absence of ERK binding (i.e., MKP3's inactive form) and postulated that the allosteric activation exhibited once ERK binds to MKP3's N-terminal domain may be due to realignment of one or two amino acid residues in the catalytic triad (composed of aspartic acid 262, cysteine 293 and arginine 299). These residues showed alternate orientations compared to the related VHR protein (which is a general phosphatase). The first of these residues was arginine 299 (R299), which was oriented in a manner that prevented its side chain from interacting with substrate phosphoryl groups. Interpretation of the x-ray crystallization data was complicated by an artefact that blocked the catalytic pocket with a chloride ion. This ion was observed adjacent to R299 and could have influenced the arginine's side chain orientation. Furthermore, no other biochemical data to date have provided supportive evidence of a role for R299 in MKP3's allosteric activation. A second, more critical observation was that aspartic acid 262 (D262) resides on a random coil, flexible loop structure ~20 Angstroms away from the active site cleft. This distance is normally too far away from the active site for the aspartic acid to participate as an acid catalyst during dephosphorylation. However, independent studies by Fjeld and Rigas (*61;63*) both measured the effect of pH on MKP3 activity through steady state analysis and pH titration. Their experiments showed that MKP3's ERK-induced k_{cat}

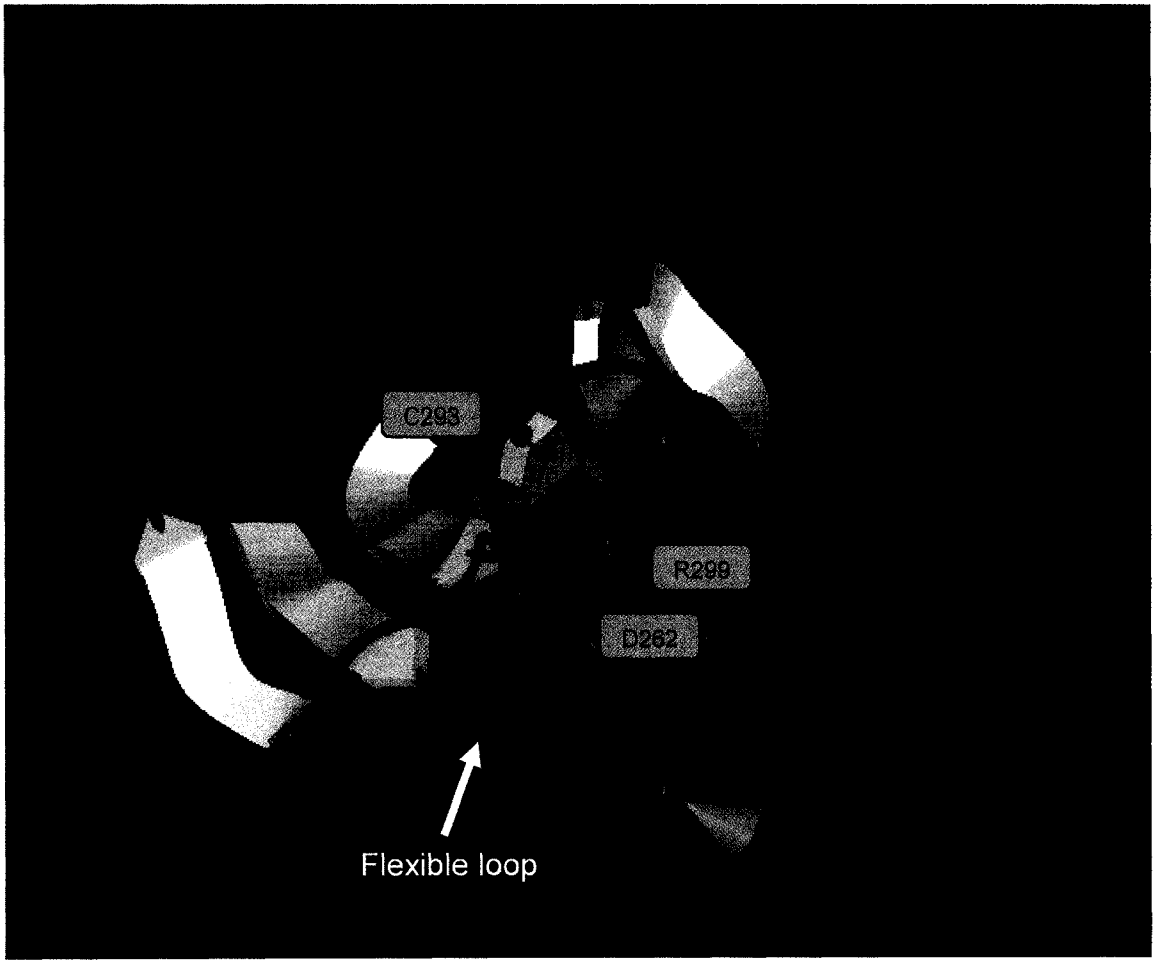
depends on an ionisable acidic amino acid residue with a pK_a consistent with the involvement of an aspartic acid.

Combining the data from the two research groups leads to the current view of MKP3 activation. Upon ERK binding to the N-terminal domain, the C-terminal catalytic domain undergoes an allosteric restructuring that allows highly efficient catalysis. The structural reorganization is believed to occur at the flexible loop that supports D262 (*61;74*) that may undergo loop closure to bring D262 into proximity with the other active site residues (R299 and C293; Figure 5, (*12;61;63;74*)). The structures for the structurally homologous phosphatases, MKP5 and DUSP5, have recently been determined and the data show that neither enzyme possesses the same distorted geometry at their catalytic aspartic acid residues that MKP3 exhibits (*50*). Furthermore, DUSP5 shows a constitutive activity and can dephosphorylate pNPP in the absence of ERK (*26*). These findings have been used as evidence that MKP3 function is related to ERK recognition and allosteric activation that involves repositioning the D262 residue (*26;50*).

1.2.5. Models of the full-length MKP3 and MKP3/ERK complex

No research group has successfully determined the structure of full-length MKP3 or MKP3 complexed with ERK, presumably due to the difficulty in obtaining a co-crystallized ERK/MKP3 complex or a structure of an artificially activated MKP3. Researchers have found other novel methods to study the N- and the C-terminal MKP3 domains. Using NMR and mass spectrometric techniques, both Farooq et al. and Zhou et al. have shown that MKP3 domains may interact with one another or with ERK (*75;77*). Using an NMR magnetization transfer method, Farooq et al. was able to show that the MKP3 N-terminal domain can bind a PAC1 catalytic domain (a homologue of the MKP3 catalytic domain). The binding was

Figure 5. Structure of the MKP3 catalytic domain. The x-ray crystal structure of the MKP3 catalytic domain was determined by Stewart et al. (PBD entry **1MKP**; (*74*)). This PDB was used to draw the 3D structure using Insight II (Accelrys Software Inc). Stewart showed that the catalytic triad was misaligned compared to constitutively activated phosphatases. The critical residue that regulates the efficiency of MKP3 catalysis is thought to be D262 (shown in blue), which resides on a flexible loop (identified by arrow). The current MKP3 activation model suggests that ERK binding, on MKP3's N-terminal domain, induces a loop closure to bring the D262 residue into proximity with the remaining C293 and R299 residues of the catalytic triad (shown in grey).


C293
R299
D282
Flexible loop

primarily through ionic interactions in oppositely charged residues and the NMR estimate of the dissociation constant (K_D) was ~100 μM (*75*).

More recently, Zhou's group has used hydrogen/deuterium (H^+/D^+) exchange experiments to map the interface between the MKP3 and ERK2 molecules. This study has provided binding data to complement the known structural data (*67;77*). The deuterium exchange method relies on the ability of labile amide or hydroxyl protons to exchange with deuterons contained in a solvent (commonly D_2O). The rate of H^+/D^+ exchange depends on solvent exposure and the lability of an amino acid residue's amide protons, which is influenced by participation in salt bridging, hydrogen bonding, or by being buried in the core of the protein. Each of the aforementioned factors can reduce the rate of deuterium exchange. Surface exposed residues, not participating in hydrogen bonds or salt bridges, will show significantly greater exchange than buried or otherwise protected residues. This differential exchange can be detected by comparing mass differences before and after a binding interaction by mass spectrometry (*67*). With MKP3, H^+/D^+ exchange was measured before and after incubation with ERK2. Not surprisingly, the ERK-binding, KIM domain (residues 65-69) displayed strong solvent protection following ERK incubation. This result indicates that the residues were likely surface exposed in the unbound MKP3 form, but lie alongside the ERK surface and hence are shielded once MKP3 and ERK2 bind (*77*). While the shielding of the KIM sequence (which binds to ERK) was expected, the study by Zhou led to the surprising finding that nearly the entire C-terminal domain sequence exhibited shifts in exchange upon ERK binding. The implication is that a more extensive ERK-induced conformational change occurs than previously believed. The main findings from the C-terminal residues were the discovery of an FXFP-like sequence (-^{364}FTAP^{367}-). This region

showed reduced deuterium exchange, suggesting that the sequence is therefore involved in ERK-induced change following ERK binding (*42;77*). Additionally, an MKP3 specific -^{101}NSSDWNE^{107}- sequence, that had not been previously identified, was shown to exhibit decreased deuterium exchange in the presence of ERK2. The -^{101}NSSDWNE^{107}- sequence has now been observed to bind opposite to the ERK kinase active site and is believed to be involved in tethering MKP3 to ERK (*77*).

1.2.6. MKP3 activity in the presence of solvents

To determine the mechanistic basis for MKP3's allosteric activity, two separate research groups have used small molecules to modify the activity of MKP3. These groups, led by Zhou (*15*) and Fjeld (*15;61*), identified glycerol and DMSO as two co-solvents that could activate MKP3 in the absence of ERK. The effect of glycerol was measured on both inactive and ERK-activated MKP3 samples. The results showed that the addition of glycerol was able to increase the catalytic efficiency (k_{cat}/K_M) of MKP3 in the absence of ERK. A 50% glycerol addition to inactive MKP3 led to a ~12 fold increase in k_{cat}/K_M over MKP3's basal activity, but the addition of 40% glycerol to ERK-activated MKP3 also led to a ~2.5 fold increase to MKP3's ERK-activated k_{cat}/K_M. While the authors of the study implied that the glycerol-induced activity in the absence of ERK was due an allosteric-like induction, the view did not fit the data showing that the ERK-activated MKP3 activity could be improved by additional glycerol. Unfortunately, the authors have not revisited the results since Zhou's first reporting. Interestingly, glycerol at a similar concentration to Zhou's experimental conditions has been shown to increase SHP1 activity (SHP1 is a member of the PTP family to which the DUSPs are related). This effect is due to the glycerol-induced unbinding of an inhibitory SH2 binding domain (*Src* homology domain) that blocks the SHP1 catalytic site.

The possibility that glycerol could affect MKP3 activity through a glycerol-mediated unbinding of an inhibitory domain was not explored by Zhou.

A second group, led by Fjeld, has observed a much greater effect with the use of DMSO. Fjeld *et al.* showed that at the ERK-activating effect on MKP3 could be replicated by using a 30% DMSO concentration. MKP3 activity was measured in the presence of increasing DMSO concentrations or DMSO in combination with ERK2. Dephosphorylation of pNPP, as a chromogenic substrate, was then used to detect MKP3 activity. At a 15% DMSO concentration, MKP3 showed less than optimal activity. This activity was increased by the addition of ERK2 to the reaction. At 30% DMSO, MKP3 activity had reached a maximal activity that was not increased even by the addition of ERK. The activity was comparable to fully activated MKP3 using ERK2 as an activator alone. This finding indicated that 30% DMSO was sufficient to induce a similar activity to ERK and that presumably a similar mechanism of MKP3 activation was involved (i.e., DMSO induced a structure similar to the ERK-induced MKP3 activation).

In both cases, either using glycerol or DMSO, MKP3 activations were said to be due to a repositioning of the critical D262 side chain, either due to glycerol's effect on hydrophobic residues or DMSO's effect as a "chemical chaperone". Unfortunately, the very different effects on MKP3 activity make it unlikely that both solvents act through the same mechanism. Glycerol was only able to induce a small fraction of the total ERK-activated MKP3 activity, while DMSO replicated the high phosphatase activity of ERK-activated MKP3.

1.2.7. In vivo regulation of MAPK pathways by MKP3

The major mechanism involved in deactivation of MAPK pathways is dephosphorylation. Therefore, due to the ERK's role in development and the ERK-specific

function of MKP3, there is an implied role for MKP3 in regulating growth and development of the cell. Studies of early embryogenesis have identified potential physiological functions (*46;55;57*) and show that MKP3 can act as a negative regulator of fibroblast growth factor (FGF) induced signalling during early embryogenesis (*27*). Specific examples of developmental regulation that are dependant on the proper expression of, and regulation by, MKP3 include limb and neural tube development in chick, mice and presumably all other vertebrates (*27;28;31;57;58;78*).

In addition to its role as a regulator of the ERK-cascade in development, current data also indicate that MKP3 affects the severity and potential outcome of pancreatic adenocarcinoma. The link between MKP3 and pancreatic adenocarcinoma is suggested by the *Mkp3* gene locus, which is in a region of frequent allelic loss in pancreatic adenocarcinomas (*24;79;80*). Furthermore, many intimal hyperplasias (which may progress to adenocarcinomas) overexpress MKP3. This can indicate a potential role for the protein in regulating cell growth prior to cellular transformation (*79*). While it is speculated that deregulation of MKP3, and by association, deregulation of the ERK pathway, could result in the onset of more invasive adenococarcinomas (*18;19;21;22;80*), the evidence for this role is relies on results showing a loss of MKP3 protein expression in adenocarcinoma. However, preliminary work by Aubin showed that MKP3 is present in high-grade pancreatic adenocarcinomas (Figure 1). To reconcile this result, experiments are required to determine a mechanism for MKP3's apparent loss of function.

1.3. Objectives

The goal of my research was to define the physical properties of MKP3 and determine how MKP3 functions as an ERK cascade regulator. This work aimed to reconcile the conflicting data that show a role for MKP3 in deactivating ERK and the unexplained overexpression of MKP3 in pancreatic adenocarcinomas that fails to deactivate ERK1 and ERK2.

A potential mechanism for the loss of function may be the uncharacterized MKP3 N- to C-terminal binding that was suggested by the ability of MKP3's N-terminal domain to bind the PAC1's C-terminal domain (*75*). If binding results in inhibition of catalytic activity, then overexpression of MKP3 may lead to unregulated binding and an overall loss of catalytic function. To test this hypothesis, binding assays were required to test if MKP3's N- and C-terminal domains could truly bind. An ERK-free assay that could activate the MKP3's catalytic domain in the absence of ERK was needed to test if N-terminal domain removal could increase MKP3 activity. In addition, a measure of activity for the catalytic domain (in the absence of ERK), with increasing concentrations of the N-terminal domain, was needed to test for inhibition. In summary the objectives for this study were:

1. Produce full-length MKP3 and N- or C- terminal domain truncated mutants,
2. Determine if binding constants between MKP3's N- and C-terminal domains are consistent with interdomain binding,
3. Develop an ERK-free MKP3 phosphatase activity assay, and
4. Measure MKP3 catalytic domain activity in the presence of the N-terminal domain to test for inhibition.

2.0 Materials

Materials used in *Mkp3* cloning, including *Mkp3* cDNA, PCR2.1 plasmid, INVαF[1] cells, carbenicillin and primers used for *Mkp3* DNA manipulations were purchased from Invitrogen (Burlington, ON). Additional reagents required for cloning include Midi prep kits (Qiagen; Mississauga, ON), Advantage HF-PCR kits (Clontech; Mountain View, CA), pET-15b plasmid (Novagen; Madison, WI), BL21(ΔE3) cells (EMD Bioscience; La Jolla, CA) and Quickchange XL mutagenesis kits (Stratagene; La Jolla, CA). Reagents used for buffers and cell media including: ammonium bicarbonate, arginine, dimethyl formamide, dimethyl sulfoxide, EDC, EDTA, ethanolamine, guanidine-HCl, iodoacetamide, LB broth, NaCl, NHS, para-nitrophenol phosphate, sodium deoxycholic acid, Tris-base, Tris-HCl and urea were obtained from Sigma-Aldrich (Oakville, ON). Additional buffer constituents: acetone, acetonitrile and methanol were supplied by EMD Chemicals (Gibbstown, NJ). DNAse I and IPTG were obtained from Roche (Mississauga, ON). His-Trap metal affinity columns, ECL detection reagent and ECL anti-mouse and anti-rabbit antibodies were obtained from GE Healthcare (Piscataway, NJ). Recombinant ERK2 and anti-$(His)_6$ RAb were acquired from Upstate/Millipore (Danvers, MA). Anti-MKP3 MAb was acquired from R&D Systems (Minneapolis, MN). Peptides including: ERK2's biphosphorylated active site (-^{179}DHTGFLT-pEY-pVATR^{190}-), ERK2's MKP3 binding motif (-^{314}EQYYDPSDEPIAEA^{327}-), and MKP3's kinase interacting motif (-^{57}AIPGIMLRRLQKGNLPVR^{74}-) were obtained from JPT Peptide Technologies (Acton, MA). The post-KIM (-KK^{77}FTRGEDRDRFTRRCGTDTVVL^{97}-) peptide was purchased from Sigma-Aldrich (Oakville, ON). JNK1 was donated by Dr. Richard Blouin (L'Université de Sherbrooke; Sherbrooke, QC). Normal human pancreatic tissues (obtained at surgery from

an anonymous organ donor) and immortalized human pancreatic duct cells (HPDE4, HPDE6) were provided by Dr MS Tsao (University of Toronto; Toronto, ON). Human pancreatic adenocarcinoma cell lines (CRL1420, CRL1469, CRL1682, CRL1687, CRL1837, HTB79, HTB80, HTB134, and HTB147) were obtained from the American Type Culture Collection (Manassas, VA).

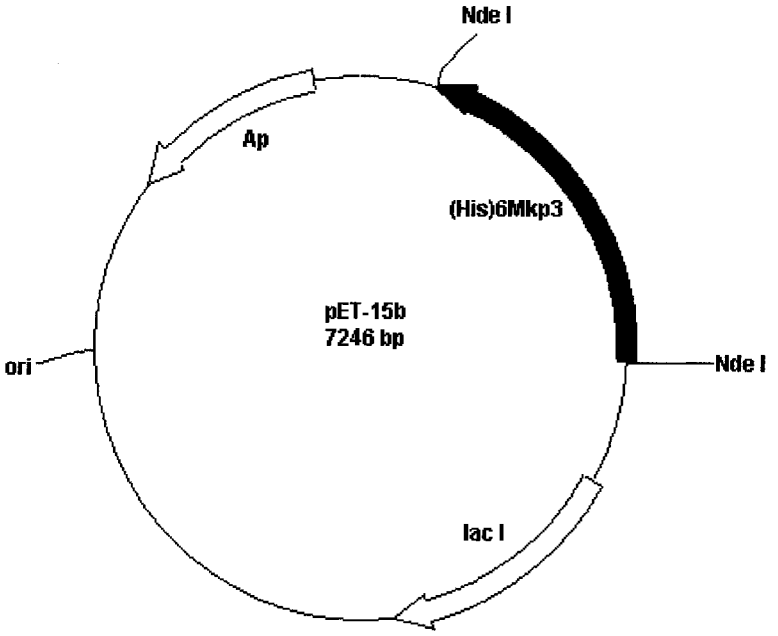
2.1 Methods

2.1.1. Cloning

A human *Mkp3* cDNA construct was purchased from Invitrogen (Genestorm collection, Genbank accession, **NM_1946.2**). The entire *Mkp3* coding sequence was excised by PCR amplification using the Advantage HF PCR kit (Clontech) with forward (5^1-GGA ATT CCA TAT GAT AGA TAC GCT CAG AC-3^1) and reverse (5^1-TAG TTT ACA TAT GTC ACG TAG ATT GCA GAG AGT CC-3^1) primers that were engineered to contain single NdeI sites (underlined). The PCR product was purified by agarose gel electrophoresis prior to ligation into a pCR2.1 vector and transformation into INVαF^1 cells for colony screening and plasmid purification (by QIAprep mini-prep kit, as per manufacturer's instructions). To express protein in *E. coli*, the *Mkp3* coding sequence was excised from the pCR2.1 vector using NdeI cleavage and subcloned into NdeI-linearized pET-15b using T4 DNA ligase. The new pET-15b:*Mkp3:WT* construct (Figure 6A) was transformed into BL21(ΔE3) cells. The resulting recombinant MKP3 protein possessed an N-terminal $(His)_6$-tag and a thrombin cleavage site conferred by the pET-15b plasmid in addition to the MKP3 primary structure (Figure 6B). Similar methods were used to generate

Figure 6. Construction of the pET-15b:*Mkp3* and MKP3 proteins. *Mkp3* cDNA corresponding to either the full-length, wild type MKP3 (*Mkp3:WT*), amino acids 1-154 (*Mkp3:NT*), or amino acids 155-381 (*Mkp3:CT*) was inserted into a pET-15b vector using the NdeI restriction enzyme cleavage sites. **A.** Plasmid diagram. The plasmid was constructed so that *Mkp3* DNA was placed between two NdeI restriction sites, with a preceding, N-terminal $(His)_6$-tag followed by a thrombin cleavage site. MKP3 expression was controlled by an IPTG-inducible T7 promoter. **B.** MKP3:WT amino acid sequence. The MKP3 amino acid sequence is shown with the preceding N-terminal (His)6-tag (in bold) and thrombin cleavage site (underlined). MKP3:NT sequence shaded in green and MKP3:CT sequence shaded in yellow. **C.** MKP3 protein sequences. The newly generated plasmids were transformed into BL21(ΔE3) cells for protein expression. The proteins that were produced were expressed as either full-length MKP3 (MKP3:WT) or as truncation mutants containing amino acid residues 1-154 (MKP3:NT) or 155-381 (MKP3:CT). Each protein was produced containing an N-terminal $(His)_6$-tag and thrombin cleavage site.

A.

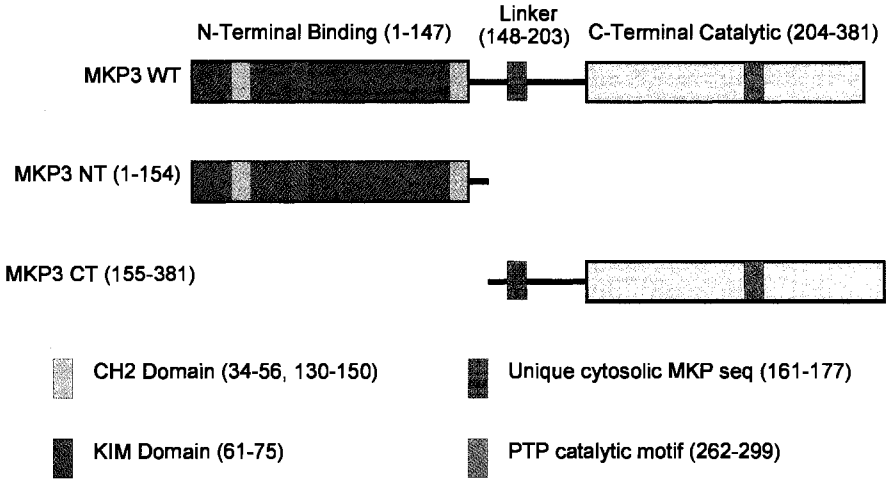


B.

MGSSHHHHHHSSGLVPRGS

PLPVLGLGGL RISSDSSSDI ESDLDRDPNS ATDSDGSPLS
NSQPSFPVEI LPFLYLGCAK DSTNLDVLEE FGIKYILNVT PNLPNLFENA
GEFKYKQIPI SDHWSQNLSQ FFPEAISFID EARGKNCGVL VHCLAGISRS
VTVTVAYLMQ KLNLSMNDAY DIVKMKKSNI SPNFNFMGQL LDFERTLGLS
SPCDNRVPAO OLYFTTPSNO NVYOVDSLOS T

C.



plasmids corresponding to MKP3's N terminal 1-154 amino acid sequence and the C-terminal 155-381 amino acid sequences (*Mkp3:NT* and *Mkp3:CT*; Figure 6C). The cDNA sequences were subcloned into a suitable pET-15b, *E. coli* vector at the NdeI restriction enzyme site and transformed into BL21(ΔE3) cells for protein expression. Each protein contained an N-terminal $(His)_6$-tag and thrombin cleavage site in addition to the described amino acid sequences.

Additional catalytic domain mutants that were produced for binding studies were generated using the Quickchange XL mutagenesis kit. The kit was used, as per the manufacturer's instructions, on the *Mkp3:CT* cDNA to generate specific amino acid substitutions at residues E248, N267 and D318. Each targeted residue was substituted with an alanine. The mutations of the *Mkp3:CT* cDNA resulted in *Mkp3:CT/E248A*, *Mkp3:CT/N267A* and *Mkp3:CT/D318A* cDNAs. Each cDNA was then inserted into a pET-15b plasmids and transformed into BL21(ΔE3) cells for protein production.

2.1.2. Protein expression

Recombinant proteins were produced from BL21(ΔE3)/MKP3 starter cultures that were established by overnight growth at 37°C in a 50 mL conical tube containing 12 mL Luria-Bertani (LB) broth and 50 μg/mL carbenicillin. The next morning, 10 mL of the culture was added to 1 L of 50 μg/mL carbenicillin supplemented, LB broth in a 2.8 L Fernbach flask. The production culture was grown in an Innova 4000 incubator (New Brunswick Scientific; Edison, NJ) with shaking (150 RPM at 37°C) until an OD_{600} between 0.4 and 0.5 was reached (~2-2.5 hours). At this point, the expression of MKP3 was induced by the addition of IPTG (1 μM final concentration). The cells were then allowed to grow for a further 2 hours before being harvested by centrifugation using an Avanti JA-25 centrifuge

(Beckman; Mississauga, ON) at 8,000x g for 30 minutes. Cell pellets were then stored frozen at -80°C.

2.1.3. Protein purification

Recombinant MKP3 proteins were isolated from harvested cells by a combination of sonication and membrane solubilization. The cell pellet from a 1 L culture was thawed on ice and then resuspended in 70 mL of Tris lysis buffer (20 mM Tris, 137 mM NaCl, 1 mM EDTA, pH 7.6). Lysozyme was added to the cell suspension to a final concentration of 0.25 μg/mL and incubated at room temperature for 30 minutes. The cells were then lysed by sonication using a Sonic Dismembranator Model 100 (Fisher Scientific; Nepean, ON). Sonication was performed on ice for 2 minutes (using 1 minute sonication pulses followed by 1 minute rests on ice) and the insoluble material was pelleted by centrifugation (20,000x g, 30 minutes, 4°C). The insoluble fraction was resuspended in 70 mL of fresh Tris buffered saline (TBS, 20 mM Tris, 137 mM NaCl, pH 7.6) containing 2% sodium deoxycholate and DNAse I (50 μg/mL). The solution was incubated at room temperature for 30 minutes and centrifuged at 25,000x g at 4°C for 30 minutes. The recovered pellet was washed once with 70 mL TBS and centrifuged at 25,000x g at 4°C for 30 minutes. Lastly, the insoluble pellet was dissolved in 10 mL of TBS containing 6 M guanidine-HCl (pH 7.6).

2.1.4. Metal affinity chromatography

Guanidine denatured $(His)_6$-tagged protein (in TBS containing 6 M guanidine-HCl, pH 7.6) was purified by metal affinity chromatography using an AKTA-FPLC system (GE Healthcare; Piscataway, NJ) with a 5 mL HisTrap™ HP metal affinity column (GE Healthcare; Piscataway, NJ). The dissolved MKP3 proteins were applied at a flow rate of 5 mL/min. Non-specific proteins were eluted using a linear gradient of 0 to 50 mM imidazole

in TBS over 15 minutes at 5 mL/min (75 mL). The $(His)_6$-tagged MKP3 was then eluted in a 5 minute, 5 mL/min (25 mL) elution with 250 mM imidazole in TBS. The isolated proteins were then refolded as described in section 2.1.6.

2.1.5. MKP3 refolding assays

MKP3 refolding conditions were determined on a 96-well plate (round-bottom, Corning Lifesciences; Lowell, MA). Ten μL of purified protein solution at 1 mg/mL were diluted into 190 μL of cold refolding buffers (1:20 dilution) and left overnight at 4°C. The refolding buffers consisted of TBS (20 mM Tris, 137 mM NaCl, 5 mM DTT, pH 7.6) supplemented with guanidine and arginine in nine different formulations:

#1. 0.3 M Guanidine, 0 M Arginine
#2. 0.3 M Guanidine, 0.4 M Arginine
#3. 0.3 M Guanidine, 0.8 M Arginine
#4. 0.6 M Guanidine, 0 M Arginine
#5. 0.6 M Guanidine, 0.4 M Arginine
#6. 0.6 M Guanidine, 0.8 M Arginine
#7. 0.9 M Guanidine, 0 M Arginine
#8. 0.9 M Guanidine, 0.4 M Arginine
#9. 0.9 M Guanidine, 0.8 M Arginine

Assessment of MKP3 precipitation was performed using 200 μl of sample in each microtitre plate well. The precipitation was assessed visually under 10 x magnification using a five point scale with one point indicating no precipitation and five points indicating extensive precipitation. Assessment of MKP3 activity was performed using the method described in section 2.1.16.

2.1.6. MKP3 refolding by dialysis

MKP3 refolding was performed using gradual dialysis. MKP3 protein in 6 M guanidine solution was adjusted to 0.25 mg/mL and then dialysed against TBS (pH 7.6) containing 1 mM EDTA, 5 mM DTT and 8 M urea. The removal of urea was performed by

changing the dialysis buffer every four hours using fresh TBS (with 1 mM EDTA and 5 mM DTT at pH 7.6) with a lower concentration of urea. The concentration of urea was decreased by half at each buffer change. After five buffer changes, the solution had reached 0.25 M urea. At this point, the concentration of urea was decreased to 0 M. The refolded MKP3 was stored at 4°C in TBS buffer containing 1 mM EDTA and 5 mM DTT at pH 7.6. Under these storage conditions, MKP3:CT showed less than 20% loss of activity after >1 month. MKP3:WT was less stable and showed a ~ 20% loss of activity after 1 week. MKP3 proteins were used for biochemical assay within one week after refolding.

2.1.7. SDS-PAGE and densitometry

SDS-PAGE (*81*) was performed using protein solutions mixed 1:1 with SDS-loading buffer and 7.5% or 12% poly-acrylamide cross-linkage. Electrophoresis was performed using a Mini-Protean 3 cell (Bio-Rad; Hercules, CA) at 200 mV. Gels were stained with Coomassie Blue. Protein purity was determined by densitometry using a Bio-Rad Gel-doc XR with Quantity One 4.5.2 software. Measurement was performed by determining the MKP3 protein band density as a percentage of the total lane density.

2.1.8. Western blotting

Western transfers onto PVDF membrane were performed overnight at 30 Volts using a Bio-Rad Mini-Trans-Blot tank and Tris-glycine buffer (25 mM Tris, 192 mM glycine, pH 8.3). For detection of recombinant MKP3 proteins, either a polyclonal anti-$(His)_6$ antibody at 1:2,000 dilution or a monoclonal MKP3-specific antibody at 1:1,000 dilution, was used as the primary antibody. HRP-conjugated ECL anti-rabbit or anti mouse antibody (1:30,000 dilution) was used as the secondary antibody in each case. Detection was performed using

the ECL chemiluminescent detection reagent as described by the supplier (GE Healthcare; Piscataway, NJ).

For Western analysis of HeLa and pancreatic cell lines, the electrophoresis was performed for 1.5 h at 100 V. The separated proteins were transferred to Immobilon-P membrane and detected as described above. The blot was then stripped with Western Re-Probe Solution (EMD; La Jolla, CA) and re-probed with an anti-hRAN antibody (Cell signalling technologies/Millipore; Danvers, MA) to detect the loading control.

2.1.9. Northern blotting

Pancreatic cell lines were lysed and 10 μg of total RNA were fractionated on a 1% agarose formaldehyde-MOPS gel, which was processed for Northern blot hybridization. The blot was first probed with a 1.25 Kbp human MKP-3 HindIII/XhoI cDNA fragment, stripped and re-probed with a human β-actin cDNA.

2.1.10. In-gel tryptic digest and mass spectroscopy

In-gel tryptic digests were performed by separating protein samples using SDS-PAGE as described above. Bands of interest were excised from the gel, diced to a uniform size (approximately 1 mm^3) and then mixed with acetonitrile. The gel slurry was dried and resuspended in 1 mL of 100 mM ammonium bicarbonate containing 10 mM DTT to cleave disulphide bonds. Iodoacetamide was added (50 mM, final concentration) to cap cysteine residues and prevent disulphide bonds from reforming. Gel pieces were washed and rehydrated in 1 mL of 50 mM ammonium bicarbonate and the proteins were digested by the addition of trypsin (20 μg). Peptide fragments were identified using an LTQ-FT mass spectrometer system (Thermo-Finnigan; Waltham, MA) coupled to an Agilent 1100 series HPLC using a Zorbax 300SB-C18 capillary HPLC column (Agilent technologies; Santa

Clara, CA). PEAKS software (Bioinformatics Solutions; Waterloo, ON) was used to process the mass spectrometry data.

2.1.11. CD spectroscopy

CD spectroscopy experiments were performed using a Jasco J-810 spectrometer (Jasco; Easton, MD). Proteins were assessed in a round cuvette (0.02 cm pathlength) at a concentration of 0.5-1.0 mg/mL in TBS (pH 7.6). Wavelength data were collected over 178-260 nm at 50 nm/min. Five scans were acquired and the average was used for analysis. Due to interference at lower wavelengths, the CD data were cropped to include only the region encompassing 190-260 nm before analysis with the Dichroweb web analysis tool (*82-84*).

2.1.12. Surface plasmon resonance

Surface plasmon resonance experiments were performed on a BIAcore 3000 (GE Healthcare; Piscataway, NJ). All runs were performed in triplicate at 25°C. MKP3 (WT, NT and CT) and MAPK (ERK2 and JNK) proteins were covalently bound to BIAcore CM5 (carboxymethyl-dextran-coated) sensor chips. Covalent cross-linking of the proteins to the sensor chip surface was performed by activating the CM5 surface with a 1:1 mixture of 0.4 M 1-ethyl-3-[3-dimethylaminopropyl] carbodiimide hydrochloride (EDC) and 0.1 M N-hydroxysuccinimide (NHS). The EDC/NHS mixture was injected at a 5 µl/min flow rate for 10 minutes. The protein of interest was then immobilized by a second injection across the activated surface at 5 µl/min until a desired concentration of bound ligand was achieved (approximately 200 RU). The remaining activated carboxyl groups were then inactivated by an injection of 1 M ethanolamine at a flow rate of 5 µl/min for 10 minutes.

SPR binding data, for ERK2 and JNK1, were obtained using MKP3:WT in the mobile phase. SPR data for MKP3:WT, MKP3:NT and MKP3:CT, were obtained using

MKP3:CT in the mobile phase. MKP3:WT or MKP3:CT was injected at 20 µl/min at increasing concentrations. Binding and dissociation curves derived from SPR signals were fitted using non-linear analysis by the BIAEvaluation 4.0 software package to obtain association and dissociation constants. SPR traces were fit using a 1:1 Langmuir binding model: A + B ↔ AB, simultaneously fitting the dissociation and association phases of the SPR data to:

$$\frac{\partial B}{\partial t} = -(k_a * A * B - k_d * AB)$$

where B_0=Rmax (max binding response), and

$$\frac{\partial AB}{\partial t} = (k_a * A * b - k_d * AB)$$

where $AB_0 = 0$ (the amount of bound AB complex at t = 0, the start of analyte A injection).

2.1.13. Binding site characterization

The MKP3 binding site was characterized using peptide fragments derived from regions of MKP3 or ERK2 corresponding to the ERK2 biphosphorylated active site (-^{179}DHTGFLT-pEY-pVATR^{190}-), the ERK2 binding motif (-^{314}EQYYDPSDEPIAEA^{327}-), the MKP3 kinase interacting motif (-^{57}AIPGIMLRRLQKGNLPVR^{74}-) and the MKP3 post-KIM region (-KK^{77}FTRGEDRDRFTRRCGTDTVVL^{97}-). The two additional N-terminal lysine residues in the post-KIM peptide were added to enhance surface immobilization in surface plasmon resonance experiments. All peptides were individually immobilized on CM5 sensor chips and MKP3 was used as an analyte to detect binding as described above. MKP3:CT and MKP3:CT mutants (MKP3:CT/E248A, MKP3:CT/N267A and MKP3:CT/D318A) were used in the mobile phase and tested at increasing concentrations as described above.

2.1.14. Determination of protein concentration

MKP3 concentrations were determined using an UV absorbance method (280 nm wavelength, (*85*)). 100-150 µl of protein solution were placed into the wells of a 96 well, UV-transparent plate (Corning Lifesciences; Lowell, MA). Theoretical extinction coefficients were determined using the ExPASy web server (*86*). The coefficients were 31,900 $M^{-1}cm^{-1}$ for MKP3:WT, 15,720 $M^{-1}cm^{-1}$ for MKP3:NT, and 16,180 $M^{-1}cm^{-1}$ for MKP3:CT. The path length of the sample was determined, depending on the volume of protein solution used, by dividing the volume of solution by the area on the bottom of the microtitre plate well (0.3165 cm^2).

2.1.15. Ionic strength calculation

Ionic strength effects on MKP3 were tested by increasing the NaCl concentration in solution using a mixture of water and 1M NaCl. The final solution ionic strength was determined using the formula:

$$I = \frac{1}{2}\sum_i z_i^2 [x_i]$$

where z is the charge of an ion and x is its molar concentration.

2.1.16. MKP3 phosphatase activity assays

Assays were performed in 96-well, flat bottom plates (Corning Lifesciences; Lowell, MA) using a version of the original assay described by Camps, *et al.* (*12*). Unless otherwise noted, all assays were performed in triplicate using a 300 µl reaction volume. Assays were performed using either 12 µg of inactive or 1.2 µg activated enzyme. These amounts correspond to final assay concentrations of 0.9 µM or 0.09 µM for MKP3:WT and 1.45 µM or 0.15 µM for MKP3:CT. The assay buffer contained 2 mM pNPP in 10 mM Tris, 50 mM NaCl, at pH 7.6. This buffer was made at a 2x concentration to allow for MKP3 and

DMSO/buffer addition. Following 60-minute incubation at 25°C, the absorbance at 405 nm was detected using a Genios plate reader running the XFluor 4 software program (Tecan, Durham, NC). An extinction coefficient of 18,200 $M^{-1}cm^{-1}$ was used to measure the production of nitrophenolate following pNPP hydrolysis. Measurement of pNPP hydrolysis was performed in the linear range of the assay (5 μg fully activated MKP3 assayed over 120 minutes, (*12*)).

2.1.17. Kinetic analysis of MKP3

Kinetic analysis was performed in the presence and absence of ERK2 or DMSO. MKP3:WT and MKP3:CT under ERK2 or DMSO activating conditions were tested using 0-10 mM pNPP in 10 mM Tris at pH 7.4 with the ionic strength adjusted to 25 mM. MKP3:WT and MKP3:CT under inactive conditions were tested using 0–50 mM pNPP 10 mM Tris buffer at pH 7.4 with the ionic strength adjusted to 160 mM. All assays were performed in a 150 μl volume and experiments were performed in triplicate.

For activated MKP3 assays, the hydrolysis of pNPP was measured following the addition of ERK2 at a 2:1, ERK:MKP3 ratio (0.19 μM final concentration) or 50 μl DMSO (33%). For inactive MKP3 assays, 50 μl of Tris buffer were added in place of ERK2 or DMSO. The resulting curves were analysed by non-linear regression in Graphpad Prizm 4 (Graphpad Software; San Diego, CA). The data was fit to the Michealis-Menton equation:

$$Y=Vmax*S/(K_M+S)$$

where K_M=Michealis-Menton constant and S=substrate concentration.

k_{cat} was determined by the equation:

$$k_{cat}=Vmax/E$$

where E=enzyme concentration.

2.1.18. MKP3:NT inhibition

Assays were performed using the standard assay conditions with 33% DMSO and 0.09 μM MKP3. MKP3 inhibition was tested using an increasing concentration of MKP3:NT (0-0.19 μM) that was added to MKP3:CT or MKP3:WT assays. To determine the K_i of MKP3:NT on MKP3:CT, the activity data were fit to a sigmoidal dose-response curve to calculate IC_{50} using the following equation in Graphpad Prizm 4:

$$Y=\text{Bottom} + (\text{Top-Bottom})/(1+10^{(\text{LogIC50-X})*(\text{HillSlope})})$$

where X is the log of MKP3:NT concentration and Y is the activity of MKP3:CT. Top was set to the highest observed activity (DMSO-activated MKP3:CT), Bottom was set to the lowest observed activity (inactive MKP3:CT), Hill slope gives the largest absolute value of the slope of the curve

Using the IC_{50} value, K_i was determined using the Cheng-Prusoff equation (*87*):

$$Ki=\frac{IC_{50}}{1+\frac{[Substrate]}{K_M}}$$

where [Substrate] was the experimental pNPP concentration and K_M was the Michealis-Menton coefficient for pNPP, which was determined in section 3.4.4 (1.43 mM).

2.1.19. Statistical analysis

Statistical analyses were performed using a student t-test. A p-value < 0.05 was used to determine significance.

3.0 Results

3.1 MKP3 production

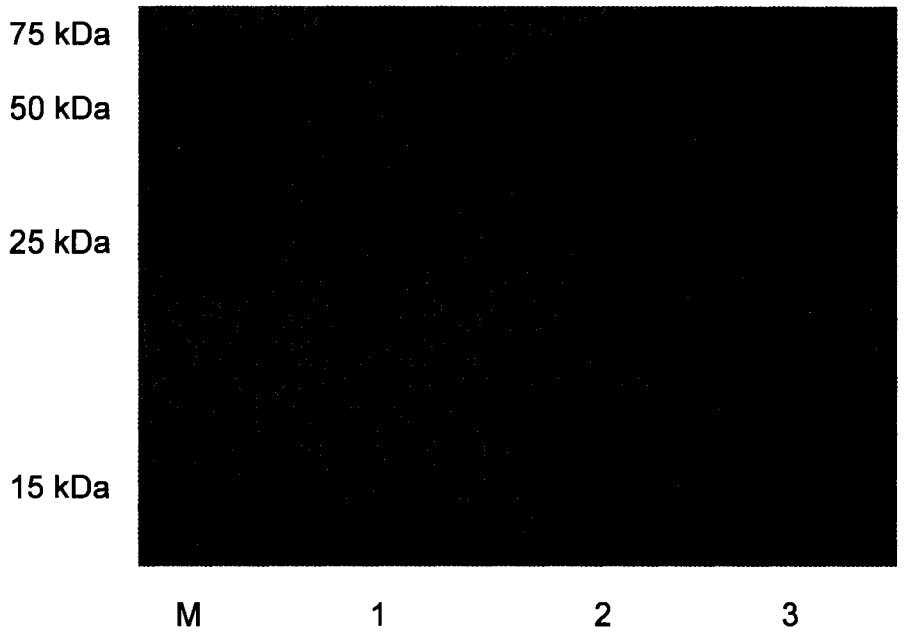
3.1.1. MKP3 cloning

To produce recombinant full-length MKP3 for this study, an *Mkp3* cDNA was inserted into the pET-15b plasmid and BL21(ΔE3) cells. Sequencing of the *Mkp3* coding sequence inserted into pET-15b (pET-15b:*Mkp3:WT*) showed a complete and correctly orientated *Mkp3* cDNA (published in (*88*)). Similar methods were then used to generate the MKP3's N-terminal (a.a. 1-154) and C-terminal domains (a.a. 155-381) truncation mutant proteins, which resulted in pET-15b:*Mkp3:NT* and pET-15b:*Mkp3:CT* plasmids. Sequencing of the *Mkp3:NT* and *Mkp3:CT* coding sequences showed that the inserted DNA was correctly oriented and unmodified. All *Mkp3*-containing plasmids were transformed into BL21(ΔE3) cells which resulted in BL21(ΔE3)/MKP3:WT, BL21(ΔE3)/MKP3:NT and BL21(ΔE3)/MKP3:CT production cells.

3.1.2. Identification and verification of recombinant proteins

To ensure that the correct MKP3 proteins were produced following cloning, the plasmid-encoded proteins were identified by Western Blot and confirmed by mass spectrometry. The recombinant *E. coli* cells that produced MKP3:WT, MKP3:NT, or MKP3:CT proteins were grown as described in Methods (section 2.1.2). After IPTG induction, samples of the total cellular protein were separated by SDS-PAGE and transferred to PVDF membrane for Western detection using an anti-$(His)_6$ antibody (Figure 7). The apparent molecular weight for each immunodetected band was calculated as 45 kDa for BL21(ΔE3)/MKP3:WT, 18 kDa for BL21(ΔE3)/MKP3:NT and 31 kDa for BL21(ΔE3)/MKP3:CT. These values compared well

Figure 7. Western detection of pET-15b-encoded MKP3. BL21(ΔE3) cells containing *Mkp3* cDNAs were grown as detailed in Methods (section 2.1.2). Cells were harvested and 12 μg of the total cellular protein were used for Western analysis using an anti-$(His)_6$ antibody, as detailed in Methods (section 2.1.8). Lane M. $(His)_6$-tagged molecular weight markers, Lane 1. BL21(ΔE3)/MKP3:CT, Lane 2. BL21(ΔE3)/MKP3:NT, Lane 3. BL21(ΔE3)/MKP3:WT.


75 kDa
50 kDa
25 kDa
15 kDa
M
1
2
3

to their theoretical molecular weights at 44.5 kDa for MKP3:WT, 19.4 kDa for MKP3:NT and 27.0 kDa for MKP3:CT. This result indicated that the desired proteins were expressed.

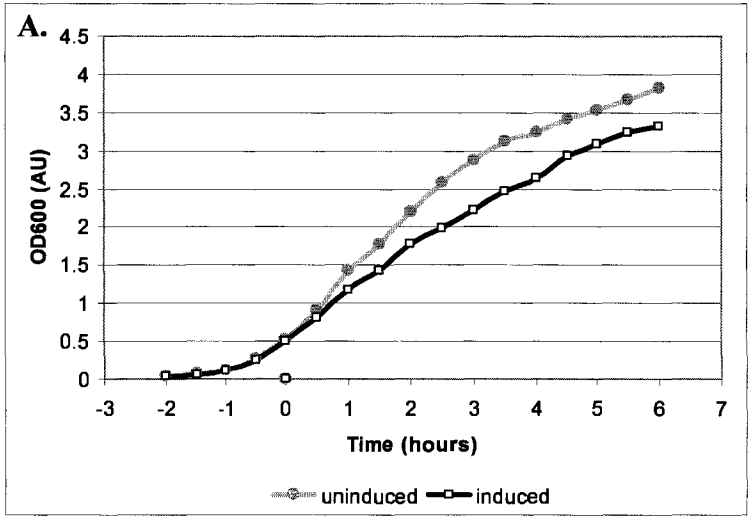
The protein identities were further confirmed by mass spectrometric analysis. Total cellular protein was separated by SDS-PAGE as in the preceding experiment. The Western detection results were then used to select bands for in-gel tryptic digestion and tandem HPLC, mass spectrometry as described in Methods (section 2.1.10). The resulting peptide fingerprint results were searched against a peptide database, which showed that the tryptic peptides matched the predicted fingerprints for each of the three cloned MKP3 variants (data not shown). The result confirmed that MKP3 proteins were correctly expressed.

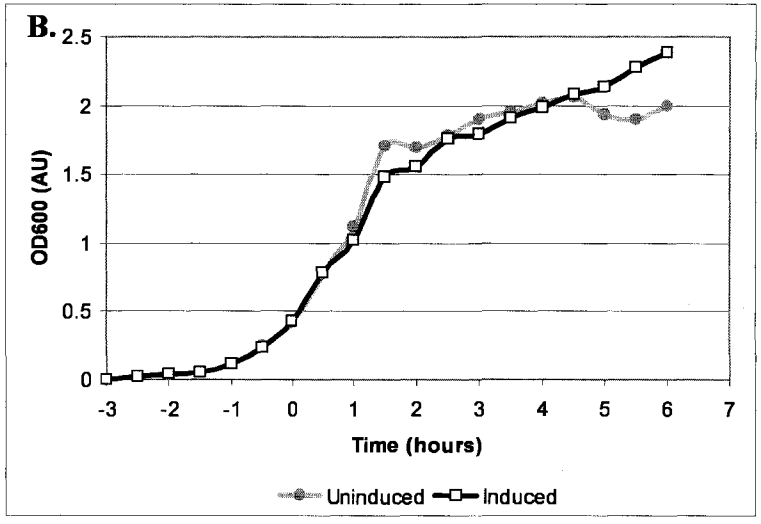
3.1.3. Optimizing expression time to improve MKP3 yield

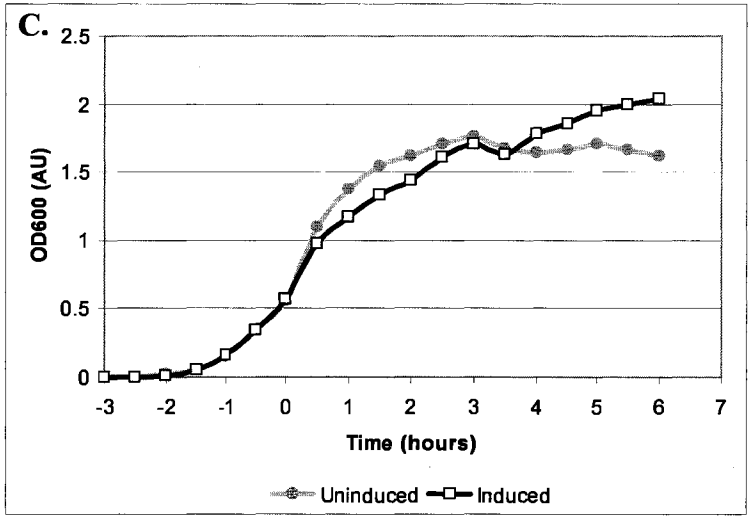
When the volumes of the *E. coli* growth cultures were increased to produce protein for biochemical analyses, the purifications (following the procedure of Zhou *et al.* (*13;15*)) resulted in poor yields. The amount of MKP3 was not adequate for the planned analyses and the growth conditions were tested to determine if the viability of cells was compromised by MKP3 production. This was done by measuring the density of BL21(ΔE3)/MKP3 cells in culture and comparing the amount of protein produced over time.

In the first experiment, MKP3-producing cells were grown as detailed in Methods (section 2.1.2). To measure cell density with time, the OD_{600} value was determined at 0.5 hour intervals. The OD_{600} readings showed that the density of *E. coli* cells increased similarly in all BL21(ΔE3)/MKP3 cell lines until IPTG induction (Figures 8A-C). Following induction, MKP3:WT expressing cells (Figure 8A) displayed a lag in growth compared to their uninduced counterpart (OD_{600} at 6 hours was 3.25 and 3.8 AU,

Figure 8. Effect of IPTG induction on growth rates. MKP3 producing cells were grown as described in Methods (section 2.1.2). The OD_{600} values for each culture were recorded at 0.5 hour intervals. At an OD_{600} of ~0.5 AU (t=0), either 10 mL of 100 mM IPTG (induced) or 10 mL water (uninduced) were added to culture media. The BL21(ΔE3) cells were then grown for an additional six hours. **A.** BL21(ΔE3)/MKP3:WT cells. **B.** BL21(ΔE3)/MKP3:NT cells. **C.** BL21(ΔE3)/MKP3:CT cells.


A.
OD600 (AU)
Time (hours)
uninduced
induced

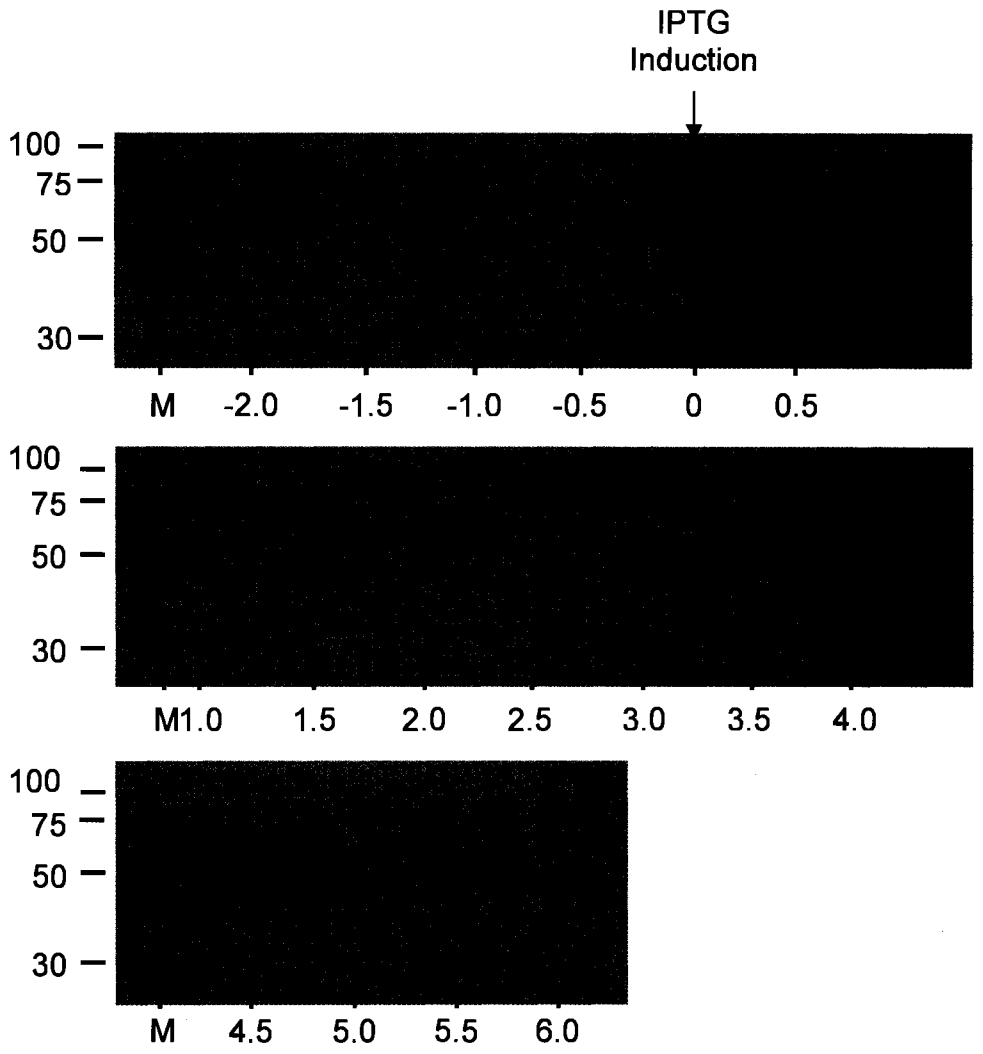

B.
OD600 (AU)
Time (hours)
Uninduced
Induced


C.
OD600 (AU)
Time (hours)
Uninduced
Induced

for induced and uninduced, respectively). The induced MKP3:NT and MKP3:CT producing cells (Figures 8B and C) did not show the same lag. The nearly identical curves in the later two samples, under induced and uninduced conditions, indicated the recombinant proteins were not detrimental to *E. coli* growth, but the addition of the MKP3:NT- and MKP3:CT-producing plasmids did appear to result in slower growing in BL21(ΔE3) cells compared to MKP3:WT-producing cells. However, in both MKP3:NT- and MKP3:CT-producing cells showed no growth when comparing the uninduced and induced conditions. All cell types showed continued growth following IPTG induction (and hence recombinant protein production), which was determined by the increase in OD_{600} readings. Therefore, it was concluded that *E. coli* cells were not grossly affected by MKP3 production.

To determine an expression time that would lead to the best MKP3 yield, MKP3 production at the various timepoints was compared using Western analysis. Figure 9 shows the results from the MKP3:WT analysis. As expected, MKP3 was not produced prior to IPTG addition. However, the protein was detected at 0.5 hours after IPTG induction. The amount of MKP3 peaked at ~1.0-1.5 hours after IPTG induction and then began to decline. MKP3 was virtually undetectable at the end of sample collection (6 hours post-induction). It is possible that the observed loss of recombinant protein with time was due to reabsorbtion of the protein by *E. coli* cells as time progressed. The results in Figures 8A and 9 used cells harvested from the same experiment. These data showed that cells continued to grow after IPTG induction, but yielded less MKP3 with extended induction times. To avoid a loss of MKP3, the data were used to select a two-hour induction period following IPTG addition.

Figure 9. Production of MKP3 as a function of time. BL21(ΔE3) cells expressing MKP3:WT were grown as detailed in Methods (section 2.1.2). One mL samples of cell culture were harvested at 0.5 hour intervals and analysed by Western blot using an anti-$(His)_6$ antibody. Figure lanes are labelled by the time when samples were obtained (in hours pre or post IPTG-induction, t=0 indicates the time of IPTG addition). Molecular weight markers are shown in lanes marked M.


IPTG
Induction
100
75
50
30
M
-2.0
-1.5
-1.0
-0.5
0
0.5
100
75
50
30
M1.0
1.5
2.0
2.5
3.0
3.5
4.0
100
75
50
30
M
4.5
5.0
5.5
6.0

3.1.4 Purification of MKP3 proteins

MKP3 production was performed using a two-hour induction time. After a preliminary fractionation, MKP3 was found to be expressed in large quantities, but produced as insoluble *E. coli* aggregates and not as soluble protein. It was decided that purification of the aggregates and a subsequent protein-refolding step was preferable over further manipulation of growing conditions to obtain soluble protein. Therefore, it was necessary to develop a protocol to solubilize, refold and verify that MKP3 was properly refolded.

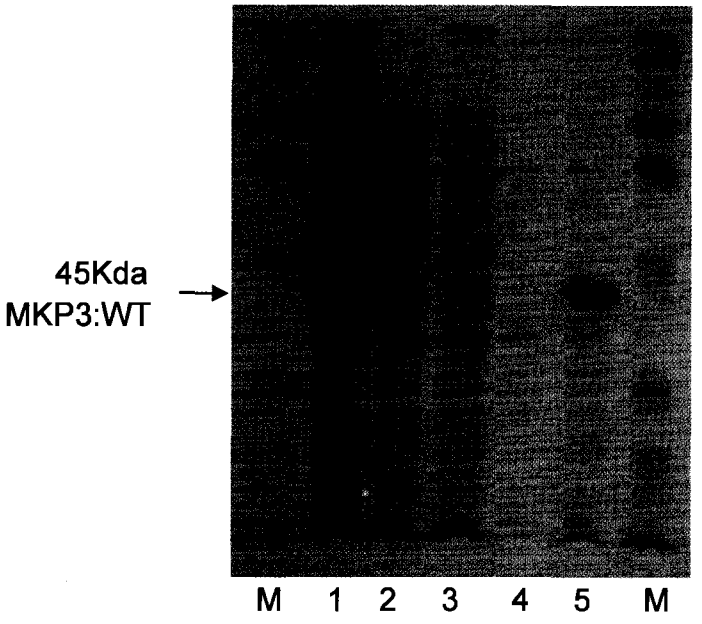
MKP3 purification was performed using a general extraction technique for protein aggregates in four steps (*89;90*): 1. Cell lysis by sonication, 2. Membrane and lipid removal using sodium deoxycholate, 3. Recovery and wash of the protein pellet, and 4. Guanidine solubilization of the insoluble protein. The results of the extraction for MKP3:WT are shown in Table II and Figure 10A. Soluble, Na^+ deoxycholate, and pellet wash fractions were excluded leaving the final MKP3 aggregate fraction, which was determined to be up to 188 mg from 546 mg total protein in 1 L recombinant *E. coli* or ~0.19 mg MKP3/mL culture.

Table II. Isolation of total MKP3:WT protein

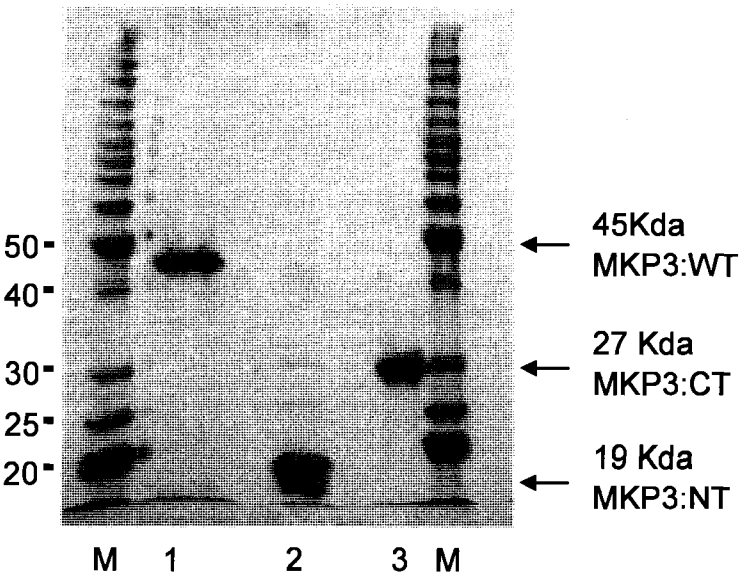
	Concentration (mg/mL)	Volume (mL)	Protein (mg)	Percent total protein
Whole cell fraction	7.8	70	546	100%
Soluble fraction	2.7	70	190	35%
Na^+ deoxycholate extraction	2.2	70	153	28%
Pellet wash	0.4	70	26	5%
Guanidine Solubilization	18.8	10	188	34%
Metal affinity Chromatography	5	15	75	14%

Figure 10. Recovery and purification of MKP3. SDS-PAGE of MKP3 purification. **A.** Insoluble protein extraction from BL21(ΔE3)/MKP3:WT. Cells were grown and processed according to Methods. 15 μl of each sample except for the final guanidine solubilized fraction (2 μl loading) was separated on 12% SDS-PAGE and stained with Coomassie blue. Lane M. Molecular weight markers, Lane 1. Non-lysed, whole cell fraction, Lane 2. Soluble protein fraction, Lane 3. Sodium deoxycholate wash, Lane 4. TBS wash, Lane 5. Guanidine-solubilized insoluble protein fraction, Lane M. Molecular weight markers. **B.** Metal affinity purification of guanidine solubilized inclusion bodies. MKP3 proteins were purified by metal affinity chromatography as described in Methods. 5 μg of each purified protein was separated on 12% SDS-PAGE. The gel was stained by Coomassie Blue. Lane M. Molecular weight markers, Lane 1. MKP3:WT, Lane 2. MKP3:NT, Lane 3. MKP3:CT, Lane M. Molecular weight markers.

A.


45Kda
MKP3:WT
M 1 2 3 4 5 M

B.


50
40
30
25
20
45Kda
MKP3:WT
27 Kda
MKP3:CT
19 Kda
MKP3:NT
M 1 2 3 M

Analysis of the guanidine-solubilized fraction by SDS-PAGE and densitometry (Figure 10A, lane 5) indicated that the insoluble protein fraction contained approximately 90% MKP3:WT protein, with the remaining 10% consisting of non-specific *E. coli* proteins.

Impurities were removed using metal affinity chromatography. Samples of each MKP3 variant, following affinity purification, are shown in Figure 10B. The SDS-PAGE results were analysed by densitometry, which showed that MKP3 proteins (WT, NT and CT) were purified to >95% purity.

3.2. MKP3 refolding

3.2.1. Rapid protein solubility and activity testing

The conditions used to recover MKP3 resulted in denatured proteins that required refolding. A method was developed to screen refolding conditions using samples of the denatured MKP3:WT. MKP3 was refolded by dilution and the samples were tested for protein solubility and DMSO-inducible phosphatase activity (both are characteristics of the properly refolded MKP3 (*61*)).

3.2.1.a. MKP3 solubility

MKP3 solubility was measured following the dilution of denatured MKP3 into refolding buffers. The buffers contained concentrations of guanidine and arginine, described in Methods (section 2.1.5). The results, listed in Table III, show that MKP3 remained soluble in buffers that contained moderate to high concentrations of guanidine and arginine (guanidine was increased down columns; arginine was increased towards the right in rows). While the results indicated what buffers (#5, 6, 8, and 9) led to increased MKP3 solubility, a second assay was still required to test for MKP3 activity.

3.2.1.b. MKP3 activity

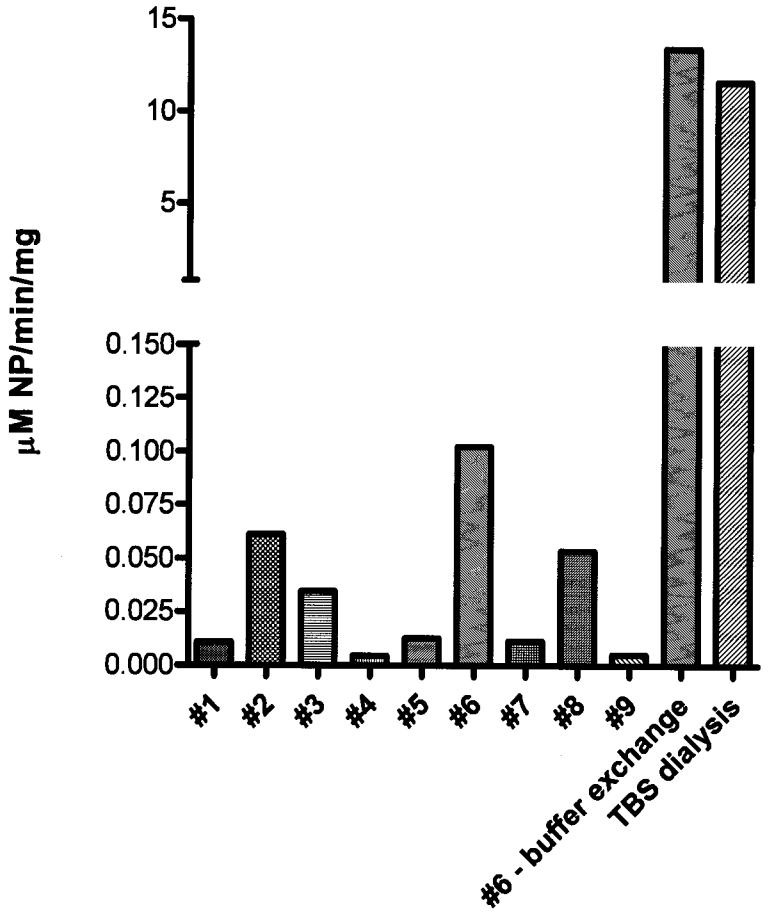
To ensure that the increased solubility of MKP3 in the preceding result was due to proper refolding, rather than the chaotropic or denaturing property of guanidine, MKP3 from each refolding condition was tested for phosphatase activity using *p*-nitrophenol phosphate (pNPP) and DMSO (as a chemical activator of MKP3 activity (*61*)). The assays were performed as detailed in Methods (section 2.1.16). The result shown in Figure 11 indicates that the MKP3 solution refolded in Buffer #6 possessed the greatest specific activity at 0.10 $\mu M\ min^{-1}\ mg^{-1}$. However, this was less than the previously reported activity of MKP3 (~20-25 $\mu M\ min^{-1}\ mg^{-1}$, (*12*)). The difference was thought to be due to the high ionic strength that resulted from residual arginine (800 mM) and guanidine (300 mM) in the MKP3 solution. To test if the residual salts were responsible for a loss of MKP3 activity, MKP3:WT in Buffer #6 was buffer-exchanged into Tris-buffered saline using centrifugal concentrators. MKP3 activity was then retested and the results are shown in Tables III and Figure 11 as "#6 – after TBS buffer exchange". Because of the arginine and guanidine removal, MKP3 showed an improved activity of 13.4 $\mu M\ min^{-1}\ mg^{-1}$, a 130-fold increase. Given the effect of arginine and guanidine on MKP3 activity, an additional method was tested to refold MKP3 in their absence. A dilute sample of MKP3 (0.25 mg/mL) in 6 M guanidine was refolded by dialysis into TBS solution. This buffer was to shown to yield a highly active protein. Refolding was assessed with the result shown in Tables III and Figure 11 under "TBS dialysis". The slow dialysis and the low concentration of MKP3 improved protein solubility without the need for guanidine or arginine. The activity of the dialysed product, 11.6 $\mu M\ min^{-1}\ mg^{-1}$, was also comparable to the sample that had been dilution-refolded and then buffer-exchanged.

Table III. MKP3 refolding screen using precipitation and activity assays		
#1. 0.3 M guanidine 0 M arginine Precipitation score: 5 Activity: 0.011	#2. 0.3 M guanidine 0.4 M arginine Precipitation score: 4 Activity: 0.061	#3. 0.3 M guanidine 0.8 M arginine Precipitation score: 2 Activity: 0.035
#4. 0.6 M guanidine 0 M arginine Precipitation score: 4 Activity: 0.005	#5. 0.6 M guanidine 0.4 M arginine Precipitation score: 1 Activity: 0.012	#6. 0.6 M guanidine 0.8 M arginine Precipitation score: 1 Activity: 0.100
#7. 0.9 M guanidine 0 M arginine Precipitation score: 3 Activity: 0.011	#8. 0.9 M guanidine 0.4 M arginine Precipitation score: 1 Activity: 0.053	#9. 0.9 M guanidine 0.8 M arginine Precipitation score 1 Activity: 0.005
#6 after TBS buffer exchange Precipitation score: 1 Activity: 13.4	TBS dialysis Precipitation score: 1 Activity: 11.6	

Table III shows each tested refolding buffer, listed by number and the final concentration of guanidine and arginine. After addition of denatured MKP3 (1:20 dilution), precipitation was assessed as described in Methods (section 2.1.5) to give a precipitation score (1-5, with increasing precipitation). Activity was measured as described in Methods (section 2.1.16) by detecting the production of nitrophenolate (NP) from pNPP hydrolysis. Activity is given as $\mu M\ min^{-1}\ mg^{-1}$ MKP3.

The final row contains two final tests: #6 after TBS buffer exchange. This shows the activity of MKP3 refolded in buffer #6 once guanidine and arginine were removed. TBS dialysis. This shows the activity of MKP3 refolded by dialysis in the absence of guanidine and arginine.

Figure 11. MKP3 phosphatase activity assay. The activities of refolded MKP3:WT samples were determined using pNPP and DMSO. MKP3 samples were refolded or buffer exchanged as detailed in Methods. 1.25 μg of each putatively refolded MKP3 sample was assayed using the phosphatase activity assay detailed in Methods. MKP3 activity was detected by the production of nitrophenol (NP). Specific activity is shown as $\mu M\ min^{-1}\ mg^{-1}$ enzyme. Activity values can also be found in Table III.


15
10
5
0.150
0.125
0.100
0.075
0.050
0.025
0.000
µM NP/min/mg
#1
#2
#3
#4
#5
#6
#7
#8
#9
#6 - buffer exchange
TBS dialysis

The results of the precipitation and activity tests showed that dialysis-based refolding was the most suitable method for refolding MKP3. Dialysis resulted in low precipitation, a high specific activity MKP3, and did not require a cumbersome buffer exchange step to remove arginine and guanidine. It was therefore chosen for refolding MKP3 samples. The final refolding protocol can be found in Methods (section 2.1.6).

3.2.2. Assessment of MKP3 secondary structure by CD spectroscopy

The preceding section used the full-length MKP3 activity in DMSO to identify refolded proteins. However, the N- and C-terminal truncation mutants do not possess a previously characterized enzyme activity. Therefore, the proteins were refolded according to the protocol developed for MKP3:WT (Methods, section 2.1.6). To test if proteins were folded, protein secondary structures were assessed by circular dichroism (CD) spectroscopy. The data were then compared against known MKP3, N- and C-terminal domain structures (PDB ID's **1HZM** and **1MKP**; (*74;75*)) to ensure proper refolding. Since a structure for the full-length MKP3 is not available for a similar comparison, a theoretical secondary structural composition was determined. The value was obtained using a summed, linear-weighted average (sum of secondary structures divided by the lengths of amino acid sequence) from the known N- and C-terminal structures.

CD spectra were obtained as described in Methods (section 2.1.11) and are shown in Figure 12. The spectra were deconvoluted to obtain linear-weighted averages for the secondary structural content of each MKP3 protein. The calculated results, along with the expected and theoretical values, are tabulated in Table IV.

Figure 12. Circular dichroism spectra from MKP3. CD spectra were obtained from samples of the refolded MKP3s. Proteins were dissolved in TBS (pH 7.6) at concentrations between 0.5-1 mg/mL. CD spectra were obtained using a Jasco J-810 CD spectrophotometer as described in Methods. To enable inter-protein comparison, the raw data was converted to mean residue elipticity ($[\Omega]_{MRW}$, measured in degrees cm^2 $dmol^{-1}$ $residue^{-1}$) and plotted as a function of wavelength. The spectral data was analysed by the Dichroweb web analysis tool to calculate the secondary structural composition. The results are listed in Table IV.

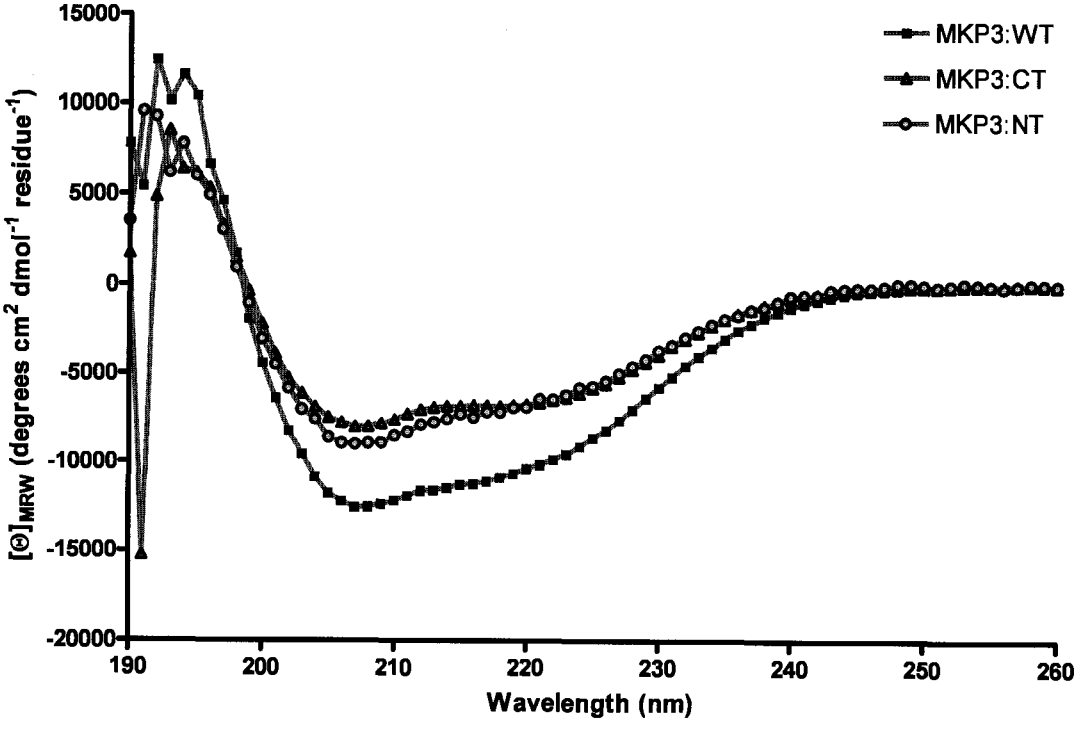

MKP3:WT
MKP3:CT
MKP3:NT
15000
10000
5000
0
-5000
-10000
-15000
-20000
190
200
210
220
230
240
250
260
Wavelength (nm)
[Θ]MRW (degrees cm2 dmol-1 residue-1)

Table IV. MKP3 secondary structure determined by CD spectroscopy						
Sample	Helix (%)		Sheet (%)		Other (%)	
	Observed	Predicted	Observed	Predicted	Observed	Predicted
MKP3:WT	36	33	14	12	50	55
MKP3:CT	43	38	24	20	33	42
MKP3:NT	40	29	19	5	41	66

The full-length wild type MKP3 protein showed CD data that indicated the protein possessed a mix of 36% α-helix, 14% β-sheet and 50% other (other consists of β- and γ-turn, and random coil). The values compared very favourable with the theoretical value, which predicted a secondary structure composition of 33% α-helix, 12% β-sheet and 55% other. The separate confirmations of enzyme activity, in the previous section (3.2.1b), and correct folding, shown by CD spectroscopy, suggest that the full-length protein was correctly refolded. This is the first example of a successful MKP3 recovery from bacterial inclusion bodies.

CD spectra were then obtained for the refolded MKP3:CT and MKP3:NT truncations. Analysis of the C-terminal, MKP3:CT showed that the protein contained 43% α-helix, 24% β-sheet, and 33% other. These values were compared to the linear-weighted average structures from the previously described inactive MKP3 C-terminal domain structure, which contained 38% α-helix, 20% β-strand, and 42% other. The CD results compared well to the predicted C-terminal values. While the refolded MKP3:CT showed more secondary structure, indicated by a greater percentage of α-helix and β-sheet and less turn or random coil, this may be explained by differences in the linear amino acid sequences used for the MKP3:CT truncation (a.a. 155-381) versus the sequence used to obtain the structure of **1MKP** (a.a. 204-347).

CD spectra of the N-terminal, MKP3:NT were obtained and showed that the protein contained 40% α-helix, 19% β-sheet and 41% other. This was in contrast to the NMR-determined N-terminal domain structure, which was reported as a 29% α-helix, 5% β-sheet, and 66% other. The CD results for MKP3:NT did not match the NMR determined secondary structures. However, these differences may have been due to the different buffer systems used between the experiments. CD spectroscopy was performed using a Tris buffered saline solution at physiological ionic strength, near physiological pH (ionic strength 157 mM, pH 7.6). NMR spectroscopy was performed at a similar ionic strength, but the buffer system was acidic as well as less hydrophilic due to the addition of urea (NMR buffer composition was 50 mM imidazole, 100 mM NaCl, 0.5 mM EDTA, 5 mM DTT, 200 mM urea, pH 6.0). In the NMR experiments, the acidic buffer was necessary for amide protonation in heteronuclear spectroscopy experiments and urea was likely used to improve protein solubility as it acts by interrupting protein-protein or protein-solvent hydrogen bonds. Both acidity and hydrophobicity can lead to loss of secondary structure. Therefore, while the N-terminal results did not show a close match to the expected secondary structure, CD spectroscopy did show that the MKP3:NT was folded into mixed alpha/beta units and not unstructured (MKP3:NT contained more secondary structural content than the NMR-determined **1HZM** structure). As a result, MKP3:NT, as well as the MKP3:WT and MKP3:CT proteins were all determined to be refolded and suitable for further analysis.

3.3. MKP3 interdomain binding

The refolded MKP3s were used in protein binding studies to characterize interdomain (N- to C-terminal) binding. This interaction was previously suggested by Farooq et al. (*75*) using the MKP3 N-terminal domain and the PAC1 catalytic domain (an MKP3, C-terminal domain analogue). In the current study, a direct analysis of N- to C-terminal binding was performed using both of MKP3's domains by a more sensitive surface plasmon resonance (SPR) method.

3.3.1. SPR-determined MKP3:WT affinity for ERK2 and JNK1

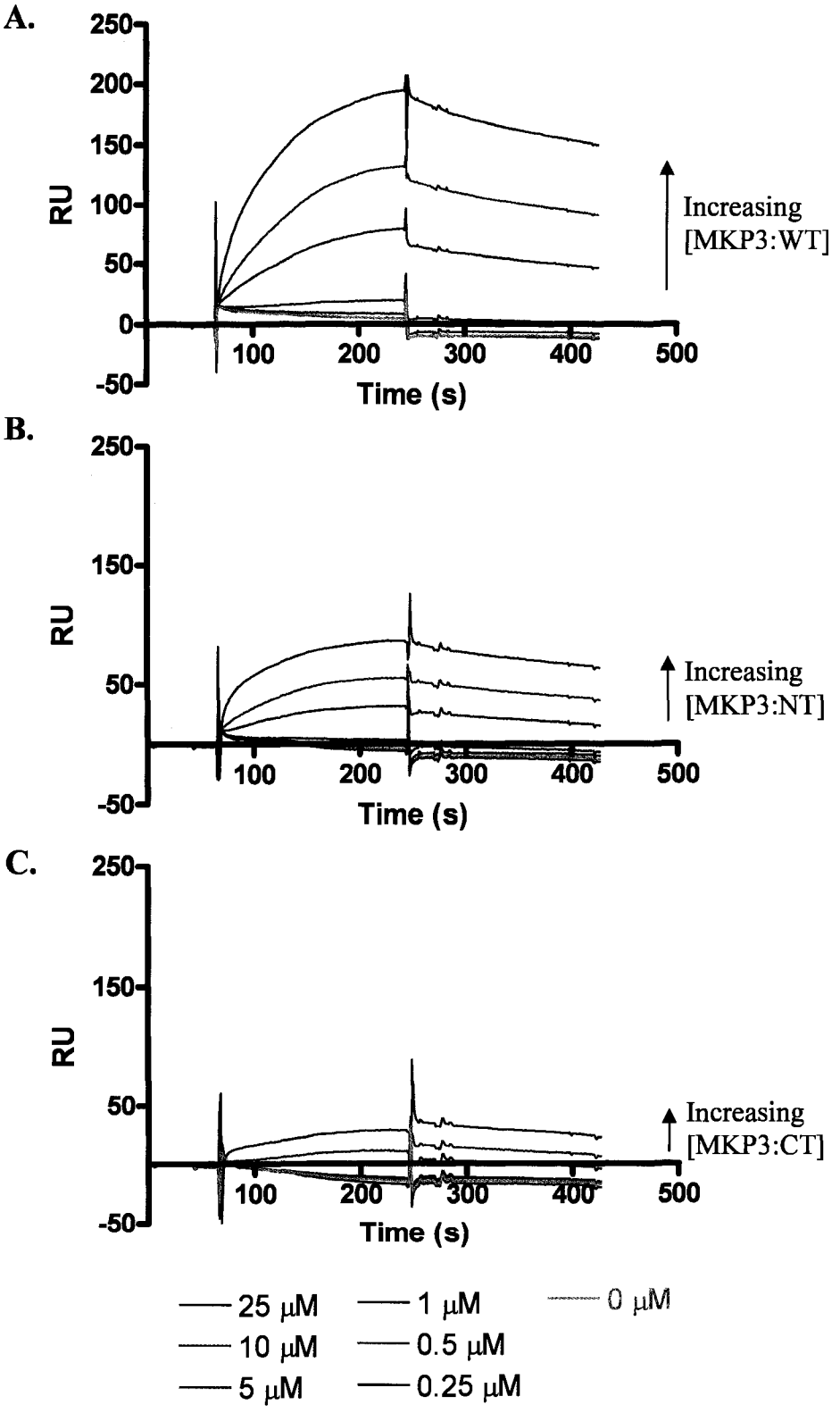
Control experiments were first performed to verify that the SPR method could detect MKP3 binding by using the full-length MKP3 and its known binding and non-binding MAP kinases (ERK2 and JNK1, respectively). Experiments were performed as described in Methods (section 2.1.12). The results, which are listed in Table V, indicated that MKP3 bound to ERK2 with an association rate (k_a) of 6.0 x 10^6 $M^{-1}s^{-1}$ and a dissociation constant (K_D) of 0.263 μM. The values were comparable to previously reported data that determined the MKP3/ERK2 K_D (using an enzymatic assay) at 0.17 +/- 0.04 μM (*41*). When MKP3:WT was tested using the JNK1 kinase (a non-binding MAPK), the data failed to show a binding response that could be fit to a 1:1 Langmuir association function. These data were also consistent with previous pull-down experiments that were unable to find an MKP3/JNK1 interaction using detection by anti-JNK1 antibody (*13;14*). The two results showed that the surface plasmon resonance was capable of determining the binding affinities in the positive and negative controls. The results also show that MKP3:WT was functional following dialysis-based refolding.

Table V. Surface plasmon resonance determined binding rates and constants				
Analyte	MKP3:WT		MKP3:CT	
	k_a	K_D	k_a	K_D
	$M^{-1}s^{-1}$	μM	$M^{-1}s^{-1}$	μM
ERK2	6.0×10^6 +/- 0.5×10^6	0.26 +/ 0.03		
JNK1	Did not bind			
MKP3:WT			456 +/- 11	0.26 +/- 0.02
MKP3:NT			574 +/- 23	1.4 +/- 0.4
MKP3:CT			23 +/- 4	48 +/- 18

3.3.2. SPR-determined MKP3:CT affinity for MKP3:WT, MKP3:NT and MKP3:CT

To characterize MKP3 interdomain binding, MKP3:WT, MKP3:NT and MKP3:CT proteins were used. Each protein was covalently immobilized on BIAcore CM5 sensor chips. MKP3:CT protein was then applied to test for binding against each immobilized MKP3 as described in Methods (section 2.1.12). The results in Figure 13A-C showed an increase in steady state binding, which was indicated by the increase in the plateau phase SPR response when the MKP3:CT analyte concentration was increased. Figures 13A and B show that MKP3:CT was able to bind both the full-length MKP3:WT and the N-terminal MKP3:NT proteins. The binding sensorgrams in Figure 13 were analysed by curve fitting to a 1:1 Langmuir model as detailed in Methods (section 2.1.12) to calculate the binding rates and constants (data are listed in Table V). The full-length MKP3 to C-terminal MKP3 interactions possessed association rate and dissociation constant of $k_a = 4.56 \times 10^2$ $M^{-1}s^{-1}$ and $K_D = 0.257$ μM, while the N-terminal MKP3 to C-terminal interaction possessed $k_a = 5.74 \times 10^2$ $M^{-1}s^{-1}$ and $K_D = 1.43$ μM. The association rates between the MKP3:WT and MKP3:NT to MKP3:CT were comparable, but the dissociation constant of MKP3:WT and MKP3:CT

Figure 13. Surface plasmon resonance binding measurement. Binding between MKP3 domains was measured using surface plasmon resonance. MKP3:CT was was covalently immobilized on BIAcore CM5 sensor chips and MKP3:WT, MKP3:NT, or MKP3:CT in the mobile phase was applied at increasing concentrations to test for binding. The concentrations used for analyses were 0, 0.25, 0.5, 1, 5, 10 and 25 μM. **A.** MKP3:WT, **B.** MKP3:NT, and **C.** MKP3:CT.


A.
250
200
150
100
50
0
-50
RU
100
200
300
400
500
Time (s)
Increasing
[MKP3:WT]
B.
250
150
50
-50
RU
100
200
300
400
500
Time (s)
Increasing
[MKP3:NT]
C.
250
150
50
-50
RU
100
200
300
400
500
Time (s)
Increasing
[MKP3:CT]
25 μM
10 μM
5 μM
1 μM
0.5 μM
0.25 μM
0 μM

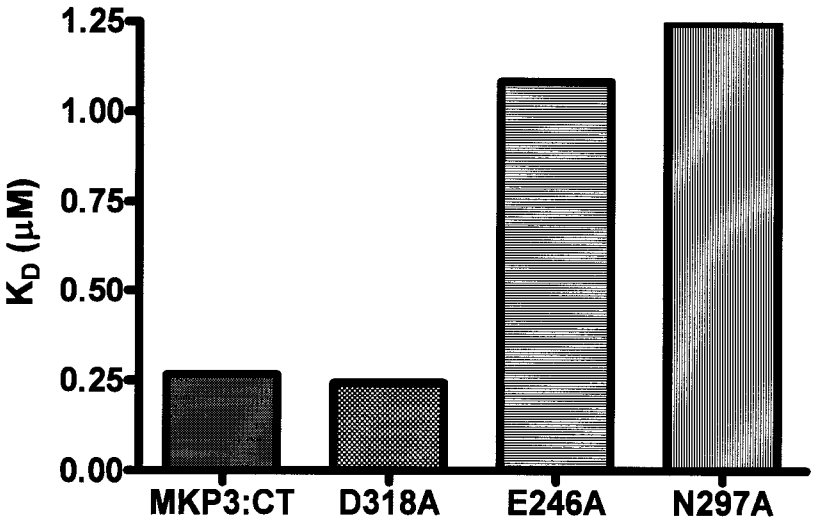
was significantly lower (an indication of a better affinity) than that of the MKP3:NT interaction. This result indicated that the full-length protein possesses additional binding motifs that are not found in the N-terminal (a.a. 1-154) sequence.

The C-terminal, MKP3:CT to MKP3:CT binding was also assessed. Sensorgram data (Figure 13C) showed that MKP3:CT protein was able to bind to MKP3:CT. However, the analysed data indicated that the MKP3:CT/MKP3:CT association was a significantly weaker interaction than MKP3:CT binding to either MKP3:WT or MKP3:NT. The analysed data gave a $k_a = 22.7\ M^{-1}s^{-1}$ and a $K_D = 47.7\ \mu M$. The MKP3:CT self-association indicates that the MKP3:WT likely possesses a protein binding site through which two C-terminal domains can bind. The detection of multiple binding sites (N- to C- and C- to C-termini) can help explain the different affinities of the C-terminal domain to MKP3:WT and MKP3:NT (which would lacks the C- to C-terminal binding ability). However, the low K_D value for MKP3:CT to MKP3:WT and specifically to MKP3:NT shows that interdomain binding would overwhelmingly favour an N-terminal to C-terminal interaction.

3.3.3 Characterization of interdomain binding

To identify the N- to C-terminal domain binding site, several peptide sequences were selected and tested against both MKP3:NT and MKP3:CT using SPR. The peptide sequences selected included 1. ERK2's MKP3 recognition motif, 2. ERK2's biphosphorylated activation loop, 3. MKP3's kinase interacting motif, and 4. MKP3's hydrophobic post-KIM region (the specific sequences can be found in Methods, section 2.1.13). The peptides were tested for binding, but all results proved negative (no indication of binding, data not shown) except for the post-KIM/MKP3:CT interaction, which is shown in Figure 14.

Figure 14. Surface plasmon resonance binding measurement using post-KIM peptide and MKP3:CT mutants. The dissociation constants (K_D) between the post-KIM peptide and MKP3:CT proteins were determined using surface plasmon resonance. Post-KIM peptide was immobilized on BIAcore CM5 sensor chips and samples of MKP3:CT, MKP3:CT/D318A, MKP3:CT/E248A or MKP3:CT/N267A was added at increasing concentrations as described in Methods. The concentrations of MKP3:CT were 0, 0.25, 0.5, 1, 5, 10 and 25 μM. Each experiment was performed once. Larger values indicate less affinity.


1.25
1.00
0.75
0.50
0.25
0.00
K_D (μM)
MKP3:CT
D318A
E246A
N297A

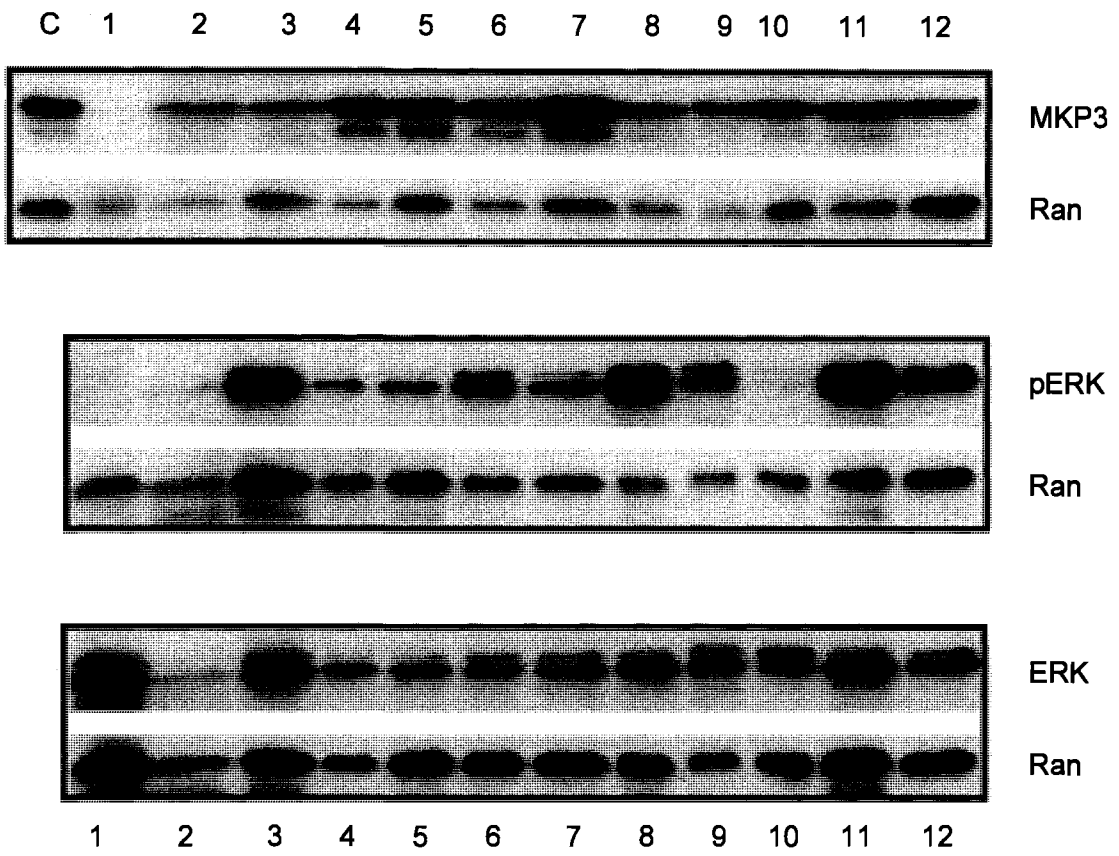
To identify the corresponding region on the MKP3:CT protein that the post-KIM peptide interacted with, mutation analysis was performed on MKP3:CT. The MKP3:CT mutants contained single amino acid substitutions at potential binding site motifs that were identified using existing magnetization transfer and hydrogen/deuterium exchange data (*67;75;77*). The three MKP3:CT mutants that were generated were E248A, N267A, and D318A (detailed in Methods, section 2.1.1). The results of SPR analysis in Figure 15 show that, compared to the MKP3:CT, the E248A and N267A mutants lost the ability to bind the post-KIM peptide. However, while E248A did show a loss of binding, the variability between injections of E248A was extremely high. This was a sign of potential sensor chip deterioration and the result must be taken cautiously. The data did show that the post-KIM peptide could bind the C-terminal catalytic domain and that mutation of N267 from a charged, hydrophilic residue to alanine, an uncharged, hydrophobic residue, resulted in a loss of the binding.

3.3.4. Immunodetection of MKP3 in pancreatic adenocarcinoma cells

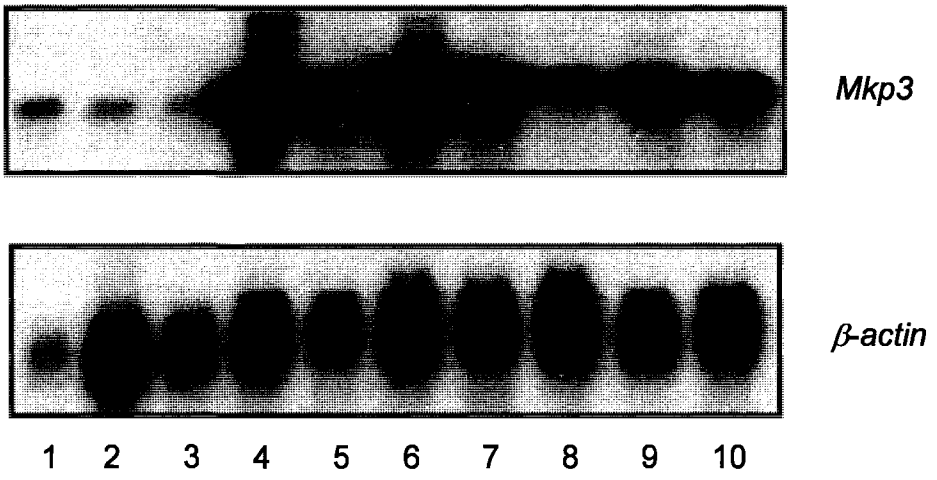
The SPR data indicated that MKP3's N- and C-terminal domains could bind (Figure 13), these data suggested that the protein could form dimers or oligomers through an interaction similar to "domain swapping" (*91-93*) or intercalation of MKP3's N- and C-termini between MKP3 subunits. To test this hypothesis, MKP3-overproducing pancreatic adenocarcinoma cells lines were identified. A panel of pancreatic adenocarcinoma cell lines were probed using an anti-MKP3 antibody, as described in Methods (section 2.1.8). The results of Western detection are presented in Figure 15A. HeLa cells (lane C), serving as a positive control, showed high levels of MKP3 as well as a lower molecular weight band that may represent the previously characterized MKP3 splicing variant that lacks exon 2 (*94*). The normal human pancreatic cells (homogenized whole pancreas), in lane 1,

Figure 15. Analysis of MKP3 and ERK1/2 expression in human pancreatic andenocarcinoma cell lines. A. MKP3 and ERK1/2 Western blot analysis. Proteins were detected using 20 μg total cellular protein from each of the indicated cell lines, as described in Methods. RAN (RAS-associated nuclear protein) was used as the protein loading control. Top panel shows MKP3, middle and lower panels show phosphorylated and total ERK levels, respectively. Lane C: HeLa cells (positive control); Lane 1: Normal pancreas (homogenized human pancreatic tissue used to detect the baseline MKP3 level); Lane 2: HPDE4 (immortalized, non-tumourigenic ductal pancreatic cells); Lane 3: HPDE6 (immortalized, non-tumourigenic ductal pancreatic cells); Lane 4-11 utilized pancreatic adenocarcinoma cell lines. Lane 4: CRL1420 (MIAPaCa-2); Lane 5: CRL1469; Lane 6: CRL1682 (BxPC-3); Lane 7: CRL1687; Lane 8: CRL1837 (Su86.86); Lane 9: HTB79; Lane 10: HTB80; Lane 11: HTB134; Lane 12: HTB147. **B.** *Mkp3* Northern blot analysis. Ten microgram of total cellular RNA were fractionated and processed for Northern blot hybridization as described in Methods. *Mkp3* was detected using a 1.25 kbp human *Mkp3* HindIII/XhoI cDNA fragment, while a human *β-actin* cDNA was used as the loading control. Lane 1. Normal pancreas, Lane 2. CRL1420 (MIAPaCa-2), Lane 3. CRL 1469, Lane 4. 1682, Lane 5. 1687 (BxPC-3), Lane 6. 1837 (Su86.86), Lane 7. HTB79, Lane 8. HTB80, Lane 9. HTB134, Lane 10. HTB147.

A.


C 1 2 3 4 5 6 7 8 9 10 11 12
MKP3
Ran
pERK
Ran
ERK
Ran
1 2 3 4 5 6 7 8 9 10 11 12

B.


Mkp3
β-actin
1 2 3 4 5 6 7 8 9 10

were found to possess low basal levels of MKP3. HPDE4 and HPDE6 cell lines (lanes 2 and 3), which are ductal pancreatic cell lines that have been immortalized (but not tumourigenic) were also included, as they are comparable to precancerous pancreatic neoplasias. These cells were observed to possess high levels of MKP3. The final lanes in the Figure 15A (lanes 4-12) contained a series of pancreatic adenocarcinoma cells lines, all of which showed the presence of MKP3 expressed at levels equal to or exceeding both normal pancreatic tissues and the immortalized ductal cells. These results were in contrast to previously published data (*22*) that showed no detectible MKP3 in pancreatic adenocarcinomas. Additional experiments were performed in parallel to detect phosphorylated ERK (pERK) and total ERK. The results indicated that pERK was overproduced in pancreatic adenocarcinoma cells compared to the ductile cell line (Figure 15A, lane 1), This result showed that MKP3 produced in the adenocarcinomas was functionally impaired or non-functional. *Mkp3* RNA levels were detected using a Northern blot, which was performed as detailed in Methods (section 2.1.9). The results of Northern analysis are shown in Figure 15B. The Northern blot indicated that human pancreas, in lane 1, possesses a low level of *Mkp3* RNA, but in seven of the nine adenocarcinoma lines (lanes 4-10) the RNA levels were highly upregulated and indicated that *Mkp3* was actively transcribed. The result confirmed the results of the MKP3 Western blot. The two adenocarcinoma cell lines (CRL1420 and CRL1469; lanes 2 and 3, respectively) that possessed lower *Mkp3* RNA levels showed high MKP3 protein levels. The disagreement between results may indicate that MKP3 was not degraded or that the protein displayed a long half-life within these cell lines. Regardless, the RNA data confirmed the Western results in Figure 14A, which indicated MKP3 was expressed in pancreatic adenocarcinomas at high levels.

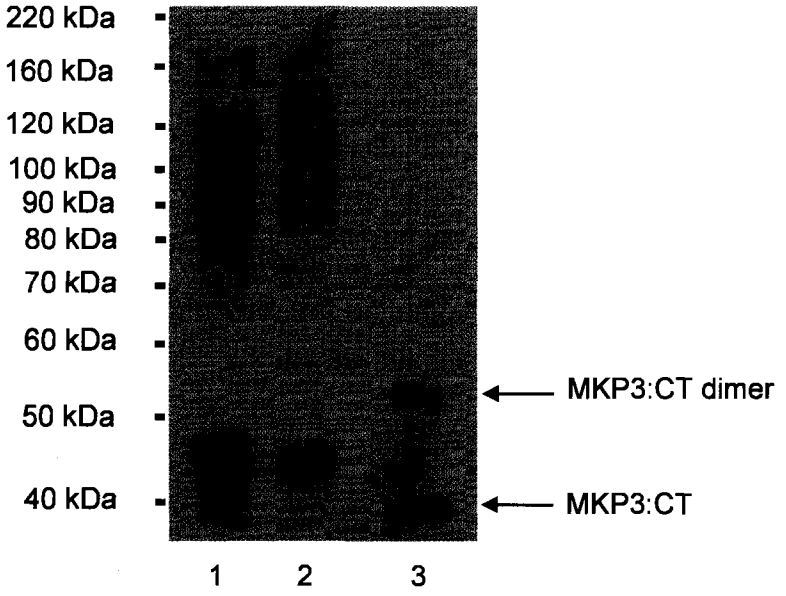
To determine if the high MKP3 concentration resulted in oligomerization, a second Western blot was performed using CRL1687 cells (chosen for their overproduction of MKP3). CRL1687 cells were lysed in non-denaturing buffers to retain protein-protein complexes. Samples of purified, recombinant MKP3:WT and MKP3:CT were prepared similarly. The results show, as predicted by surface plasmon resonance data (Figure 13C), MKP3:CT (~30 kDa) could bind to itself and was detected as a homo-dimer at ~60 kDa (Figure 16A, lane 3). The purified, recombinant MKP3:WT protein (~45 kDa; Figure 16A, lane 2) similarly showed homo-dimer formation, but also homo-trimer formation. The series of MKP3:WT monomer, dimer, and trimer was compared to bands detected in CRL1687 cell lysates (Figure 16A, lane 1). The immunodetected bands from both MKP3:WT and CRL1687 lanes were identical and indicated that MKP3 formed oligomers in CRL1687 cells. Lastly, to compare the result across MKP3-overproducing pancreatic adenocarcinomas, the Western blot was repeated with additional cell lines (Figure 16B). Each of the lines showed immunodetected bands that were identical to the CRL1687 lane (Figure 16B, lane 5), which were attributed to MKP3 oligomerization. The result suggested that oligomerization of MKP3 is a common occurrence in pancreatic adenocarcinomas.

3.4. Development of an MKP3 phosphatase activity assay

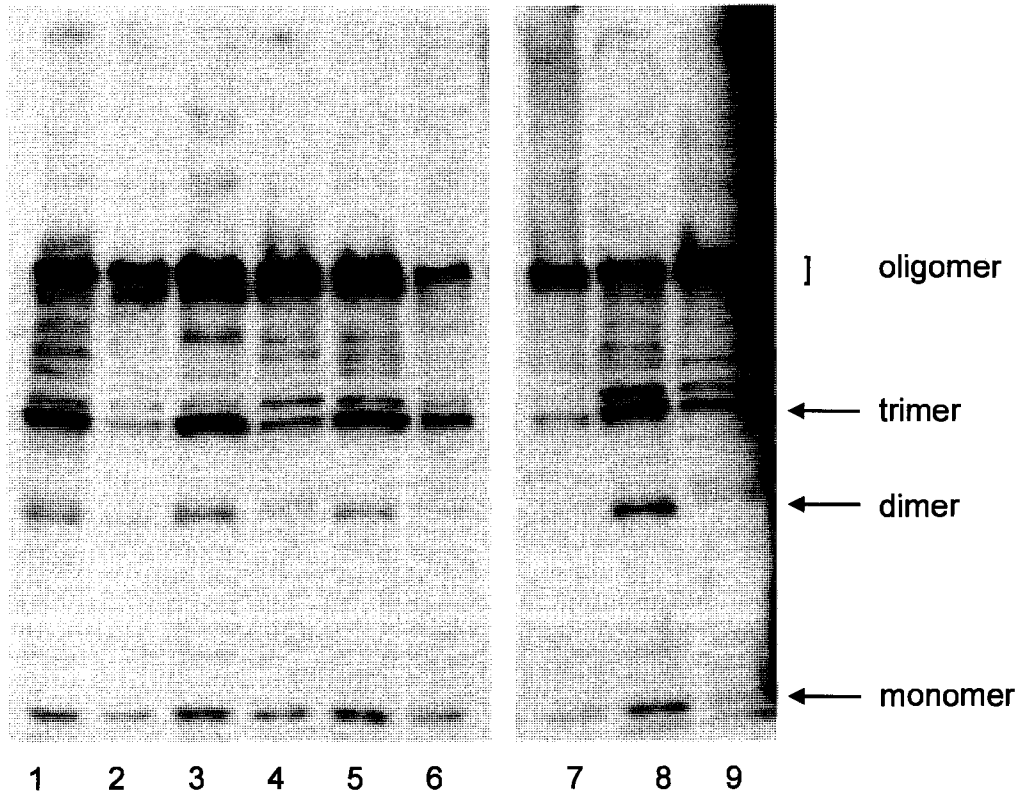
To determine if interdomain binding could affect MKP3's phosphatase activity, an assay was needed that could test the MKP3:CT protein in its active state, either in the presence or absence of MKP3:NT. For the full-length MKP3, activity is typically measured using a *p*-NPP assay in the presence of ERK. Addition of ERK is necessary to bind and allosterically activate MKP3, which then dephosphorylates pNPP. The same method cannot be used for the truncated MKP3:CT because it lacks the ability to bind and become

Figure 16. Western analysis of MKP3 under non-denaturing condition. Western analysis was performed to test for MKP3 oligomerization. Proteins were prepared in non-denaturing RIPA buffer to retain any bound complexes. Cell lysate (10 μg) or purified, recombinant MKP3 (0.4 μg) were separated by 7.5% PAGE prior to Western detection. **A.** Western-detected MKP3 proteins. Lane 1: CRL1687, Lane 2: purified MKP3:WT, Lane 3: purified MKP3:CT. **B.** Western-detected MKP3 proteins in pancreatic adenocarcinoma cell lysates. Lane 1: HPDE6; Lane 2: CRL1420; Lane 3: CRL1469; Lane 4: CRL1682; Lane 5: CRL1687; Lane 6: CRL1837; Lane 7: HTB80; Lane 8: HTB134, Lane 9: HTB147.

A.


220 kDa
160 kDa
120 kDa
100 kDa
90 kDa
80 kDa
70 kDa
60 kDa
50 kDa
40 kDa
MKP3:CT dimer
MKP3:CT
1
2
3

B.


]
oligomer
trimer
dimer
monomer
1
2
3
4
5
6
7
8
9

activated by ERK. Since DMSO was shown to increase the activity of MKP3:WT in ERK-free assay conditions (in section 3.2.1b) it was explored as a candidate activator for MKP3:CT in an ERK-free pNPP assay.

3.4.1. Effect of solvents on MKP3 activity

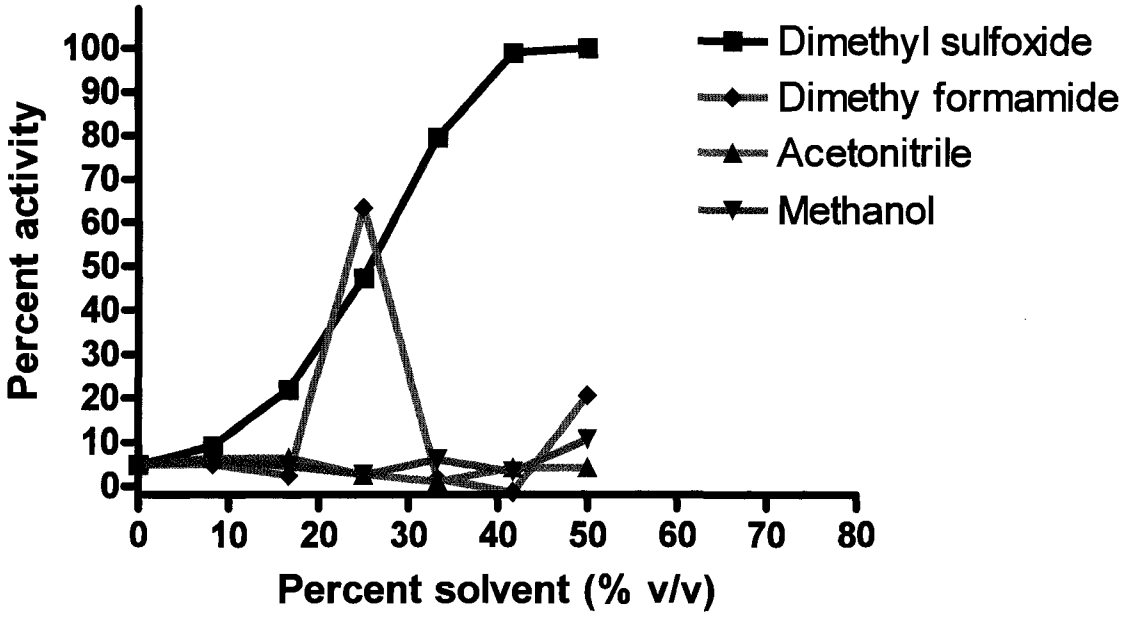
DMSO could increase the activity of the full-length MKP3 protein. To test if the property was specific to DMSO, solvents that possessed a structural or chemical similarity (Table VI) were selected for testing as co-solvents in a pNPP assay. Each solvent was tested using the assay conditions detailed in Methods (section 2.1.16) at 0-50% (v/v) concentrations.

Table VI. Properties of solvents

Solvent	Formula	Boiling point (°C)	Dipole moment (μ)	Dielectric constant (κ)
Water	H_2O	100	1.82	80
Dimethyl sulfoxide	$(CH_3)_2SO$	189	3.96	47.2
Dimethyl formamide	$(CH_3)_2NCHO$	153	3.82	38.3
Acetonitrile	CH_3CN	81	3.92	36.6
Methanol	CH_3OH	68	1.7	33
Acetone	$(CH_3)_2CO$	56	2.88	20.7

During the testing, precipitation was observed at the higher solvent concentrations, particularly in acetonitrile and methanol. When precipitation was noted, the activity was adjusted to the concentration of MKP3 that remained in solution (measured by UV absorbance, as detailed in section 2.1.14). When acetone was tested at concentrations greater than 25% (v/v), the solvent began to dissolve the microtitre plate, making measurement of activity impossible (thus data not shown). To compare the effects of the remaining solvents on MKP3 activity, the hydrolysis of pNPP with each solvent concentration was plotted. The

Figure 17. Effect of co-solvents on MKP3 activity. The effect of co-solvents on MKP3's phosphatase activity was measured by pNPP hydrolysis in the presence of DMSO, DMF, acetonitrile, methanol, and acetone as described in Methods. The activities of MKP3 in the indicated concentration of co-solvents are shown as a percentage of the maximum DMSO-induced MKP3 activity and are adjusted for protein precipitation.


Percent activity
100
90
80
70
60
50
40
30
20
10
0
0
10
20
30
40
50
60
70
80
Percent solvent (% v/v)
Dimethyl sulfoxide
Dimethy formamide
Acetonitrile
Methanol

activity and concentration data are shown in Figure 17 and plots activity as a percentage of the 50% DMSO concentration (50% DMSO resulted in the greatest increase in pNPP hydrolysis and was set as 100% activity). The results showed that, aside from DMSO, only DMF had an effect on MKP3 activity. However, DMF increased activity at only one tested concentration and not a range of concentration as shown by DMSO. The results of the methanol, acetonitrile and acetone data could not be used to draw a relationship between a general dipole or dielectric constant effect and DMSO's ability to increase MKP3 activity. And since DMF, methanol, acetonitrile, and acetone showed little or no effect on MKP3 activity, their use was not pursued further. DMSO, which increased MKP3 activity, was selected for ERK-free assay development.

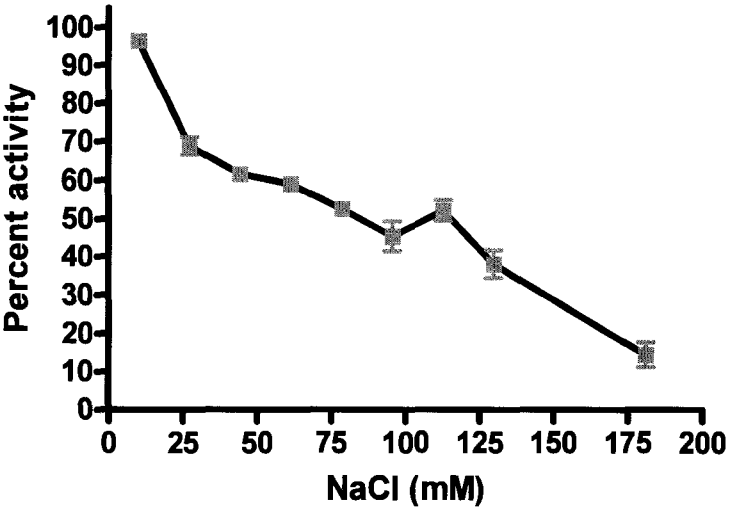
3.4.2. Effect of ionic strength on MKP3 activity

The effect of ionic strength on MKP3's phosphatase activity was tested. MKP3 was assayed using pNPP, with and without DMSO (as described in section 2.1.16) using NaCl to adjust the ionic strength of the assay buffer. Figure 18A shows that an increase in ionic strength led to a significant decrease MKP3 activity when assayed in the presence of DMSO. Figure 18B shows the opposite, an increase in ionic strength above 100 mM resulted in a small, but significant increase in MKP3 activity when assayed in the absence of DMSO.

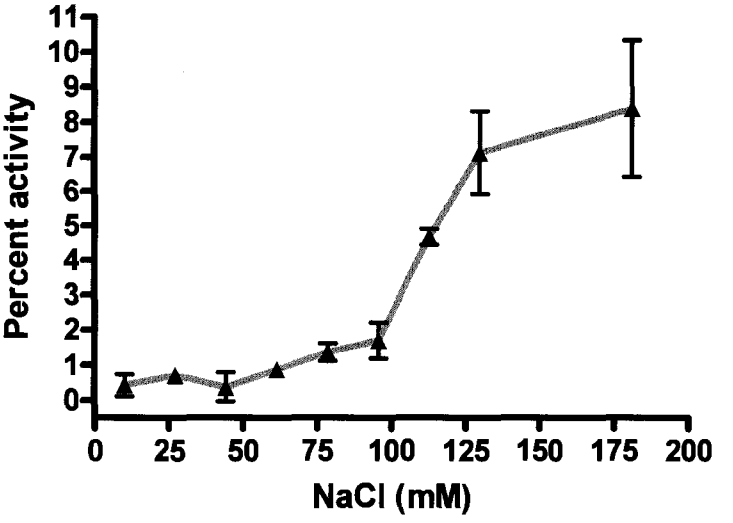
The different effects on activity depended on the presence or absence of DMSO. The most pronounced effect was in the presence of DMSO, where higher ionic strength from the addition of NaCl concentration led to a diminished MKP3 activity. This effect was similar to what was observed in refolding assays where high concentrations of charged salts (guanidine or arginine) resulted in decreased MKP3 activity (section 3.2.1b). Since both Tris and pNPP, which would be used for any assay, contributed to the ionic strength of buffers, these data were used to select appropriate pNPP concentration ranges for the kinetic analyses

Figure 18. Effect of ionic strength on MKP3 activity. The effect of ionic strength on MKP3 phosphatase activity was tested by increasing the NaCl concentration in pNPP assay buffer. Experiments were performed in a 150 µl volume as described in Methods (section 2.1.16). Data is shown as a percentage MKP3 activity at 10 mM ionic strength (100% activity is set to DMSO-activated MKP3, 0% is basal MKP3). Buffer ionic strength for each assay is indicated on the x-axis. **A.** Effect of NaCl concentration on MKP3 in the presence of DMSO, **B.** Effect of NaCl concentration on MKP3 in the absence of DMSO (y-axis showing the 0-10% activity range), **C.** MKP3 in the presence and absence of DMSO (y-axis showing 0-100% activity range).

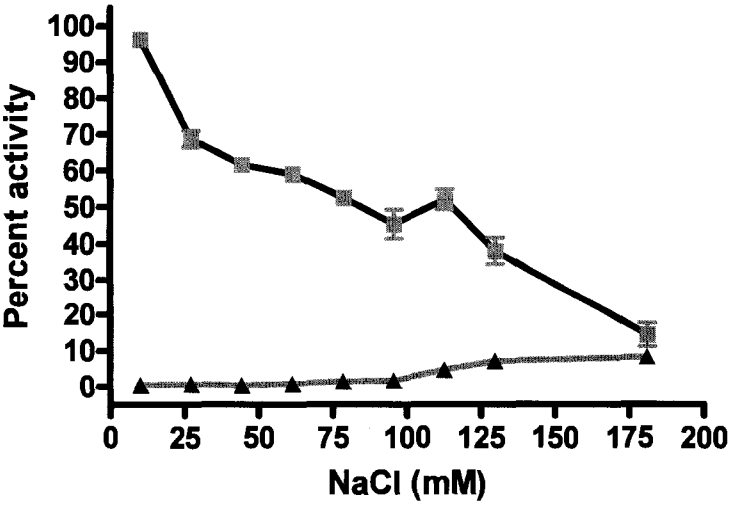
A.


Percent activity
100
90
80
70
60
50
40
30
20
10
0
0
25
50
75
100
125
150
175
200
NaCl (mM)

B.


Percent activity
11
10
9
8
7
6
5
4
3
2
1
0
0
25
50
75
100
125
150
175
200
NaCl (mM)

C.


Percent activity
100
90
80
70
60
50
40
30
20
10
0
0
25
50
75
100
125
150
175
200
NaCl (mM)

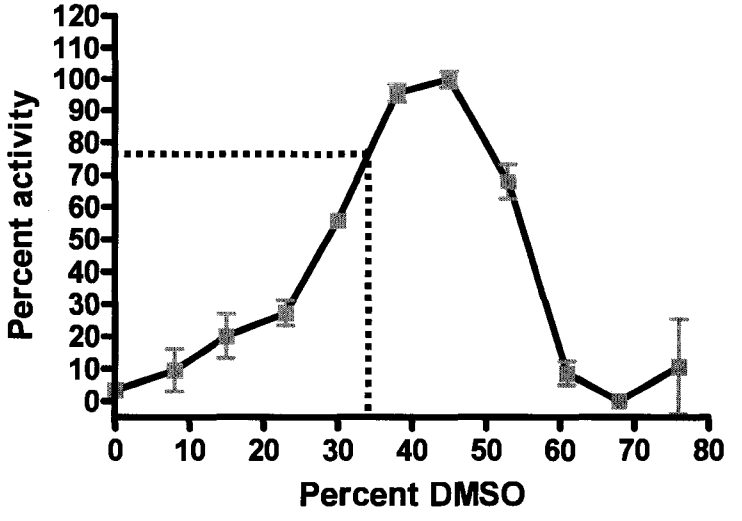
performed in section 3.4.4. The first range was 0-10 mM pNPP, with buffer ionic strength adjusted to 25 mM. The range was used to test MKP3 in the presence of DMSO where the low ionic strength could maximize MKP3 activity. The second range was 0-50 mM pNPP, with buffer ionic strength adjusted to 160 mM. This range was used to test MKP3 in the absence of DMSO. The higher pNPP concentrations were required to improve the kinetic measurements for basal (unactivated) MKP3 activity. While different ionic strengths were chosen for testing MKP3 in the presence or absence of DMSO, this was deemed an acceptable compromise. The 160 mM ionic strength used in tests without DMSO may have led to a slightly elevated activity, but as shown in Figure 18C, this was a small increase in relation to the MKP3 activity in the presence of DMSO, which was tested at a 25 mM ionic strength.

3.4.3. Effect of DMSO concentration on MKP3 activity

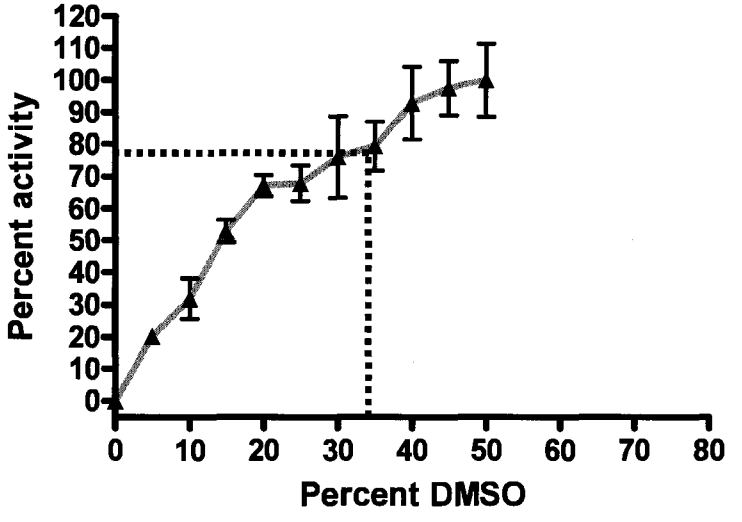
The preceding experiments were performed to optimize solvent and ionic strength conditions for an MKP3 activity assay, but all tests were performed on full-length MKP3. To test if the conditions also affected the truncated MKP3:CT, both MKP3:WT and MKP3:CT were tested using increasing DMSO concentrations at 25 mM ionic strength, as detailed in Methods (section 2.1.16). The results from MKP3:WT assays are shown in Figure 19A and confirmed that DMSO led to increased MKP3:WT activity. This result was similar to the previous DMSO data shown in Figure 17 (using 0-50% DMSO concentrations). However, this experiment used an extended range of DMSO concentrations (0-76%) and was an average of three experiments. The curve displayed an optimum and showed a sharp decline in activity above 45% DMSO. This loss of activity may have been a result of protein denaturation or a lack of acid/base ionization due to the loss of water at the higher solvent concentrations.

Figure 19. Effect of DMSO concentration on MKP3:WT and MKP3:CT activities. MKP3:WT and MKP3:CT were assayed at increasing DMSO concentrations. Assays were performed in triplicate as described in Methods (section 2.1.16) using either 1.2 μg MKP3:WT or MKP3:CT. DMSO concentrations were increased from 0-76% in MKP3:WT experiments and 0-50% in MKP3:CT experiments. Data is reported as a percentage of the maximum activity given by each protein. **A.** MKP3:WT **B.** MKP3:CT.

A.


Percent activity
120
110
100
90
80
70
60
50
40
30
20
10
0
0 10 20 30 40 50 60 70 80
Percent DMSO

B.


Percent activity
120
110
100
90
80
70
60
50
40
30
20
10
0
0 10 20 30 40 50 60 70 80
Percent DMSO

MKP3:CT was tested using a 0-50% DMSO concentration range. Figure 19B shows that MKP3:CT activity could also be increased using DMSO. This represented the first reported activation of MKP3's truncated catalytic domain and showed that DMSO might be able to simulate the allosteric activation of MKP3 by ERK, even though MKP3:CT lacked the N-terminal ERK-binding domain.

Using the data in Figure 19, a 35-45% DMSO concentration was deemed optimal for both MKP3:WT and MKP3:CT assays, but at a 33% concentration, the relative increase in activity was comparable (~75-79%) for both MKP3:WT and MKP3:CT (Figures 19A and B). Since a 33% DMSO concentration had been used in previous assays, both in our lab (*88*) and in previous literature (*15*), and the relative increase in activity with the concentration was equivalent, 33% DMSO was used in later experiments and this ensured continuity with the previous data.

3.4.4. Comparison of ERK- and DMSO-induced MKP3 activities

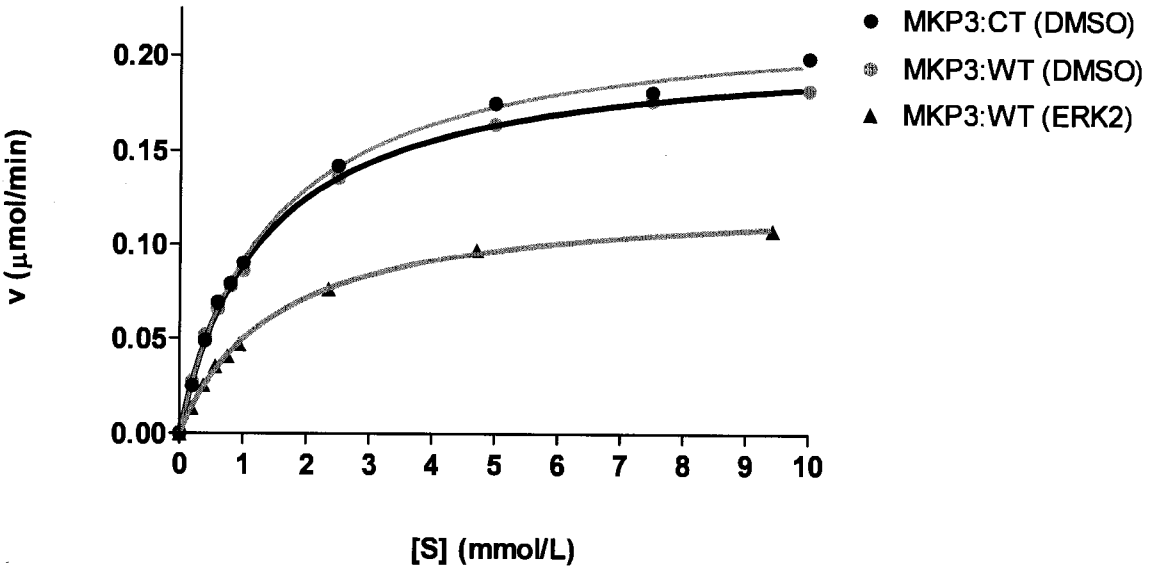
Data from the experiments in sections 3.4.1-3.4.3 were used to develop an ERK-free assay to measure MKP3 activity. This assay was compared to the accepted pNPP assay using ERK as an activator. The comparison was used to determine if the increased activity, resulting from DMSO addition to assay buffers, was similar to ERK's allosteric activation of MKP3. Experiments were performed using steady-state kinetic analysis to characterize MKP3's k_{cat} and K_M under basal, ERK-activated, and DMSO-activated conditions. Experiments were performed using both MKP3:WT and MKP3:CT.

3.4.4.a. ERK2-induced MKP3 activity

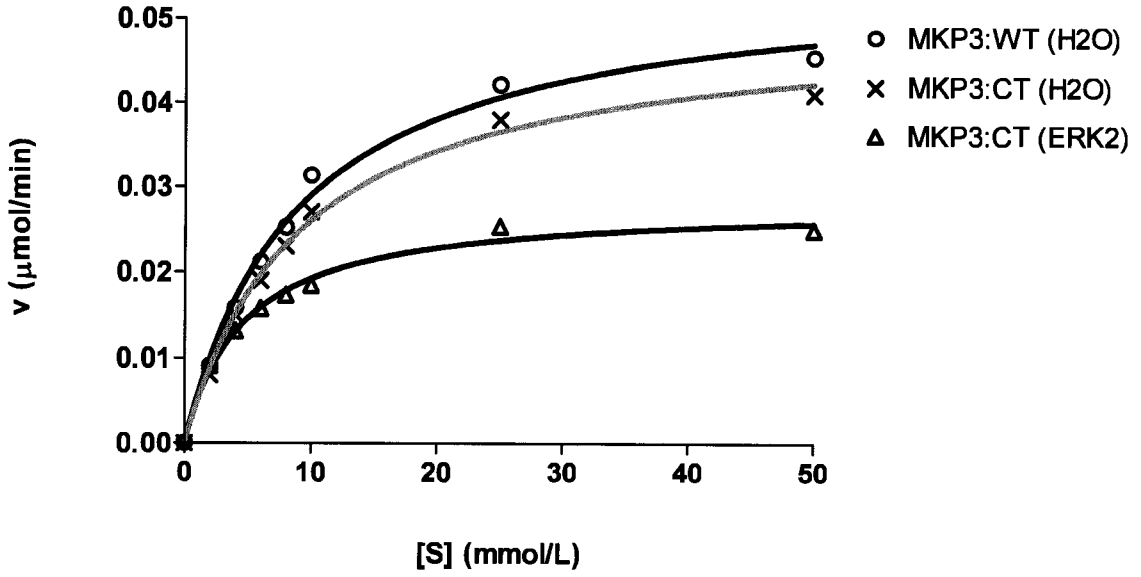
Kinetic rates and constants were determined (as detailed in section 2.1.17) using pNPP assay buffers with and without ERK. Representative steady state data from all tests and curve fittings is shown in Figure 20A (activated MKP3 assays) and B (unactivated

Figure 20. Steady-state analysis of MKP3 activity. Kinetic analysis was performed on MKP3:WT and MKP3:CT using ERK2, 33% DMSO, or H2O as a control. Experiment were performed as described in Methods (section 2.1.17), in triplicate and using 1.2 μg (high activity MKP3 test) or 12 μg MKP (low activity MKP3 test). **A.** Data from ERK and DMSO-activated MKP3 samples. Data was obtained using a 0-10 mM substrate concentration. **B.** Data from unactivated MKP3 samples. Data was obtained using a 0-50 mM substrate concentration. All data was fit to the Michealis-Menton equation using a non-linear curve fit in Graphpad Prizm.

A.


MKP3:CT (DMSO)
MKP3:WT (DMSO)
MKP3:WT (ERK2)
v (μmol/min)
0.20
0.15
0.10
0.05
0.00
0
1
2
3
4
5
6
7
8
9
10
[S] (mmol/L)

B.


MKP3:WT (H2O)
MKP3:CT (H2O)
MKP3:CT (ERK2)
v (μmol/min)
0.05
0.04
0.03
0.02
0.01
0.00
0
10
20
30
40
50
[S] (mmol/L)

MKP3 assays). The analysed data is tabulated in Table VII. Results from the steady-state analysis show that addition ERK resulted in an increased MKP3:WT hydrolysis of pNPP compared to the MKP3:WT control. The data indicated that allosteric activation resulted in an improved catalytic efficiency that was composed of two parts: an increase in k_{cat} and a decrease in K_M. Compared to the pre-activated values (MKP3:WT + Water), k_{cat} showed a three-fold increase, while K_M showed a five-fold decrease.

Table VII. Kinetic analysis of activated or unactived MKP3

	k_{cat} (s^{-1} x 10^{-3})	K_M (M x 10^{-3})	k_{cat}/K_M ($M^{-1}s^{-1}$)
MKP3:WT + Water	9.9 +/- 0.41	7.87 +/- 0.99	1.25 +/- 0.21
MKP3:WT + ERK2	29.6 +/- 6.1	1.61 +/- 0.14	18.4 +/- 2.0
MKP3:WT + 33% DMSO	39.6 +/- 3.7	1.32 +/- 0.04	30.1 +/- 3.6
MKP3:CT + Water	12.0 +/- 0.4	7.60 +/- 0.82	1.59 +/- 0.23
MKP3:CT + ERK2	14.0 +/- 0.9	4.50 +/- 0.68	3.11 +/- 0.66
MKP3:CT + 33% DMSO	58.1 +/- 0.4	1.44 +/- 0.07	40.3 +/- 2.1

The kinetic assay was repeated using MKP3:CT with and without ERK2. While the catalytic domain showed approximately a two-fold increase in k_{cat}/K_M, the increase was solely due to a decreased K_M value. While the result indicated that MKP3:CT could interact with ERK2 (leading improved substrate affinity), the loss of the N-terminal domain prevented full allosteric activation of the truncated mutant. A similar result was reported by Stewart *et al.* (*74*), using a shorter MKP3 truncation (a.a. 204-347) that showed a two-fold decrease in K_M, without an increase in k_{cat}, when the protein was assayed in the presence of ERK2.

The results of the experiment show that both MKP3:WT and MKP3:CT were able to bind or associate with ERK2. This interaction led to respective decreases in the enzyme's K_M (the apparent first-order substrate binding constant) for pNPP. However, only MKP3:WT was capable of undergoing complete allosteric activation, where the binding of ERK2 also led to an increased k_{cat}.

3.4.4.b. DMSO-induced MKP3 activity

A parallel experiment was performed using 33% DMSO in place of ERK2 and compared to the previous results. The calculated k_{cat} and K_M data are shown in Table VII. The addition of DMSO led to an increase in MKP3:WT activity. The resulting DMSO-induced catalytic efficiency (k_{cat}/K_M) was determined to be 30.1 +/- 3.6 $M^{-1}s^{-1}$, which was 24-fold greater than unactivated MKP3:WT and even 1.6-fold greater than MKP3:WT activated by ERK2. The higher k_{cat}/K_M values observed in the assays with DMSO were due to both an increase in k_{cat} and a decrease in K_M, which was similar to assays in which ERK was added. This showed that DMSO could chemically mimic ERK's allosteric activation of MKP3.

MKP3:CT activity was then tested with DMSO. The calculated results showed that, both the k_{cat} and K_M were altered by DMSO. First, K_M was decreased to 1.44 +/- 0.065 μM, a value comparable to MKP3:WT activated by ERK or by DMSO. Secondly, the DMSO-activated k_{cat} was also increased. However, the increase in k_{cat} was significantly larger than either the increase detected using ERK2 and MKP3:WT (~2-fold greater) or when DMSO and MKP3:WT were tested (~1.5-fold greater). The result showed that DMSO was capable of mimicking ERK's allosteric effect by increasing k_{cat} and decreasing K_M. The results also showed that under the same pNPP assays condition, when DMSO was added, MKP3:CT

possesses a greater catalytic efficiency than the full-length MKP3:WT. This suggested an inhibitory function for the N-terminal domain, which is removed in the MKP3:CT protein. The inhibition was a novel finding that could not be detected by the standard pNPP assay utilizing ERK. This was because the MKP3 catalytic domain cannot be fully activated by ERK, as it is with DMSO.

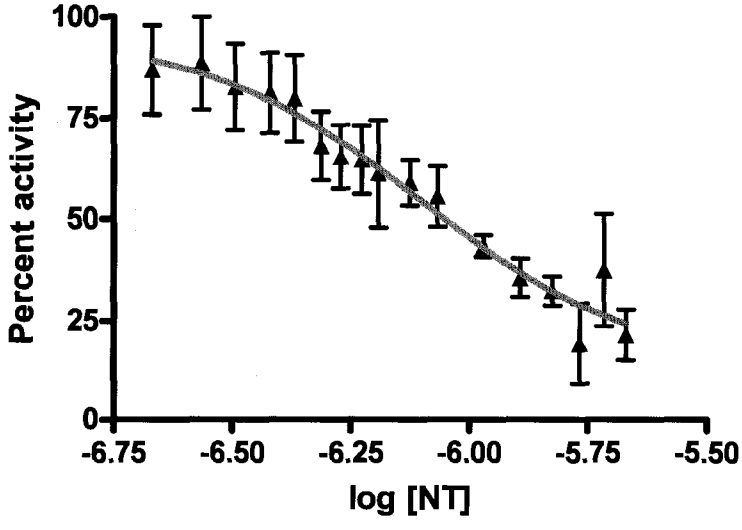
3.5. Auto-inhibition of MKP3

3.5.1. Effect of MKP3:NT concentration on MKP3 activity

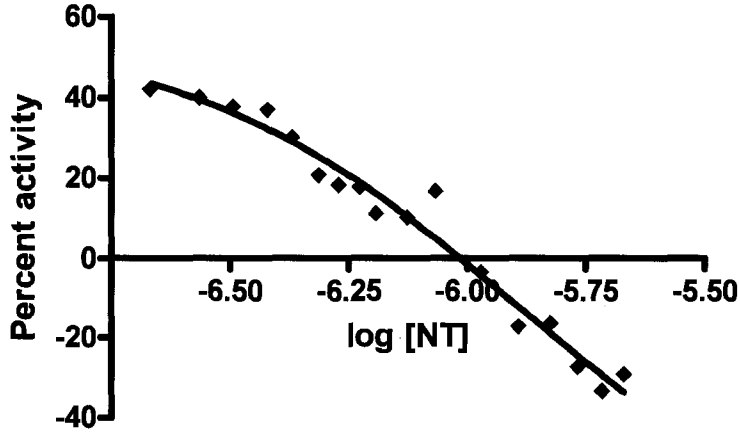
To characterize the effect of increasing MKP3:NT concentrations on MKP3:CT and MKP3:WT activities, the standard pNPP assay with 33% DMSO was used as detailed in Methods (section 2.1.16 and 2.1.18). To measure inhibition, data were plotted as a percentage of activity, with the activity of MKP3:CT (with 33% DMSO) in the absence of MKP3:NT set to 100%. The activity of the unactivated MKP3:CT (no DMSO) in the absence of MKP3:NT was set to 0%. Figure 21A shows that when MKP3:NT was added up to a 2.2-fold excess, the result was a >80% inhibition of the DMSO-activated MKP3:CT activity. A curve was fitted to the data using non-linear regression and a sigmoidal-dose response model (Methods, section 2.1.18). Analysis of the curve fit indicated that MKP3:NT's half-maximal inhibition concentration (i.e., IC_{50}) was ~0.97 μM MKP3:NT. An inhibition constant (K_I) was then determined using this IC_{50} value, which is related to K_I by the Cheng-Prusoff equation (described in Methods, section 2.1.18). K_I was determined to be 0.68 +/- 0.37 μM MKP3:NT. The result shows that MKP3's N-terminal domain can inhibit MKP3's C-terminal domain activity.

Figure 21. MKP3:NT inhibition of MKP3:CT and MKP3:WT. The effect of increasing MKP3:NT concentration on MKP3:CT and MKP3:WT activities was tested as described in Methods (section 2.1.18). Activities are shown as percent activity, with the activity of MKP3:CT in the presence of DMSO (without MKP3:NT addition) assigned a value of 100%. The basal activity of MKP3:CT in the absence of DMSO (without MKP3:NT addition) was assigned a value of 0% activity. MKP3:NT concentrations were added to the assay at the indicated concentrations, plotted using a log scale. **A.** MKP3:CT activity in the presence of increasing MKP3:NT concentration. The experiment was performed in triplicate. **B.** MKP3:WT activity in the presence of increasing MKP3:NT concentration. Negative values indicate MKP3:WT activity was less than the basal MKP3:CT activity. The experiment was performed once.

A.


100
75
50
25
0
Percent activity
-6.75
-6.50
-6.25
-6.00
-5.75
-5.50
log [NT]

B.


60
40
20
0
-20
-40
Percent activity
-6.50
-6.25
-6.00
-5.75
-5.50
log [NT]

MKP3:NT inhibition was also tested on the full-length MKP3:WT. This experiment was performed only once, to confirm the previous experiment, The result in Figure 21B, which is plotted using the same scale as Figure 21A, showed that MKP3:NT could decrease MKP3:WT activity. This was similar to MKP3:NT inhibition of the MKP3:CT activity. The MKP3:NT concentration was not increased to the point that the MKP3:WT activity was fully inhibited and K_I could not be determined. However, the data did show that it was possible to inhibit MKP3:WT activity more than MKP3:CT, which is shown by the decrease in MKP3:WT activity below 0% (the basal MKP3:CT activity). The greater effect is presumably because MKP3:WT contains an attached N terminal domain in addition to the MKP3:NT that was added during the inhibition assay.

The calculated K_I for MKP3:NT's inhibition of MKP3:CT was in a similar range to the SPR-determined K_D for MKP3:NT and MKP3:CT binding. This indicates that MKP3:NT was capable of binding MKP3:CT and that the binding resulted in a decrease in MKP3:CT activity. This shows that MKP3's N-terminal domain possesses an inhibitory function on the C-terminal, catalytic domain. The data was consistent with kinetic results that showed the full-length MKP3 protein was less active in the presence of DMSO than the MKP3:CT, which lacked the N-terminal domain (Table VII).

4.0 Discussion

Soluble MKP3 was difficult to produce in large quantities in an *E. coli* system. MKP3 was found to be expressed as inclusion bodies using the optimized growth conditions determined in section 3.1.3. While inclusion bodies are common in bacteria, previous studies using *E. coli* to produce MKP3 (in the same plasmid/cell combination) have described its production as a soluble protein, with a lower total yield (*13;15;95*). A review of the literature indicates that other research groups may have encountered similar problems when producing MKP3 as they were unable to achieve yields comparable to that shown in Table II using the same pET-15b and BL21(ΔE3) combination. Part of the difficulty may be that the pET-15b and BL21(ΔE3) combination led to a decrease in MKP3 production if the cells were harvested beyond a two hour induction time (shown in Figures 8 and 9). Using the growing conditions that were determined in Figure 9, a large production of MKP3 (approximately 33% of the total cell protein, Table II) was obtained. Unfortunately, under these conditions, MKP3 is produced as inactive inclusion bodies, but the potentially large yield from the insoluble fraction made it an appealing source of MKP3, if a refolding protocol could be determined.

To maximize the yield of active protein, a way to screen various refolding protocols was needed. The difficulty was that there was no existing assay to measure MKP3 activity in the absence of ERK. While ERK could have been added to the assay conditions used to test each refolded MKP3, for my purposes, ERK samples were purchased in very small quantities and were an extremely valuable resource. To avoid using ERK, a novel two-part refolding assay was developed using precipitation screening and chemical activation of MKP3. MKP3 was refolded by dilution and screened according to Methods (sections 2.1.5 and 2.1.16),

which showed that MKP3 possessed poor activity in high ionic strength buffers and required buffer exchange into TBS for full activity (Table III and Figure 11). To avoid a cumbersome, multi-part refolding method, a dialysis-based refolding method was tested. This method exhibited very little precipitation and showed a comparable activity to the buffer exchanged sample. Following a successful characterization of MKP3's refolding, dialysis was adopted for all other MKP3 refolding.

Circular dichroism spectroscopy was used to assess secondary structure in refolded proteins. The data showed that all proteins exhibited a mixed α-helix and β-sheet composition, which was expected for the properly refolded MKP3s. However, there were some differences between the experimental and the reference data (Table IV). The major difference was in the N-terminal domain. Examination of both the CD data in Table IV and previously reported NMR structure (*75*) indicated a greater percentage of β-sheet and turn/coil motifs in the NMR structure. This difference was likely due to the buffer system, which contained 200 mM urea. Urea, as a chaotrope, acts by disrupting hydrogen bonding. This effect may have stabilized a partially unfolded MKP3 structure during NMR data accumulation. The NMR-derived PDB, showed that 66% of the protein existed as β- or γ-turn or random coil. The lack of secondary structure could indicate a "molten globule" or partially unfolded state. This result was in contrast to CD data obtained in TBS buffer (isotonic, pH 7.4) without urea. Under these conditions, the N-terminal domain showed a greater percentage of α-helix and β-sheet structure, which indicated a more folded protein.

A second observation, from CD spectroscopy, was in the C-terminal domain structure (Table IV). The MKP3:CT protein used in this study contained an extended amino sequence compared to the referenced x-ray crystal structure (*74*). The differences were in amino acid

regions 155-203 and 348-381, which were absent in the x-ray structure. These regions have been modelled in previous studies as linear, flexible chains (*75*), but as the CD data showed, these regions contributed to the overall ordered structure of MKP3:CT. Therefore, the sequences are unlikely to be unstructured as previously envisioned. If the 155-203 and 348-381 regions had contributed exclusively to turn and random coil motifs, the predicted secondary structure content of the MKP3:CT protein would be 24% helix, 13% sheet and 64% other motifs. The MKP3:CT CD data showed 43% helix, 24% sheet and 33% other motifs, which indicated that the two regions exhibit an organized structure in solution. This finding is intriguing because the current models presents the connecting region between MKP3's N- and C-terminal domains (approximately amino acid region 155-203) as an unstructured, flexible linker (*74*). This is unlikely. It is known that the linking region possesses a nuclear export signal (-^{164}VLGLGGLRISSD^{175}-) as well as an ability to increase the affinity of MKP3's N-terminal domain for ERK, through an FXPF-like sequence (-^{364}FTAP^{367}-). These properties suggest that the two amino acid regions possess characteristic structures, presumably required for their function or presentation on the surface of the MKP3 protein. Unfortunately, the CD data cannot determine if 155-203 and 348-381 regions comprise distinct domains, NMR or x-ray spectroscopic studies will be required for that determination, but the data did show that the regions possessed an organized secondary structure, which is opposite to the existing model.

Following MKP3 production and characterization of its refolding, MKP3 binding was analysed. Control SPR experiments were performed, which showed that the full-length MKP3 was capable of binding ERK2, but not JNK1 (Table IV). These experiments were followed by experiments to test for MKP3:CT binding. MKP3:CT was tested against MKP3:WT, MKP3:NT and MKP3:CT (Table IV, Figure 13). The results of the experiments

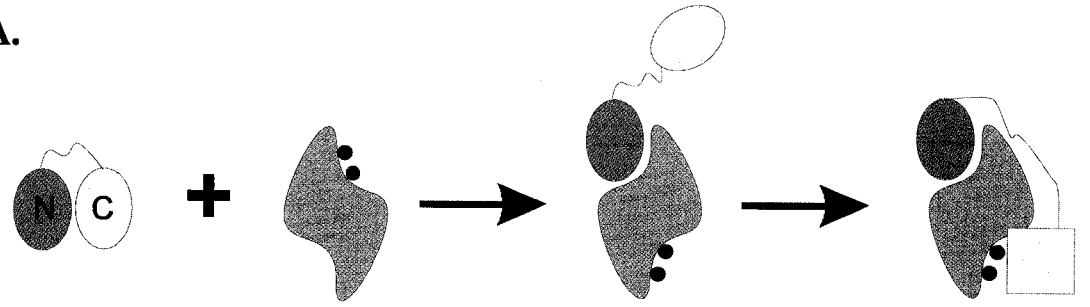
showed that MKP3's N-terminal domain bound to the C-terminal domain. While previous studies have suggested that this interdomain binding exists, the reported values were determined with a relatively insensitive NMR technique using the MKP3`s N-terminal domain and the PAC1 catalytic domain (an MKP3 catalytic domain analogue). The current study calculated the actual value for dissociation constant between MKP3's two domains, which was found to be approximately 70-fold greater than the NMR estimate (*75*). This was likely due to the differences in amino acid sequence between MKP3 and PAC1.

The SPR-assessed strength of the N- to C-terminal interaction (K_D=1.43 +/- 0.41 μM) implies that MKP3 exists in an internally bound state under normal physiological conditions. When compared with the MKP3 to ERK2 interaction (K_D=0.26 μM), the data show that ERK2 possessed a greater potential for binding. This may indicate that ERK2 can displace the N-terminal binding domain from a C-terminal domain due to competitive binding. It also suggests that if N- to C-terminal domain binding results in shielding of the active site, then competitive N-terminal to ERK binding that can release the C-terminal domain, would allow substrate access. This model may partially explain MKP3`s lack of activity in the absence of ERK (Figure 22A).

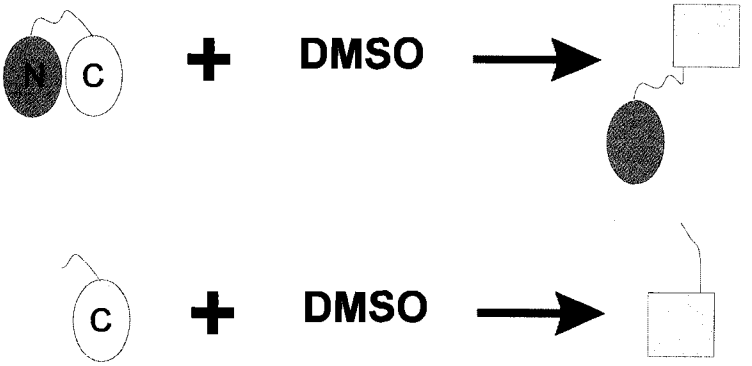
The N- to C-terminal binding was further characterized using peptides chosen due to their roles in MKP3 regulation (i.e., the MKP3 and ERK binding peptides and KIM region described in section 2.1.13) or selected using data obtained from NMR titration (*75*) and H^+/D^+ exchange experiments (*77*) (the post-KIM region). When tested with MKP3:CT, only the post-KIM region peptide exhibited an ability to bind (Figure 14). The post-KIM sequence was originally selected as a potential binding site due to H^+/D^+ exchange data (*77*) that

Figure 22. Models of MKP3 activation in the presence of ERK and DMSO. The results of the binding studies and DMSO-activation indicate that MKP3 likely exists in an internally bound state due to an interdomain binding site. This binding can shield the MKP3 active site and prevent non-specific hydrolysis in the absence of ERK. The following figures illustrates the methods of MKP3 activation (ERK binding and DMSO activation) and a method for inhibition. **A.** ERK activation. In the presence of ERK, competitive binding may lead to two steps: 1. The displacement of the unactive C-terminal domain (yellow oval) from the N-terminal domain (green oval) by ERK binding 2. The free C-terminal domain undergoes allosteric activation (shown as yellow square) induced by the bound ERK. **B.** DMSO activation. DMSO added to either the full-length MKP3:WT or the truncated C-terminal domain leads to activity similar to ERK-activated MKP3. This implies that DMSO mimics the ERK-induced allosteric activation of MKP3's catalytic domain (shown as a yellow square). **C.** MKP3:NT inhibition. Inhibition studies showed that the MKP3 N-terminal domain could inhibit the C-terminal domain activity, when assayed in the presence of DMSO. This can either occur due to shielding of the active site residues in the catalytic domain or by stabilizing the unactivated MKP3 structure (represented by the interchange between oval and square conformations).

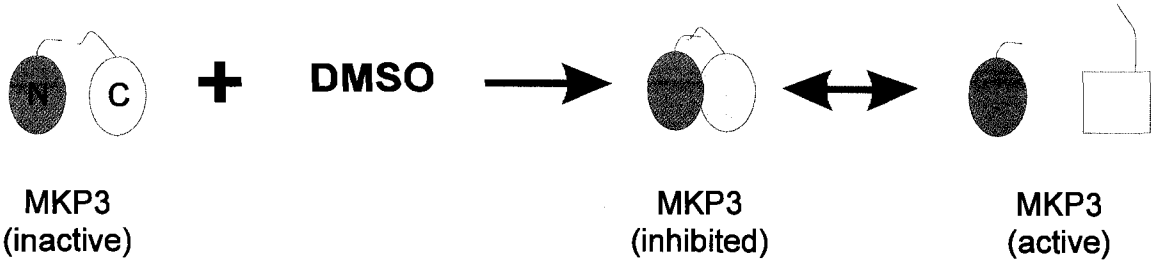
A.


N
C
MKP3
(inactive)
ERK
MKP3/ERK
binding
MKP3
(active)

B.


N
C
DMSO
C
DMSO
MKP3
(inactive)
DMSO activated
MKP3

C.


N
C
DMSO
MKP3
(inactive)
MKP3
(inhibited)
MKP3
(active)

indicated a significant increase in exchange occurred at the region, when MKP3 was bound to ERK2. The region also contained an unusual cluster of positive and negative charges leading to the expectation that the hydrophilic residues would be surface exposed at all times. If true, then no net H^+/D^+ exchange would occur whether MKP3 was bound to ERK or not. Additionally, when these data was compared to the known structure of the N-terminal domain, the post-KIM region was found to compose the α5, surface-exposed helix on the **1HZM** structure (Figure 23). The increase in deuterium exchange can only be explained if this region is normally solvent shielded, but becomes exposed when MKP3 binds ERK2. Together with the result showing that the post-KIM peptide was able to bind to the MKP3 C-terminal domain, these data indicate that post-KIM region contains the MKP3-interdomain binding site.

To determine the region on the corresponding C-terminal domain that the post-KIM peptide might bind to, published NMR transfer experiments between the MKP3 N-terminal domain and a PAC1 domain (*75*) were examined. Amino acids on the PAC1 structure that showed the greatest chemical shift changes (>0.30 ppm) in the presence of the MKP3 N-terminal domain were identified as potential interdomain contact points (Table VIII). Since the PAC1 and MKP3 catalytic domains show 48.6% sequence identity, an identity-weighted amino acid alignment was generated to find the matching residues on MKP3 (Figure 24). MKP3:CT mutants were created by altering the residues that showed the greatest magnetization transfer (Table VIII, shaded in grey). The mutants were tested for binding to the post-KIM peptide by SPR analysis, which showed that mutation of an asparagine (N267) to an alanine resulted in a loss of binding. This implied that the post-KIM and N267 on MKP3:CT comprised most, if not all, of the MKP3 interdomain binding site. Data also

Figure 23. 3D structure of the post-KIM peptide and MKP3:CT amino acid substitutions. A. The post-KIM structure was shown using the NMR structure of the MKP3's N-terminal domain, which was determined by Farooq et al. (PDB entry **1HMZ**; (*75*)). The peptide was located in the MKP3's N-terminal domain structure and then displayed using InsightII (Accelyrs Software Inc.). The post-KIM peptide was identified as the region that formed the α5 helix in the MKP3 N-terminal domain. The nature of the charged amino acid sequence and its helical structure indicate that the region should be surface exposed. **B.** The C-terminal catalytic structure was determined by Stewart et al. (PDB entry **1MKP**; (*74*)) The catalytic triad consisting of D262, C293, and R299 are shown. Additionally, amino acid homologues of PAC1 that were shown to possess unusual magnetization transfer characteristics (from Table VIII) are labelled.

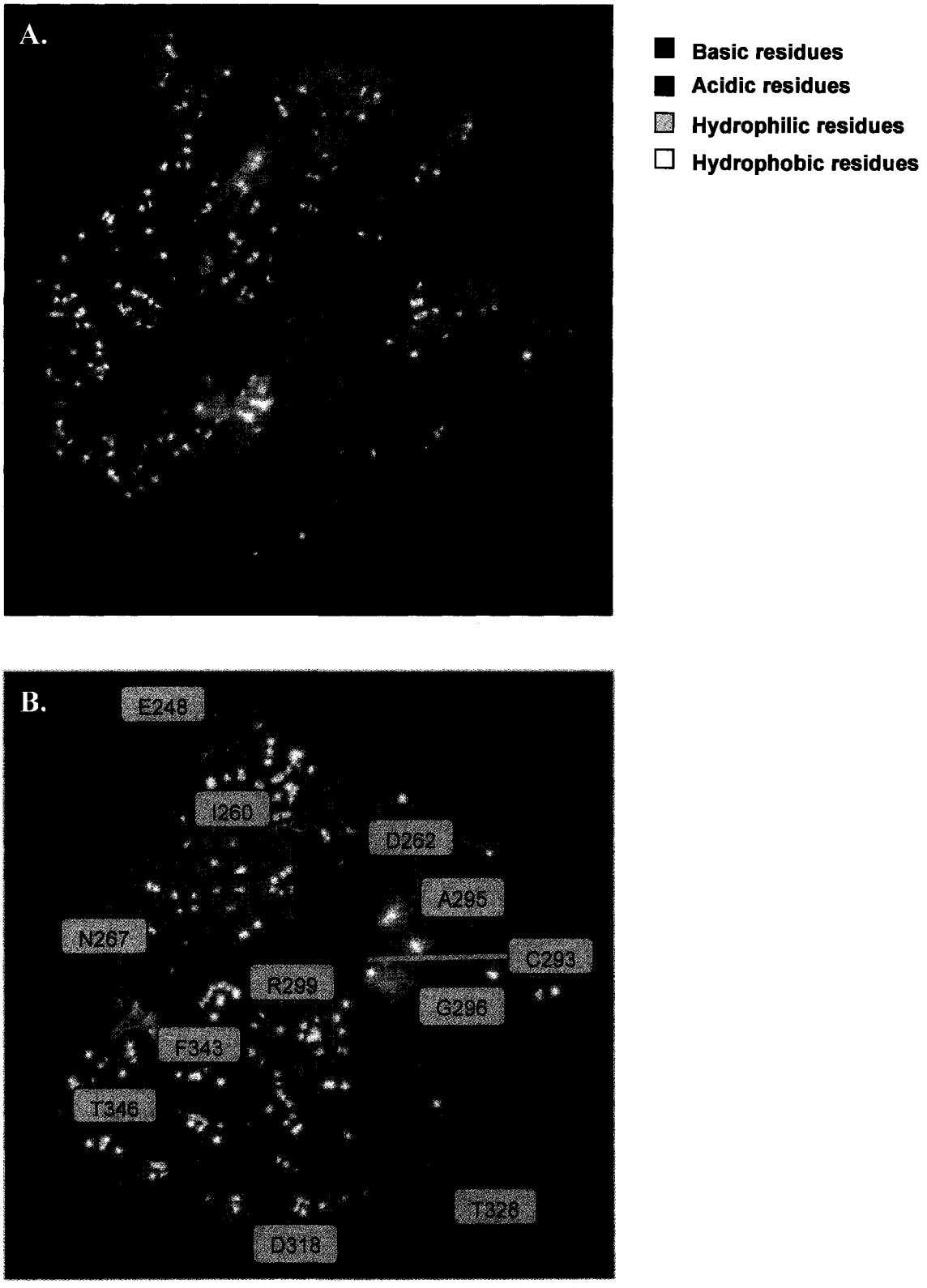

A.
Basic residues
Acidic residues
Hydrophilic residues
Hydrophobic residues
B.
E248
I260
D262
A295
N267
C293
R299
G296
F343
T346
T326
D318

Figure 24. PAC1 and MKP3 catalytic domain sequence alignment. A sequence alignment was generated to compare PAC1 and MKP3. The C-terminal domain amino acid sequences and structures of the two proteins were used to obtain an identity-weighted sequence alignment using the Modeller package in InsightII. Sequence alignment shows: Identical residues (green) and conservative substitutions (orange).

Sequence alignment:

```
P_1M3G: QGGPVEILPYLFLGSCSHSSDLQGLQACGITAVLNVSASCPNHFE--GLF
P_1MKP: ASFPVEILPFLYLGCAKDSTNLDVLEEFGIKYILNVTPNLPNLFENAGEF

P_1M3G: RYKSIPVEDNQMVEISAWFQEAIGFIDWV--KNSGGRVLVHSQAGISRSA
P_1MKP: KYKQIPISDHWSQNLSQFFPEAISFIDEARGKNCG--VLVHSLAGISRSV

P_1M3G: TICLAYLMQSRRVRLDEAFDFVKQRRGVISPNFSFMGQLLQFETQVLCH
P_1MKP: TVTVAYLMQKLNLSMNDAYDIVKMKKSNISPNFNFMGQLLDFERTL---
```

showed a loss of binding when glutamic acid (E246) was replaced, but the result must be taken cautiously due to a potential sensor chip failure that occurred during data acquisition.

Table VIII. PAC1 chemical shift and analogous MKP3 residues

	PAC1	MKP3
Δδ (ppm) > 0.30	E214	E248
	V224	I260
	E231	N267
	A259	A295
	G260	G296
	E282	D318
	G292	S328
	F301	F343
	T309	T346

Following the binding studies, the potential patho-physiological effects of MKP3 binding were explored. First, pancreatic adenocarcinoma cells, which overexpressed MKP3 were analysed using a non-denaturing Western blot (Figure 16). This result showed that high molecular weight oligomers of MKP3 existed in the cells as well as recombinant MKP3 controls. Given the data obtained from SPR binding studies, which showed that the N- and C-terminal domains could bind, the formation of oligomers is likely caused by the intercalation of N- and C-terminal domains between MKP3 proteins (i.e., "domain swapping", (*96-98*)).

The MKP3 oligomerization may also provide a reason for the differences in MKP3 detection that were observed between this thesis and previous literature (Figure 15 and (*22*)). The Western blot results in this study, that positively identified MKP3 in pancreatic adenocarcinomas, used a monoclonal antibody raised against the entire MKP3 amino acid sequence. Previous studies have used a goat polyclonal antibody raised against a 20 amino

acid sequence from MKP3's 351-381 a.a. region (exact sequence not divulged by the manufacturer, Santa Cruz Biotech). Using the PAC1 catalytic domain structure as a template, the MKP3 351-381 region is predicted to form a helix near the active site of C-terminal domain. If N-terminal binding shields the active site, as predicted in the results of this thesis, the size of the N-terminal domain would sterically block antibodies from binding the epitope. The inability of the antibody to detect MKP3, and the different results in the current thesis and previous literature, might be explained by MKP3 N- to C-terminal binding and oligomerization in the adenocarcinomas that were tested.

To test the potential of effect of interdomain binding on MKP3 activity, an ERK-free phosphatase assay system was developed. The initial tests used various solvents at increasing concentrations to identify the property of DMSO that allowed it to activate MKP3. The results in Figure 17 showed that among the tested solvents, only DMSO was capable of chemically activating MKP3. It was not possible to determine how DMSO affected MKP3's activity by using the additional co-solvents. While the solvents were chosen for their structural or chemical similarity to DMSO (Table VI), it did not appear that any trend in MKP3 activation existed. While a root cause for the activity increase could not be determined, the DMSO-activated MKP3 activity is not without precedence. Several other enzymes have been shown to exhibit increased catalytic rates in the presence of DMSO (*99;100*).

The effects of ionic strength on MKP3 function were characterized under unactivated and DMSO-activated conditions. In both tests, the ionic strengths showed an effect on MKP3 activity. Increasing ionic strength led to a decrease in DMSO-activated enzyme activity, but unexpectedly, the unactivated MKP3 samples showed a gain in activity with increasing ionic strength. These results suggest that ionic interactions are partially responsible for MKP3's

deactivation in the absence of ERK. By disrupting ionic interactions, with the increased Na^+ and Cl^- ions, MKP3's activity was increased. Typically, ionic interactions are involved in stabilizing intra- or intermolecular binding through salt bridging. As purified MKP3:WT was assayed in the absence of other proteins, this result indicated that an intra or intermolecular bond may aid in maintaining MKP3's low basal activity. The increase in MKP3 activity that was observed by increasing ionic strength may be similar to the effect that was observed by Zhou and coworkers who showed that MKP3's activity could be increased by using increasing concentrations of glycerol (*15*). Their data showed a maximum ~12-fold increase in catalytic efficiency over the basal value. This effect was likely not related to ERK-allosteric activation, since the addition of ERK resulted in a further increase to catalytic efficiency. The lack of full activation by glycerol alone may indicate that the actual effect observed by Zhou (*15*) was a disruption of inter- or intra-MKP3 binding through disruption of hydrogen bonding. This may be similar to the effect that was observed using high ionic strengths in Figure 18B, which resulted in a similar increase in MKP3 activity.

The ERK-free, activity assay was developed to test both MKP3:WT and MKP3:CT activities. The result in Figure 19 is the first example of an ERK-binding deficient MKP3 protein (MKP3:CT) in an activated state. DMSO was used to chemically reproduce the ERK activation in both full-length and truncated MKP3 proteins without the need for either ERK or even the N-terminal ERK-binding domain. When compared with MKP3:WT activities measured in the presence of ERK2, the DMSO-based assay showed very similar k_{cat} and K_M values, implying a similar mechanism of pNPP hydrolysis and therefore allosteric activation. A model of the potential interaction of MKP3 and DMSO is shown in Figure 22B.

pNPP assay using DMSO as an activator showed that under similar assay conditions, MKP3:CT possessed a greater k_{cat}/K_M than MKP3:WT (Table VII). This result indicated that

the removal of the N-terminal domain from MKP3:WT led to significantly greater activity in the remaining catalytic domain (MKP3:CT became a better enzyme). The effect could not be detected by the standard ERK-based assay, which do not activate the truncated MKP3 catalytic domain. However, the novel DMSO-based assay showed that the MKP3, N-terminal domain possessed an inhibitory function.

The DMSO-based activity assay was also used to measure the activity of MKP3:CT in the presence of various MKP3:NT concentrations. The result of the experiment (Figure 21A) showed that interdomain binding possessed an inhibitory effect. When the K_I was determined using the MKP3:CT and MKP3:NT activity data, the K_I was found to be comparable to the MKP3:NT and MKP3:CT dissociation constant (K_D) determined by SPR studies. Since K_I refers to the equilibrium dissociation constant for an inhibitor (*87*) and K_D refers to the equilibrium dissociation constant for a ligand, if K_D and K_I are similar, the inhibition of MKP3:CT by MKP3:NT must be due the same binding interaction or as modelled in Figure 22C, MKP3:NT binding directly inhibits MKP3:CT function.

These results are different from previous inhibition studies using MKP3's N-terminal domain. While the N-terminal domain has been shown to inhibit wild type MKP3 activity (*12*), the study was performed to test ERK binding. The authors of that study chose to observe the activity of ERK-activated MKP3 in the presence of increasing concentrations of the N-terminal MKP3 domain protein. This experiment led to a decreased activity due to competitive binding between the MKP3, N-terminal domain and full-length MKP3 for ERK2. In this study, we have exploited the ability of DMSO to uncouple the activity of MKP3 from the presence of ERK. The assay used in this study has allowed the testing of MKP3 binding domain (MKP3:NT) on the DMSO-activated catalytic domain (MKP3:CT),

which showed that the N-terminal domain plays a role in modulating the C-terminal domain activity.

The possible *in vivo* function for interdomain binding may be due to MKP3's role as a tightly controlled regulator of ERK1 or ERK2. Previous data shows that when either ERK isoform is available, MKP3 undergoes a structural reorganization at its catalytic domain, which leads to full enzyme activity (*12*). Allostery is a common method used to develop enzyme specificity, but the results of this study suggest inhibition of MKP3 in the absence of ERK is a second, previously uncharacterized, specificity-generating technique. Consistent with the findings is a model of MKP3 catalytic domain (MKP3:CT) deactivation through binding to the MKP3 non-catalytic domain (MKP3:NT). The binding may either stabilize the inactive enzyme structure or block the MKP3:CT active site, both of which can inhibit non-specific activity. In the presence of ERK2, competitive binding and a lower K_D value, which favours ERK and MKP3:NT binding, could result in the release of the MKP3 catalytic domain. This can enable the subsequent structural reorganizations that are necessary for high specific activity hydrolysis.

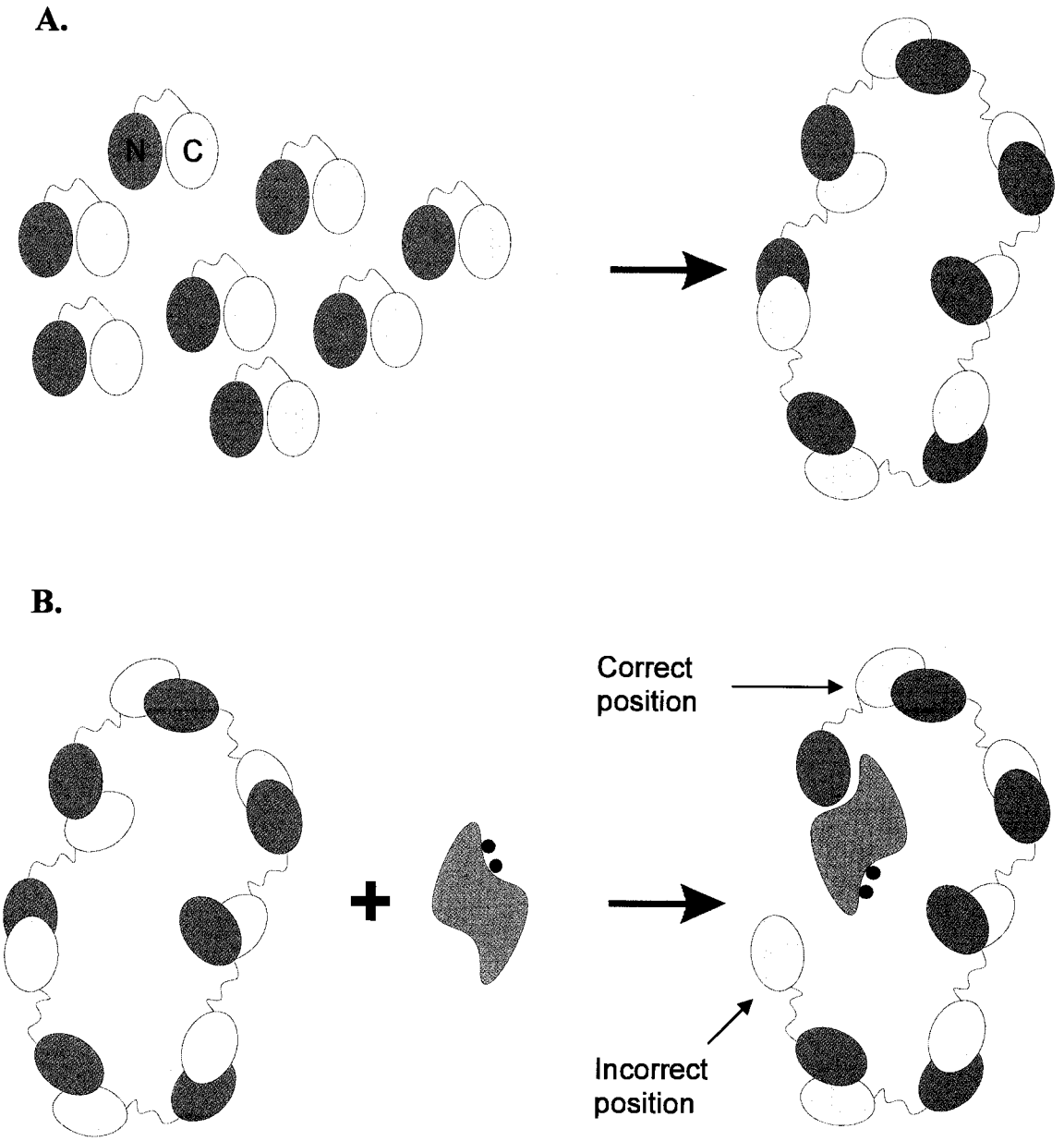
The effects of control by this interdomain binding will likely affect pancreatic adenocarcinomas. The results in Figures 15 and 16 show that the wild-type MKP3 is overproduced in pancreatic cancers that have been shown to produce constitutively active phospho-ERK (*22*). A new model for MKP3 deactivation may explain this apparent contradiction: If MKP3 overproduction reaches a critical concentration; MKP3 may itself begin to out-compete ERK for the MKP3 binding. This can lead to MKP3 oligomer formation as observed in Figure 16. The oligomers are predicted be totally inhibited and possess no phosphatase activity due to perpetual N- to C-terminal binding. The proposed model is depicted in Figure 25A, and shows oligomerization of MKP3 through N- to C-

terminal binding, a classical "domain swap" interaction (*101-103*). The N-terminal inhibitory effect that occurs will lead to the loss of MKP3 function. In this state, and even if ERK could subsequently bind to an MKP3, N-terminal domain, the interaction will not necessarily lead to hydrolysis. This is because the MKP3, C-terminal domain that would be released by ERK would be at an incorrect position to bind ERK's phosphorylated "activation motif" (Figure 25B). Additionally, if ERK could bind an N-terminal domain, the N-terminal's conjoined C-terminal domain would still be bound to another MKP3 N-terminal domain and remain inhibited (Figure 25B). This model thus explains MKP3's presence in pancreatic adenocarcinomas and its loss of function.

The self-inactivation mechanism proposed in this thesis is similar to one recently described for src-homology 2 domain-containing tyrosine phosphatase, which possesses an autoinhibitory domain that blocks the phosphatase catalytic site (*104*). It is also similar to a domain-domain interaction that was recently described for protein tyrosine phosphatase α (PTPα) where interaction between internal domains favour the stabilization of a dimeric state until ligand binding disrupts this interaction and leads to activation (*104;105*). The data from this study suggests that a similar scheme might be applicable to other members of the dual specificity and multi-domain phosphatases (*106*). In MKP3, the dimerization/oligomerization could provide an effective means for restricting the attenuation of RAS-MAPK signalling to within physiologically acceptable limits. The added level of regulation complements the feedback loop elicited by downstream ERK2 activation and nuclear translocation (*12*). This proposal is consistent with significantly elevated MKP3 protein levels seen in the tumour cell lines used in this study and in the dysplastic pancreatic adenocarcinoma specimens reported previously (*22*). Thus, elevated MKP3 does not appear to exercise intrinsic tumour suppressor activity. Self-inactivation by oligomerization becomes all the more relevant in

this context since loss-of-function mutations in MKP3 have not yet been found in cultured cancer cells or clinical specimens (*19*).

Figure 25. Model for MKP3's loss of function in pancreatic adenocarcinomas. The results of the binding and autoinhibition studies supported a model of MKP3 deactivation due to oligomerization of MKP3. The effect may result in a loss of MKP3 function and subsequent overproduction of phosphorylated ERK2. Two steps are proposed. **A.** MKP3 overproduction leads to oligomerization. In situations of MKP3 overexpression, the interdomain binding site in MKP3 may lead to intercalation of the N- and C-terminal domains between MKP3 subunits, or domain swapping. **B.** Oligomerization prevents substrate-induced MKP3 function. The high molecular weight oligomers may fail to undergo allosteric activation in the presence of ERK due to improper release of interdomain bound MKP3. Due to oligomerization, release of a catalytic domain will likely be in the wrong position to bind and dephosphorylate ERK. The correctly positioned catalytic domain would be inhibited by remaining bound to an MKP3, N-terminal domain.


A.
N
C
B.
Correct
position
Incorrect
position

PART II

5.0. Introduction:

5.1. SARS coronavirus

The second project contained in this thesis is a study of the SARS (severe acute respiratory syndrome) nucleocapsid protein (SARS N protein), a component of the SARS coronavirus. SARS was first described as an atypical pneumonia in the Guangdong province of China in November of 2002 (*107-109*). The infection spread rapidly and by February 2003 the world health organization had identified 305 affected individuals, with 30% being heath care workers (*109*). Following its identification in late February, SARS had captured international attention and led to a global alert on March 15, 2003. Eventually on April 16 2003, the causative agent of the SARS illness was identified as the SARS coronavirus (*108*). This virus belongs to the coronaviradae family of enveloped, positive-sense RNA viruses. The family is notable as coronaviruses contain a sizeable amount of viral RNA (27-32 kb) and employ a unique, discontinuous mechanism for RNA transcription. The method leads to a high frequency of recombination and mutation (*110-112*), which is further enhanced since viral RNA polymerases lacks the inbuilt repair mechanisms found in analogous DNA polymerases.

The major breakthrough that followed the identification of the SARS coronavirus was its sequencing (*107*) that resulted in the identification of four open reading frames corresponding to the N, M, E and S structural proteins. The SARS N protein has been found to act as a nucleocapsid protein that binds to a specific packaging signal on the viral RNA. It is the interaction of this complex with M (membrane), E (envelope) and S (spike) proteins that leads to budding through the cell membrane and virus maturation (*113-115*).

5.1.1. SARS N protein characteristics

The SARS N protein is a 422 amino acid, 46 kDa protein with a pI = 10.1 (*116*). The high pI is characteristic of many DNA or RNA binding proteins and it is due to the high percentage of basic amino acid residues in the primary structure (54%). Other unusual characteristics of the N protein are the low percentage of hydrophobic amino acids (which constitute only 17% of the total sequence) and the total absence of cysteine residues (*117*). The atypical composition of the primary structure suggests that the SARS N protein will lack a substantial hydrophobic core as well as stabilizing disulfide bonds. This feature is unusual, even when compared to other N proteins within the coronavirus family (*118*). While the peculiar composition of the protein may lead to the SARS N protein's RNA binding properties, the potential lack of intramolecular interactions and binding could indicate that the SARS N protein possesses an unstable structure. This was proved in recently published reports using fluorescence spectroscopy to measure the thermal denaturation of SARS N protein. The results of the studies showed a lack of protein stability, which was exhibited as a partial unfolding or conformational change at relatively low temperature and low pH (*116;118*). The study by Wang *et al. (116)* indicated that the SARS N protein's melting temperature is low: the protein begins to unfold at 35°C and is completely unfolded at 55°C, with a melting temperature, T_m = 43°C. These data can be compared to that of Duan *et al.* (*119*), who showed that heat induced attenuation of the virus occurs at 56°C. For comparison, other comparable virus proteins unfold at much higher T_m (50-60°C, (*120-123*)).

Wang's results showed that mildly acidic pH was able to affect protein stability. In a second experiment, an acid titration was used in combination with fluorescence spectroscopy to observe the effect of pH on the SARS N protein. The results show that denaturation began at a pH of ~5 and SNP was completely denatured by pH 2.7 (*116;116;118*). This result

supports data presented by Huang *et al.* (*124*) who showed that the SARS coronavirus enters the cell through an endosome-mediated process and that inhibitors of endosomal acidification prevent SARS infectivity. Their data supported a role for acidic pH on the release of viral RNA from the SARS N protein, which is likely due to an acidic pH-induced SARS N protein unfolding *(116)*. More recently, 3D structural data has been obtained from both the N protein's multimerization and RNA binding domains (PDB entries **2GIB** (*125*) and **2OG3 (*126*)**, respectively). These data showed that the nucleocapsid contains a compact C-terminal dimerization domain (a.a. 270-370), but an unstructured N-terminal RNA binding region (a.a. 50-174). Very little is known about the intervening linker region.

In addition to structural instability, the SARS N protein may display a protein lability, which has been described as a susceptibility to cleavage by caspases (*127*). Proteomic analysis on the SARS N protein produced in eukaryotic cells showed N protein-derived cleavage products in the 20 and 30 kDa ranges on SDS-PAGE. This cleavage was attributed to apoptosis upon Vero (monkey) cell infection and thought to be due to a caspase-3 dependant proteolysis (*127*).

5.1.2. SARS binding motifs

The SARS N protein exhibits extensive multimerization that is believed to be involved in the SARS virus packaging and maturation. Several potential intermolecular binding sites have been identified. One of the first to be characterized was the highly hydrophilic serine and arginine rich -^{184}SSRSSSRSRGNSR^{196}- region. This region has been shown to be involved in protein dimerization and the formation of higher molecular weight "dimers of dimer" (tetramers, (*114;115;117;128*)). With confocal microscopy, the native SARS N protein is normally visualized as a halo around the nucleus/nucleolus, but deletion of the 184-196 region abrogates SARS N protein self-association and results in protein

monomers that exhibit a diffuse localization throughout the cytoplasmic compartment (*129*). The loss of epinuclear localization and diffusion away from source of viral RNA transcription will affect SARS RNA binding and packaging.

5.2. Rational and objectives

N protein expressing pQE2:*snp* plasmid in M15 *E. coli* cells was supplied by Dr. Sean Li for use in developing a SARS N protein antibody. Production and refolding characterization was performed, but the experiments led to an unexpected co-purification of smaller molecular weight proteins. Based on the SDS-PAGE estimated molecular weights, the proteins were similar to the proteolysed products that were observed by Ying *et al.* (*127*). While Ying's group attributed the N protein cleavage to caspase-3, *E. coli* do not contain caspases and experiments were designed to determine if the cleavage was due to proteolysis by a non-specific bacterial protease contamination or if it was due to autocatalysis.

To make this determination, several questions needed to be answered:

1. Are the detected protein bands the results of N protein cleavage or non-specific bacterial proteins,
2. Where is the cleavage occurring, what is the cleavage motif, and
3. Determine if cleavage occurs in the presence or absence of contaminating proteases.

6.0. Materials

Materials used in SARS N protein cloning, including the original *snp* cDNA used as a template for bacterial SARS N protein were supplied by the lab of Dr. Xuguang Li (Health Canada, Ottawa, ON). Additional reagents used for *snp* cloning including, M15 cells, pQE-2 plasmid and Mini prep kits, were acquired from Qiagen (Mississauga, ON). BL21(ΔE3) cells were obtained from EMD Bioscience (La Jolla, CA) and IPTG was purchased from Roche (Mississauga, ON). Buffer reagents including Tris base and Tris-HCl, sodium phosphate mono and dibasic, guanidine-HCl, urea, ammonium bicarbonate, LB media, sodium deoxycholic acid (NaDOC), iodoacetamide and trifluoro-acetic acid (TFA) were obtained from Sigma-Aldrich (Oakville, ON). Acetonitrile was obtained from EMD chemicals (Gibbstown, NJ). Material used for mass spectrometry included NP20 SELDI-chips, which were purchased from Ciphergen Biosystems (Fremont, CA). Sequencing grade trypsin was acquired from Promega through their Canadian representative Fisher Canada (Nepean, ON). Material for protein purification including His-trap metal affinity and Superdex G75 size exclusion columns were obtained from GE Healthcare (Piscataway, NJ). A Vydac C8 reverse phase column was purchased from Grace Canada (Ajax, ON). Materials for protease assays and refolding including, Quanticleave ™ protease assay kits and Snakeskin ™ dialysis tubing, were acquired from Pierce Biotechnology through their Canadian representative, Fisher Canada (Nepean, ON). Purified ovalbumin and RNAse A were acquired from Sigma-Aldrich (Oakville, ON). FRET-labelled peptides, which included EDANS/DABCYL-conjugated $-^{168}$LPKGFYAEGSRGGSQASS^{185}- and -^{181}SQASSRSSSRSRGNSRNSTP^{200}-SARS N protein peptides, were purchased from JPT Peptide Technologies (Acton, MA).

6.1. Methods

6.1.1. SARS N protein expression and purification

M15 *(E. coli)* cells that contained the SARS N protein cDNA (in a pQE2 plasmid, pQE2:snp) were provided by Dr. Xuguang Li. The pQE2:snp plasmid was subsequently purified using a Qiagen Plasmid Midi-Prep Kit as per the manufacturer's instructions and transformed into BL21(ΔE3) cells.

Both M15 and BL21(DE3) cells containing the pQE2:snp plasmid were used to produce SARS N protein using the same protocol. A 10 mL starter culture was grown overnight in a 50 mL conical tube. This culture was used to inoculate 1 L of LB broth in a 2.8 L Fernbach flask. Cells were then grown with shaking at 125 RPM at 37°C in an Innova 4000 shaker (New Brunswick Scientific). Cells were induced for gene expression at an OD_{600} of 0.5 by the addition of IPTG (1 mM final concentration). The culture was then grown with shaking for an additional three hours at 37°C before harvesting by centrifugation at 8,000 x g. Cell lysis was performed by sonication using two, one minute pulses separated by a one minute rest period on ice. The insoluble fraction was collected by centrifugation at 20,000 x g for 30 minutes and then cleared of membrane material using a wash solution containing 2% sodium deoxycholate in Tris buffered saline (20 mM Tris, 137 mM NaCl, pH 8.4). The mixture was again centrifuged at 20,000 x g and the remaining insoluble material was recovered. A final TBS wash and centrifugation step was performed before the recovered inclusion bodies were solubilized in TBS supplemented with 6 M guanidine-HCl (final concentration).

Further purification of SARS N protein was performed under denaturing conditions using a His-trap HP metal affinity and Superdex G75 size exclusion columns as per the manufacturers' instructions.

6.1.2. SARS N protein refolding

Purified N protein was refolded by dialysis using Pierce Snakeskin ™ dialysis tubing (7,500 Da molecular weight cut off). The protein was first dialysed against an 8 M urea-supplemented refolding buffer (10 mM Tris, 100 mM sodium phosphate, 150 mM NaCl, 8 M urea, pH 8.0) for four hours. The urea was then removed by a stepwise replacement of the refolding buffer with a Tris/Phosphate buffer (10 mM Tris, 100 mM sodium phosphate, 150 mM NaCl, pH 8.0). At each step, the urea concentration was reduced by half and then allowed to equilibrate for a minimum of four hours. Following dialysis of the protein solution into a 0.25 M urea, two final buffer changes, into 0 M urea, Tris/Phosphate buffer, were performed to eliminate any residual denaturant.

6.1.3. $(His)_6$-tag cleavage

To cleave the N-terminal $(His)_6$-tag, the pH of the buffer was changed to 7.0 by dialysis and then Qiagen DAPase (dipeptidase) was added to the protein solution as per the manufacturer's instructions. The DAPase addition was only performed for the first SARS N protein purification. Subsequent preparations of N protein were used for analyses following dialysis into pH 7.0 buffer, but without the addition of Qiagen DAPase.

6.1.4. SDS-PAGE

Proteins were separated on SDS poly-acrylamide using the 12% cross-link method described by Laemmli (*81*) or the 20% Tricine method described by Schagger (*130*). Gel electrophoresis was performed using a Bio-Rad Mini-Protean 3 tank (Hercules, CA). Staining was performed using a 30 minute fixing step in 10% methanol and 14% glacial

acetic acid and then either one hour staining by Coomassie blue or overnight staining in Sypro Ruby Red stain (Invitrogen, Burlington, ON). Gels were then placed in 10% methanol and 14% acetic acid solution for a 30 minute to two hour destaining with shaking. Visualization and scanning of the gels was performed under white (Coomassie stain) or UV (Sypro stain) trans-illumination using a Bio-Rad Gel-Doc XR.

6.1.5.SELDI-TOF/MS

Surface enhanced laser desorption ionization, time of flight mass spectrometry (SEDLI-TOF/MS) was performed on a Ciphergen (Fremont, CA) PBS IIC instrument. Instrument calibration was performed using Ciphergen All-in-1 peptide standards adhered to a normal phase, NP20 protein array according to the supplier's instructions. One μg of SARS N protein was applied to each of the remaining sample channels for analysis. SELDI-time of flight spectra were then generated by laser desorption/ionization using an average 130 laser shots with an intensity of 190-200 (arbitrary units) and detector sensitivity of 8. Molecular weight detection was set for optimal sensitivity in the 10,000-30,000 Da range with a detection maximum of 50,000 Da.

6.1.6. MALDI-TOF/MS

SDS-PAGE was performed on N protein samples as previously described. Bands of interest were excised and diced to a uniform size (approximately 1 mm^3) then washed with water before resuspension in acetonitrile. The gel slurry was dried and resuspended in 1 mL, 100 mM ammonium bicarbonate, which contained 10 mM DTT to cleave disulphide bonds. To prevent disulphide bonds from reforming, cysteine residues were capped by the addition of iodoacetamide (50 mM final concentration). Gel pieces were washed and rehydrated in 1 mL of 50 mM ammonium bicarbonate and proteins were digested using 20 μg sequencing

grade trypsin. Peptide fragments were analysed by Matrix assisted laser desorption ionization, time of flight mass spectrometry on a Micromass MALDI-LR instrument (Waters, Mississauga, ON). The data were processed using Waters' MassLYNX 3.5 analysis software. To identify the source of tryptic peptides, the MALDI-TOF/MS data was searched against the Mascot peptide fingerprint database (Matrix Science, Boston, MA).

6.1.7. LC/MS

Protein mixtures were separated by reverse phase HPLC (rpHPLC) using a Waters CapLC instrument outfitted with a Waters NanoEase 75 μm x 100 mm, Atlantis dC18 3.5 μm column. Chromatography was performed at 7 μl/min with all solvents acidified with 0.2% formic acid. Peptides were eluted over 60 minute with a 2%-40% acetonitrile gradient and residual peptides removed by capillary purging at 98% acetonitrile. Peptides were analyzed online with nanoflow electrospray ionization quadropole time-of-flight (nESI-QTOF) mass spectrometry using a Waters/Micromass QTOF Ultima Global instrument.

6.1.8. Reverse phase HPLC

Peptide fragments were purified and concentrated prior to mass spectrometric analysis using a Vydac C8 reverse phase column connected to a Thermo Spectra System HPLC (Fisher Canada, Nepean, ON) running ChromQuest Chromatography Data System software (Fisher Canada, Nepean, ON). The SARS N protein fragments that adhered to the C8 column were eluted using 0.01% trifluoro-acetic acid (TFA) and a 10%-90% acetonitrile gradient.

6.1.9. Quanticleave ™ protease assay

The Quanticleave protease assay (using flourescein isothiocyanate, FITC-conjugated casein) was used with samples of the refolded, $(His)_6$-tagged SARS N protein as described by the manufacturer's instructions. Fluorescence was detected using 485/538 nm excitation/emission wavelengths in a Genios plate reader running the XFluor 4 software program (Tecan, Durham, NC).

6.1.10. EDANS/DABCYL-conjugated SARS N protein assay

EDANS/DABCYL-conjugated peptides were designed to correspond to regions – ^{168}LPKGFYAEGSRGGSQASS^{185}- and –^{181}SQASSRSSSRSRGNSRNSTP^{200}- of the SARS N protein, the sites of a putative cleavage motif. Cleavage of either peptide leads to unquenching of the EDANS reagent by DABCYL and results in a 40-fold enhancement of fluorescence. The ability of the assays to detect proteolysis was characterized using the peptide substrates dissolved in PBS at pH 7.2 (0.5 mg/mL final concentration). Peptide solution (190 µl) was added to a 96-well plate to which 10 µl of trypsin was added (at $5 \times 10^{-1} - 5 \times 10^{-8}$ mg/mL, 5 µg – 500 fg/200 µl assay). Fluorescence was recorded using 360/465 nm excitation/emission wavelengths at 1 minute intervals over 60 minutes at 25°C. Appendix A shows that the limit of detection for trypsin was 500 pg/ 200 µl assay or ~2.5 ng/mL.

Assays using N protein solutions were performed using a 1 mg/mL peptide substrate concentration in PBS. 100 µl of the solution was then diluted by mixing with 100 µl N protein solution at a 0.25 mg/mL concentration (25 µg protein). The reaction mixture was then transferred to a 96 well plate and fluorescence was monitored as described above.

6.1.11. Protein mixing experiments

Non-specific protease activity was tested using a modified refolding protocol in which a 10-fold excess of either ovalbumin or RNAse A was simultaneously refolded with the SARS N protein. Ovalbumin and RNase A were prepared by permanent cleaving and capping of disulfide bonds with iodoacetamide, as described in Section 6.1.6. The test proteins were then denatured by addition to 6 M guanidine buffer containing SARS N protein with the concentrations of the added proteins were adjusted, relative to the N protein, to a 10:1 w/w ratio (2.5 ovalbumin or RNAse A and 0.25 mg/mL SARS N protein). Both proteins (SARS N protein and ovalbumin or SARS N protein and RNase A) were refolded in the same dialysis tube using the method described in section 6.1.2. The results of the mixing experiment were analysed by SDS-PAGE separation.

7.0. Results

7.1. SARS N protein production and purification

The SARS N protein was produced using a SARS N protein cDNA that was inserted into a pQE-2 vector. Proteins were produced in M15 or BL21(ΔE3) cells as insoluble inclusion bodies and purified as described in Methods (section 6.1.1). SDS-PAGE separation of proteins from the insoluble fraction of M15 cells is shown in Figures 26A. The Inclusion body fraction, following a metal affinity purification is shown in Figure 26B. The SARS N protein was recovered with >90% purity.

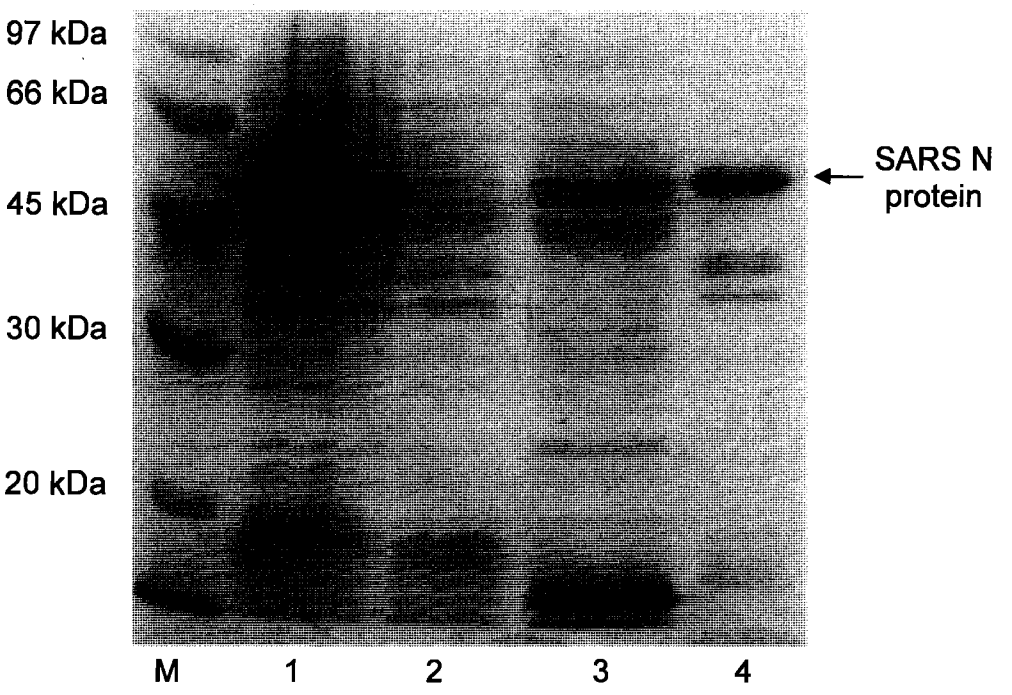
7.1.1. SARS N protein refolding and cleavage

The SARS N protein that was obtained from M15 cells in Figure 26B was refolded by dialysis as detailed in Methods (section 6.1.2). As the protein was refolded, samples of the protein solution were obtained for SDS-PAGE analysis. The SDS-PAGE result in Figures 27, lanes 1-3 show that SARS N protein was a single band under denaturing conditions (lane 1, 8 M urea) up until the dialysis was changed to 0 M urea (lane 3). However, in lanes 4-7 (after the second 0 M urea dialysis change onwards), two additional bands at approximately 29 (called band A) and 25 kDa (called band B) became visible. Because the A and B bands were not observed prior to refolding when proteins were denatured in guanidine (Figure 26B) or in urea (Figure 27, lane 1 and 2), the results indicated proteolysis of the SARS N protein followed refolding.

A potentially confounding factor in this first purification of the N protein was that the appearance of the bands coincided with the completion of refolding and adjustment of the buffer to pH 7.0 (lane 4). This was necessary to cleave an N-terminal $(His)_6$-tag (the

Figure 26. Inclusion body isolation and purification of the SARS N protein. M15 cells, containing the pQE2:*snp* plasmid, were grown according to Methods (section 6.1.1). Samples at each step of SARS N protein purification were mixed with SDS-loading buffer (1:1 ratio) and 10 µl of the mixture was loaded on 12% SDS-PAGE. Staining was by Coomassie blue stain. **A.** Inclusion body isolation. Samples of the SARS N protein obtained during protein recovery were analysed. Lanes M. Molecular weight markers, Lane 1. Whole cell fraction, Lane 2. Soluble protein fraction, Lane 3. Sodium deoxycholate wash, Lane 4. Guanidine-solubilized, insoluble protein fraction. **B.** Purification by metal affinity chromatography. Guanidine-solubilized SARS N protein was purified by metal affinity chromatography, using a 0-250 mM imidazole gradient. Lane M. Molecular weight markers, Lanes 1-4. Metal affinity purified, SARS N protein collected following imidazole elution.

A.


97 kDa
66 kDa
45 kDa
30 kDa
20 kDa
SARS N protein
M
1
2
3
4

B.

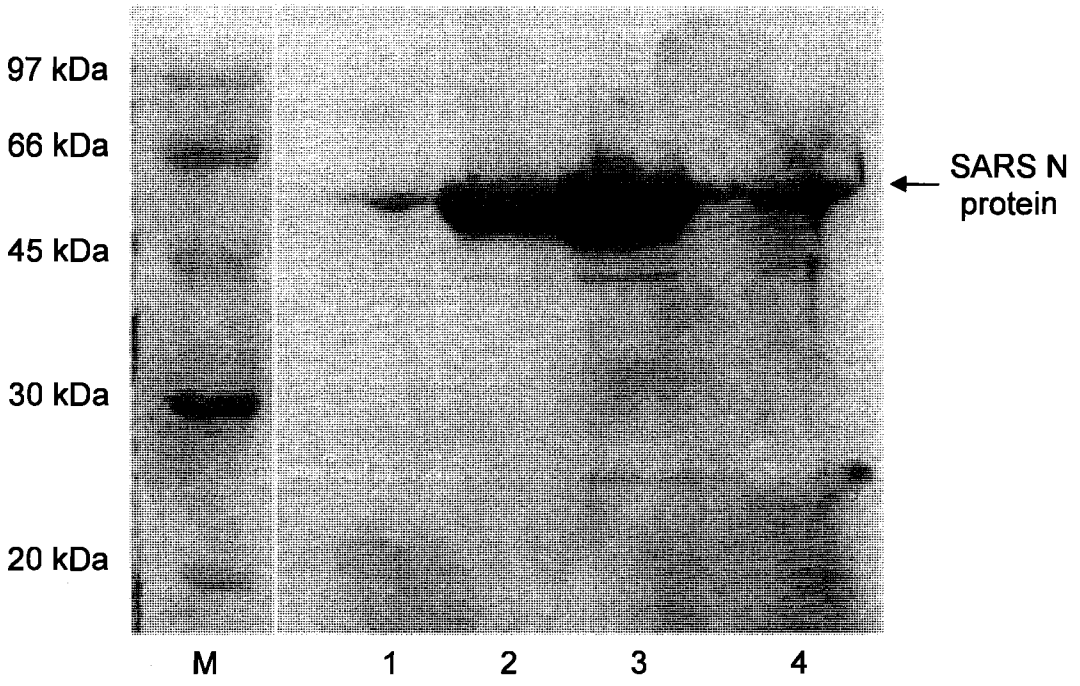
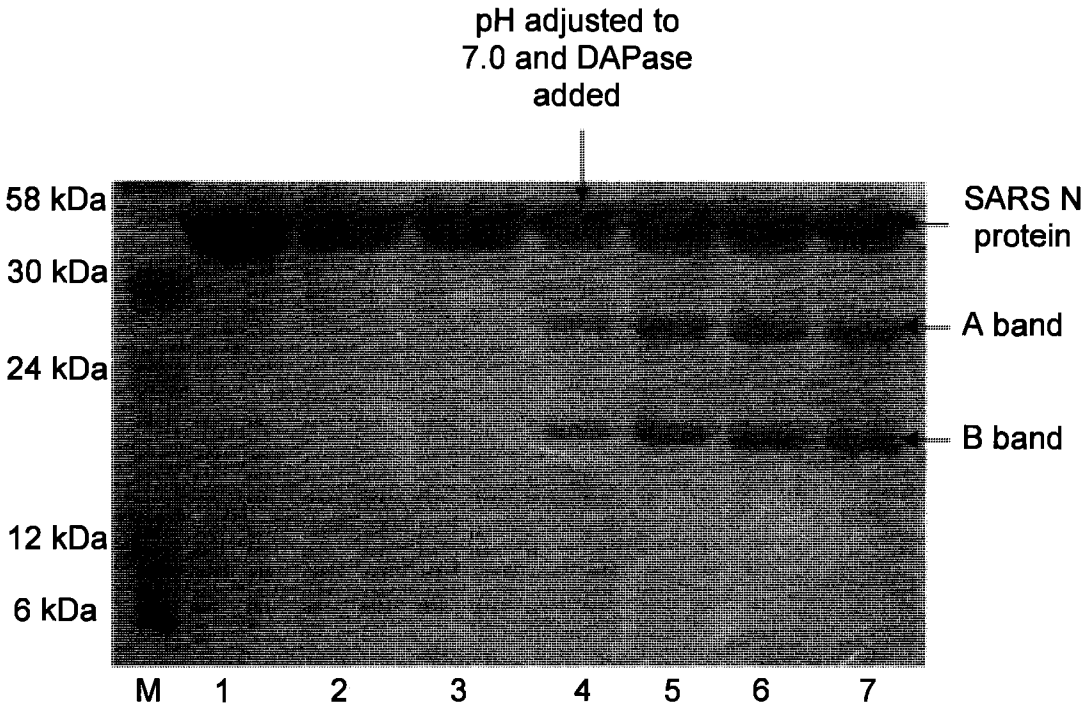

97 kDa
66 kDa
45 kDa
30 kDa
20 kDa
SARS N protein
M
1
2
3
4

Figure 27. SARS N protein refolding and proteolysis. SARS N protein was produced in M15 cells and purified. The protein was refolded by gradual dialysis as described in Methods (lanes 1-3). The engineered $(His)_6$-tag was removed using Qiagen DAPase (lanes 4-7). At each step, protein samples were obtained and mixed with SDS-loading buffer (1:1 ratio). 10 μl of the protein mixtures were loaded onto 12% SDS-PAGE and stained using Coomassie blue. Lane M. Molecular weight markers, Lane 1. SARS N protein in 8 M urea, Lane 2. 0.5 M urea, Lane 3. 0 M urea, Lane 4. 0 M urea (pH adjusted to 7.0 and DAPase added), Lane 5. 10 minutes following DAPase addition, Lane 6. 20 minutes following DAPase addition, Lane 7. 30 minutes following DAPase addition.


pH adjusted to
7.0 and DAPase
added
58 kDa
30 kDa
24 kDa
12 kDa
6 kDa
SARS N
protein
A band
B band
M
1
2
3
4
5
6
7

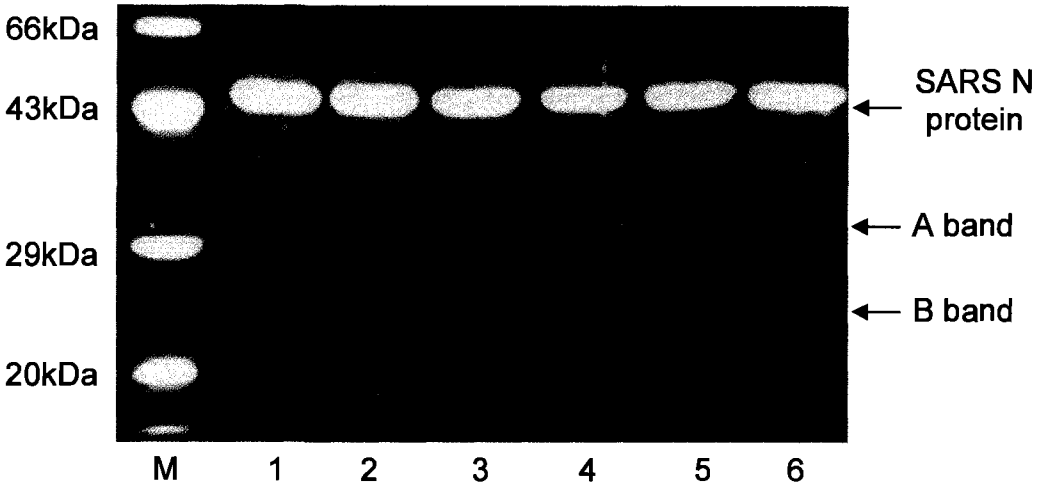
Qiagen, TAGZyme™ tag) with a proprietary dipeptidase (DAPase). The DAPase was added to the SARS N protein during the second 0 M urea buffer change (Figure 27, lane 4). The DAPase functions by cleaving dipeptides from a protein's N-terminal towards the C-terminal end and it is unlikely that the appearance of cleaved product was due to specific action of this enzyme.

To confirm that the A and B bands were not due to DAPase proteolysis during refolding, the N protein purification was repeated without pH 7.0 adjustment, without DAPase addition, and using N protein produced in BL21(ΔE3) cells. Figure 28 shows SDS-PAGE results from samples of the SARS N protein that were obtained following each dialysis buffer change during refolding. The SDS-PAGE was stained using Sypro Ruby Red to detect low abundance proteins (the Sypro stain is ~50-fold more sensitive than Coomassie blue). In Figure 28A, the increased sensitivity of the fluorescent stain showed the presence of co-purified bacterial proteins in the SARS N protein sample (observed between 43 and 29 kDa), but also showed the appearance of A and B bands during 0.5 M and 0 M urea dialysis (lanes 5 and 6). This result reproduced the data shown in Figure 27, which showed that the A and B bands became visible at 0 M urea. The appearance of the A and B bands in both Figures 27 and 28 implied that the protein cleavage was not due to the effect of DAPase and that cleavage was not limited to SARS N protein recovered from the M15 strain of *E. coli.*

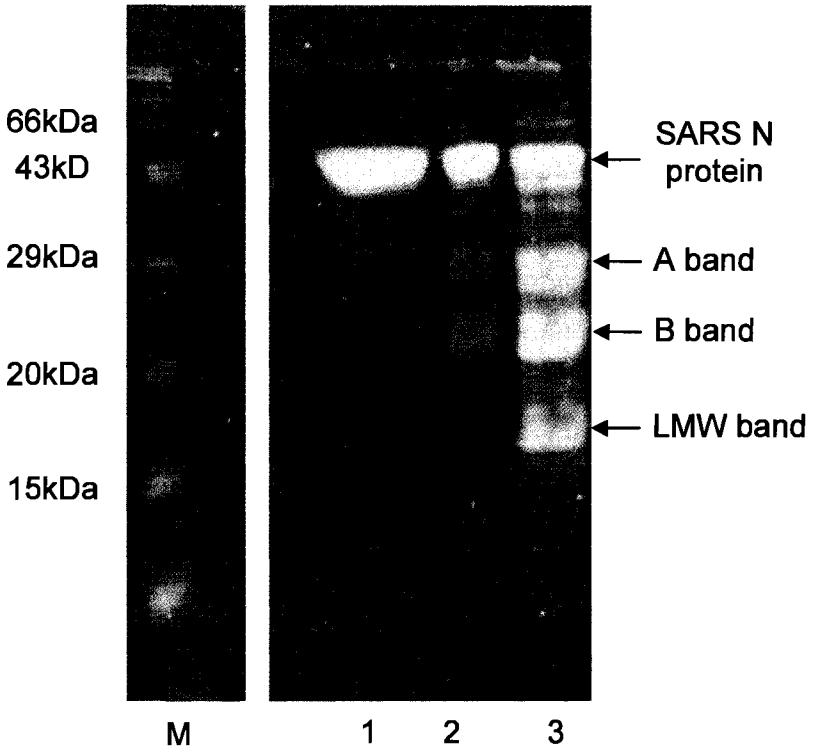
The samples that were obtained from BL21(ΔE3) cells were stored at 4°C for 18 hours (overnight) or 120 hours (5 days) at pH 8.0. These sample and a sample of the SARS N protein that was taken immediately after protein refolding (t=0) were separated on a Tricine gel (20% cross-link). This results in Figure 28B show that the SARS N protein

Figure 28. SARS N protein cleavage. SARS N protein was produced in BL21(ΔE3) cells and refolded as described in Methods (section 6.1.1 and 6.1.2). SARS N protein samples were obtained during protein refolding and mixed with SDS-loading buffer (1:1 ratio). 10 μl of protein mixtures was loaded onto 12% Laemmli SDS-PAGE (**A**) or 20% Tricine gel (**B**). Gels were stained by Sypro Ruby Red stain. **A.** SARS N protein cleavage during refolding. The SARS N protein samples were obtained following the indicated dialysis buffer changes. Lane M. Molecular weight markers, Lane 1. 8 M urea, Lane 2. 4 M urea, Lane 3. 2 M urea, Lane 4. 1 M urea, Lane 5. 0.5 M urea, Lane 6. 0 M urea. **B.** SARS N protein cleavage as a function of time. Samples of the BL21(ΔE3)-produced N protein were obtained following dialysis into 0 M urea and storage at 4°C for the indicated lengths of time. Lanes M. Molecular weight markers, Lane 1. t=0, Lane 2. After 18 hours, Lane 3. After 120 hours.

A.


66kDa
43kDa
29kDa
20kDa
SARS N protein
A band
B band
M
1
2
3
4
5
6

B.


66kDa
43kD
29kDa
20kDa
15kDa
SARS N protein
A band
B band
LMW band
M
1
2
3

underwent cleavage with time (Figure 28B, lane 1 vs Lane 2). Lane 3 shows that the SARS N protein underwent additional cleavage into smaller molecular weight fragments (LMW band) with prolonged storage. This result implies that additional cleavage occurs as a function of time.

7.2. Mass spectrometric characterization of protein fragments

Mass spectrometry was used to identify the A and B bands by determining their molecular weights. The result shown in Figure 28 indicated that cleavage of the tag by DAPase was not responsible for the detected proteolysis, but to avoid any addition of extraneous protease, the recombinant N protein was analysed without $(His)_6$-tag cleavage. Therefore, all samples of N protein analysed by mass spectrometry contain an additional 11 amino acid, N-terminal MKHHHHHHHMQ- sequence (resulting in an a.a. sequence of 1-433 versus 1-422 in the untagged protein). Unless otherwise noted, the N protein sequences and the numbering in the following sections reflect the addition.

7.2.1. SELDI-TOF/MS of SARS N protein and A and B bands

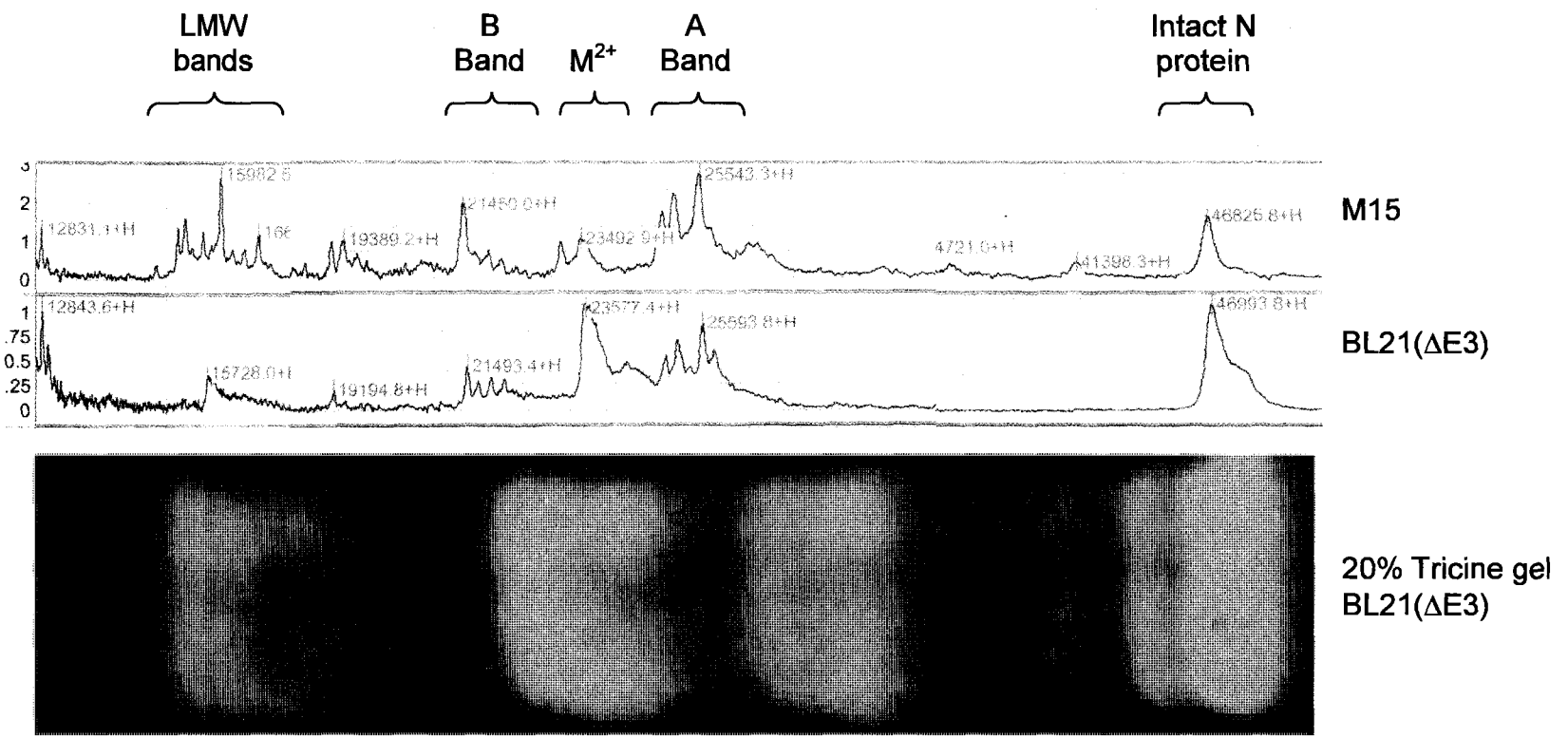
Surface enhanced laser desorption ionization (SELDI) mass spectrometry was used to measure the molecular weights of all proteins in the SARS N protein solution. SARS N protein preparations from both M15 and BL21(ΔE3) cells were prepared for SELDI-analysis and the molecular weights were determined as described in Methods (section 6.1.5). While the A and B bands were resolved on SDS-PAGE as two discrete bands at approximately 29 and 25 kDa, SELDI-TOF mass spectrometry showed that the bands were composed of several proteins possessing similar molecular weights (Figure 29). Analysis indicated that these A and B bands contained several proteins within the molecular weight ranges: 24.0-

26.0 (band A) and 21.5-23.0 kDa (band B). Each set of peaks in a range consisted of four or five peaks and each peak was separated from its nearest neighbour by ~200-250 Da (approximately two amino acids). The SELDI-determined molecular weights for the A and B bands were summed. The results showed: 26 + 21.5 = 47.5 kDa and 24 + 23 = 47 kDa. The molecular weights were consistent with the A and B bands being the hydrolysed fragments of the full-length SARS N protein (47 kDa). When the masses between the M15 and BL21(ΔE3)-produced samples were compared, the data indicated that both samples possessed the same multiple peaks in the A and B band molecular weights ranges. This result was consistent with a specific cleavage of the intact SARS N protein into the same molecular fragments.

7.2.2. MALDI-TOF/MS analysis of in-gel tryptic digests

Matrix assisted laser desorption ionization, time of flight mass spectrometry (MALDI-TOF/MS) analysis was performed on in-gel tryptic digests of the proteins to confirm the identities of the A and B bands. The three major bands that were identified by SDS-PAGE (the recombinant N protein, A band and B band) were excised from a gel and digested as described in Methods (section 6.1.6). Peptide masses were determined using MALDI-TOF/MS. The MASCOT peptide database was used to identify the digested proteins by their tryptic peptide fingerprints. The MASCOT database indicated that the 47 kDa protein possessed peptides consistent with the SARS N protein amino acid sequence. When the A and B bands were assessed, both were also identified as the SARS N

Figure 29. Molecular weight analysis of SARS N protein and A and B bands by SELDI-TOF/MS. Samples of the SARS N protein solution following refolding were used for molecular weight analysis by SELDI-TOF/MS. The SELDI-TOF/MS determined molecular masses are shown on the x-axis. The y-axis indicates the relative ion intensity. The top panel shows the refolded SARS N protein solution produced by M15 cells. The middle panel shows the refolded SARS N protein solution produced by BL21(ΔE3) cells. These data were compared to the bands identified on a 20% Tricine gel (bottom panel, taken from Figure 28B). The regions that were identified included the **LMW bands** (small molecular weight proteins observed after extended storage), **B band, M^{2+}** (Doubly charged intact N protein), **A Band,** and **Intact N protein.**


LMW bands
B Band
M2+
A Band
Intact N protein
M15
BL21(ΔE3)
20% Tricine gel
BL21(ΔE3)

protein. To identify the region of the SARS N protein that the A and B bands corresponded to, a manual search of the molecular masses was performed by comparing the MALDI-TOF/MS data (Figure 30) to the expected molecular weights of tryptic peptides determined by an *in silico* digest of the SARS N protein amino acid sequence. The results, summarized in Table IX show the experimentally observed molecular weight (MW), the predicted molecular weight, mass difference between the observed and predicted masses (Δ MW), and peptide sequence that the data corresponded to. The results in Figure 30 and Table IX show that all the MALDI-identified peptides in the A band corresponded to the C-terminal region of the SARS N protein, while all MALDI-identified peptides in the B band corresponded to the N-terminal region. The small difference between the observed and predicted values (ΔMW) indicates accurate peptide identification.

The result identified the A band as the C-terminal end and the B bands as the N-terminal end of the SARS protein (Figure 30). The result of SELDI-TOF/MS analysis showed the A and B band possessed nearly equal molecular weights (Figure 29). Combined, the results indicated that the recombinant N protein underwent cleavage near the center of its amino acid sequence during refolding, which resulted in two extra bands on SDS-PAGE.

Figure 30. MALDI-TOF/MS analysis of tryptic peptides. MALDI-TOF/MS analysis was performed on in-gel tryptic digests. The SARS N protein band, the A band and the B band were excised from SDS-PAGE and prepared as detailed in Methods (section 6.1.6). MALDI-TOF/MS was performed and the detected peptide masses were compared to a predicted SARS N protein digest. The masses corresponding to the C-terminal end of the protein are highlighted in pink. Masses from N-terminal end of the protein are highlighted in green. **A.** Molecular weights peptides from 1000 – 1500 Da **B.** Molecular weight peptides from 1400 – 2200 Da. The summarized data can be found in Table IX.

A.

B.

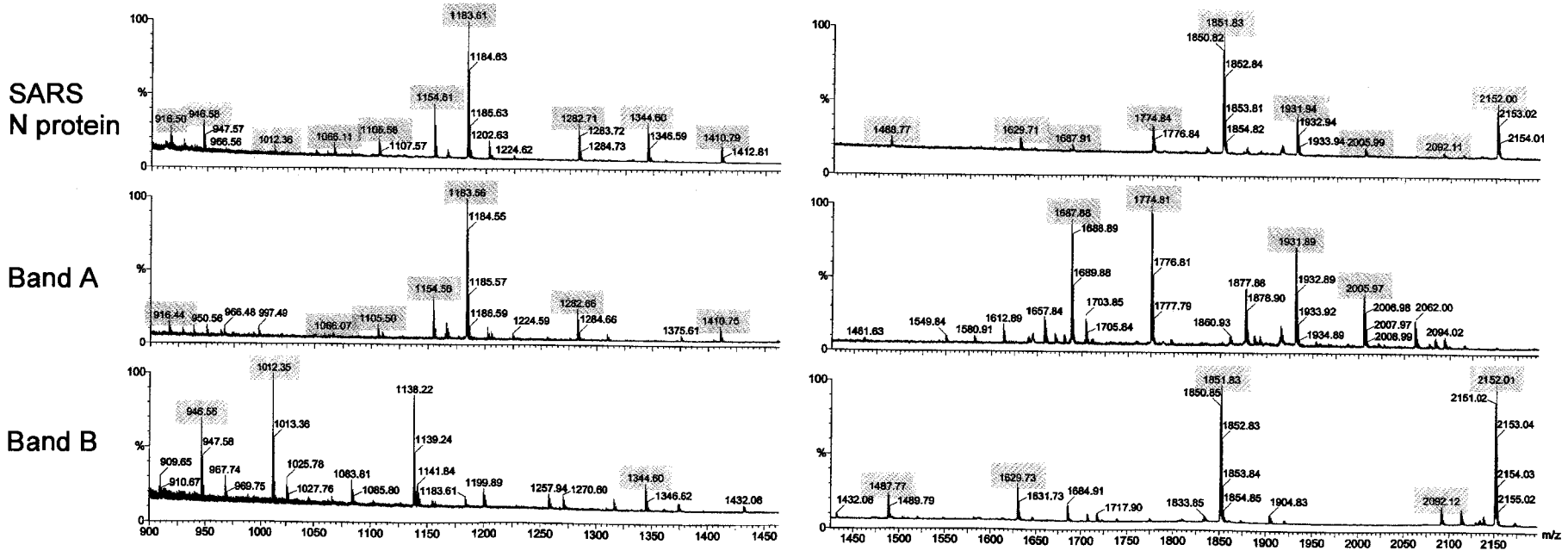
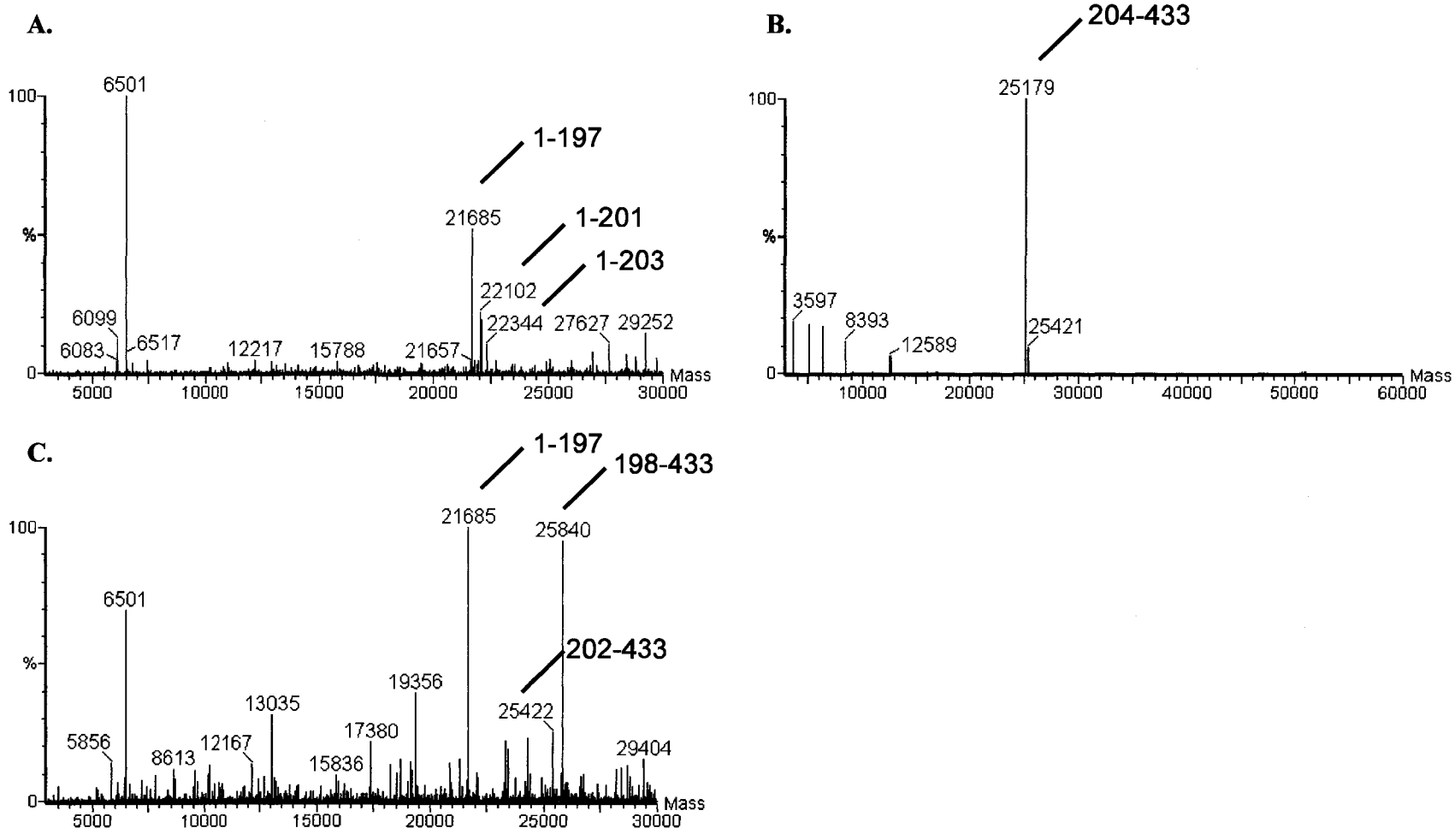


Table IX. MALDI-TOF/MS detected SARS N protein fragments				
	Observed MW	Predicted MW	Δ MW	$(His)_6$-tagged SARS N protein sequence
A Band	916.4	916.5	0.1	374-381
	1105.5	1105.6	0.1	351-359
	1154.6	1154.6	0.0	388-397
	1183.6	1183.6	0.0	279-288
	1282.7	1282.7	0.0	387-397
	1410.8	1410.7	0.1	388-399
	1687.9	1687.9	0.0	222-238
	1774.8	1774.8	0.0	290-305
B Band	946.6	946.5	0.1	74-80
	1850.9	1850.8	0.0	27-44
	2092.1	2091.1	1.00	162-181
	2151.0	2151.0	0.0	81-100

7.2.3. LC/MS analysis of protein molecular weight

To determine the exact location of SARS N protein cleavage, tandem LC/MS mass spectrometric analysis was performed. With the identities of the A and B bands known, LC/MS mass data could be compared against the theoretical masses of peptides from the SARS N protein amino acid sequence. LC/MS was performed on a BL21(ΔE3)-produced sample as described in Methods (section 6.1.7).The results shown in Figure 31 were matched against theoretical peptides mass from the $(His)_6$-tagged SARS N protein sequence. Table X shows the summarized data, which indicated that the unknown A and B proteins in solution corresponded to the SARS N protein, if cleavage occurred in the 197-204 amino acid range (this is the same region as the 186-193 sequence in the SARS protein when the $(His)_6$-tag is excluded). The results of cleavage in this region were consistent with SELDI/TOF-MS

Figure 31. LC/MS analysis of SARS N protein solution. LC/MS was performed on a SARS N protein obtained from BL21(ΔE3) cells to identify the molecular weights of the proteins in solution. Protein mixtures were separated by in-line capillary reverse phase HPLC prior to mass spectrometry as detailed in Methods (section 6.1.7). The three major elution peaks are shown as **A, B** and **C** panels. Each elution contained a mixture of proteins. The masses that could be identified as SARS N protein fragments are labelled with the corresponding amino acid sequences. The intact $(His)_6$-tagged SARS N protein contained amino acids 1-433. The summarized data can be found in Table X.


A.
100
%
0
6501
6099
6083
6517
12217
15788
21657
21685
22102
22344
27627
29252
1-197
1-201
1-203
Mass
5000
10000
15000
20000
25000
30000
B.
100
%
0
25179
3597
8393
12589
25421
204-433
Mass
10000
20000
30000
40000
50000
60000
C.
100
%
0
6501
5856
8613
12167
13035
15836
17380
19356
21685
25422
25840
29404
1-197
198-433
202-433
Mass
5000
10000
15000
20000
25000
30000

(Figure 29) and MALDI-TOF (Figure 30) determined results for the A and B bands molecular weights.

Close inspection of the LC/MS data showed that the cleavages in the 197-204 region occurred between residues 197/198, 201/202 and 203/204. These data correspond to cleavage after -SR- motifs at the SR-rich region of the N protein sequence. These data were consistent with the result from SELDI-TOF/MS that indicated the successive cleavage between products possessed a gap of approximately two amino acids (Figure 29). Therefore, the SELDI-observed series of peaks (or ragged-end cleavages) were due to cleavage at this repeating consensus sequence.

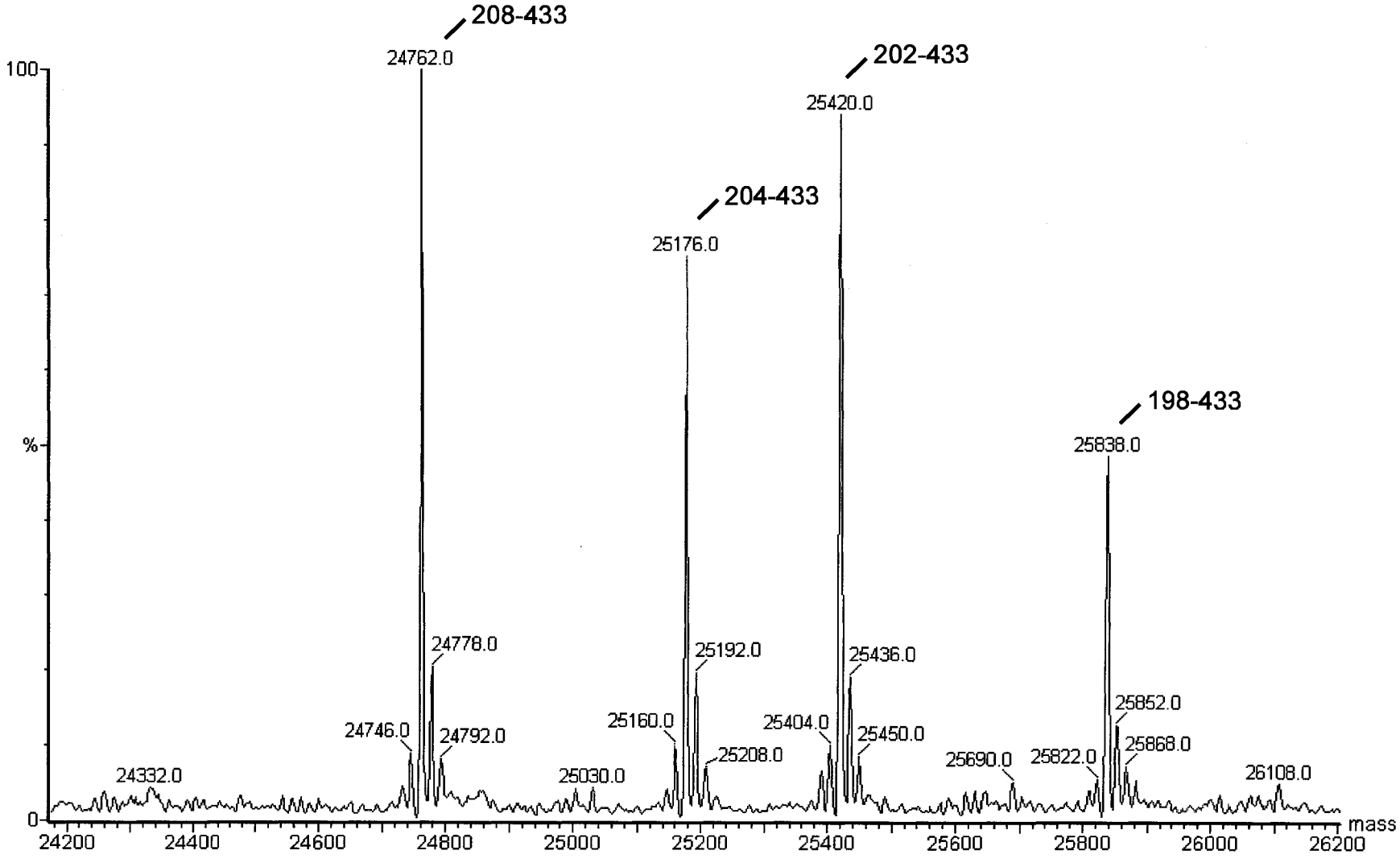
Table X. LC/MS detected SARS N protein fragments

Observed MW	Predicted MW	Δ MW	$(His)_6$-tagged SARS N protein sequence
21685	21684	1	1-197
22102	22101	1	1-201
22344	22345	1	1-203
25179	25179	0	204-433
25422	25423	1	202-433
25840	25840	0	198-433

7.2.4. HPLC, LC/MS

To improve the detection of SARS N protein fragments by LC/MS, the mass spectrometry experiment was repeated on a protein sample that had been purified by a preparatory reverse phase chromatography (rpHPLC) step. The step was used to eliminate buffer salts and improve protein ionization during a subsequent MS analysis. SARS N protein was purified as described in Methods (section 6.1.8). The eluted samples were analysed by LC/MS as in the preceding section. The result shown in Figure 32 shows that the rpHPLC step improved the sensitivity of the mass spectrometric detection and reduced the

Figure 32. Analysis of reverse phase purified, SARS N protein fragments by LC/MS. Protein analysis using reverse phase HPLC as a preparatory purification step prior to LC/MS analysis. A SARS N protein sample was separated using a Vydac C8 column and a 10-90% acetonitrile gradient (section 6.1.8). Eluted proteins were collected and assessed by LC/MS as described in Methods (section 6.1.7). Masses were matched to amino acid sequences from the $(His)_6$-tagged SARS N protein and are shown by the figure labels. The tabulated data can be found in Table XI.


208-433
24762.0
202-433
25420.0
204-433
25176.0
198-433
25838.0
100
%
0
24778.0
25192.0
25436.0
24746.0
24792.0
25160.0
25404.0
25450.0
25208.0
25852.0
25868.0
24332.0
25030.0
25690.0
25822.0
26108.0
mass
24200
24400
24600
24800
25000
25200
25400
25600
25800
26000
26200

background noise. The improved LC/MS results were analysed to identify N protein fragments, which identified four C-terminal, SARS N proteins. The molecular weights matched the SARS N protein sequence following cleavages in the SR-rich 195-207 amino acid region. The data showed that cleavage was specific for the SR motif within the specified region illustrated in Figure 33.

This set of LC/MS data was similar to the previous data (Figure 31 and Table IX) and identified the same C-terminal proteins including additional one at (a.a. 208-433). Unfortunately, the corresponding N-terminal proteins could not be detected. This may indicate that the reverse phase step did not use a suitable acetonitrile gradient to elute these proteins and they either were eluted in the column flow though or were totally retained. This will be a direction for future studies. However, the current result shows that all C-terminal fragments, which were expected to elute simultaneously due to the shared amino acid sequence, could be identified.

Table XI. HPLC, LC/MS detected SARS N protein fragments

Observed molecular weight	Predicted MW	Δ MW	$(His)_6$-tagged SARS N protein sequence
24762	24764	2	208-433
25176	25178	2	204-433
25420	25421	1	202-433
25838	25839	1	198-433

7.3. SARS N protein protease assays

To test for the presence of non-specific proteases, two fluorescence-based proteolytic assays were used: Commercial Quanticleave ™ assay using FITC-labelled casein for detection of non-specific proteolysis and EDANS/DABCYL conjugated SARS peptides corresponding to the N protein cleavage region.

Figure 33 SARS N protein cleavage sites. The mass spectrometry determined cleavage sites are shown in red. Cleavage occurred following the SR motifs labelled in red at sequence position 195-207 (-^{184}SSRSSSRSRGNSR^{196}- in the untagged SARS amino acid sequence). A $(His)_6$-tag and TAGZyme™ cleavage site is shown in bold. SR motifs that showed no evidence of cleavage are labelled in green.

~21KDa
a.a. 1-197

MKHHHHHHHMQMSDNGPQSNQRSAPRITFGGPTDSTDNNQNGGRNGAR
PKQRRPQGLPNNTASWFTALTQHGKEELRFPRGQGVPINTNSGPDDQI
GYYRRATRRVRGGDGKMKELSPRWYFYYLGTGPEASLPYGANKEGIVW
VATEGALNTPKDHIGTRNPNNNAATVLQLPQGTTLPKGFYAEGSRGGS
QASSR/

SSSR/ SR/ GNSR/

~25KDa
a.a. 208-433

NSTPGSSRGNSPARMASGGGETALALLLLDRLNQLESKVSGKGQQQQG
QTVTKKSAAEASKKPRQKRTATKQYNVTQAFGRRGPEQTQGNFGDQDL
IRQGTDYKHWPQIAQFAPSASAFFGMSRIGMEVTPSGTWLTYHGAIKL
DDKDPQFKDNVILLNKHIDAYKTFPPTEPKKDKKKKTDEAQPLPQRQK
KQPTVTLLPAADMDDFSRQLQNSMSGASADSTQA

7.3.1. Protease detection using the Quanticleave ™ assay

FITC-labelled casein was used to detect protease activity. The assay was prepared according the manufacturer's instructions using samples of the SARS N protein. The result from the experiment showed no increase in fluorescence (data not shown).

The Quanticleave assay is based on FITC-labelled casein as described by Twinning (*131*). Cleavage of the substrate leads to an increase in fluorescence. Detection of ~1 ng trypsin is possible, but the linear range for protease detection by this method is between 5-100 ng trypsin in a 50 µl reaction volume. The low sensitivity means that it is unlikely that a non-specific protease would not be detected. This indicated the absence of general proteases or that the contaminating protease was not capable of cleaving casein.

7.3.2. Protease detection using EDANS/DABCYL-conjugated SARS N protein peptides

The choice of casein as the substrate for the assay did pose one problem. Casein lacks an -SR- motif, which might have resulted in the observed negative result. To determine if the absence of FITC-labelled substrate cleavage was due to the lack of an -SR- motif in the casein sequence, a specific proteolysis assay was developed using the -SR- rich region that corresponded to the SARS N protein cleavage site. Peptides were synthesised according to the data in Figure 33. One peptide contained the entire SR-rich cleavage site (^{181}SQASSRSSSRSRGNSRNSTP^{200}) and one peptide contained the amino acid sequence immediately preceding it (^{168}LPKGFYAEGSRGGSQASS^{185}), which contained an SR-motif that was not identified as being cleaved under mass spectrometric analysis. Both peptides were conjugated with EDANS and DABCYL for fluorescent resonance energy transfer (FRET) detection (assay characteristics are shown in Appendix A). The FRET-labelled peptide substrates were incubated with recombinant SARS N protein, trypsin, or buffer

controls as detailed in Methods. Figure 34A (EDANS/DABCYL-conjugated, ^{181}SQASSRSSSRSRGNSRNSTP^{200}) and Figure 34B (EDANS/DABCYL-conjugated, ^{168}LPKGFYAEGSRGGSQASS^{185}) showed that positive controls (trypsin) showed an increase in fluorescence, but there was no increase in fluorescence in the presence of SARS N protein over a 60 minute incubation. This result indicated that no cleavage of the substrates occurred in the SARS N protein tests.

The results of the FITC-casein assay showed that the SARS N protein cleavage was not the result of a general contaminating *E. coli* protease. The EDAN/DABCYL-conjugated peptide results showed that proteolysis was not due to an SR-specific protease that could cleave the peptide substrates. Instead the result indicated that cleavage was specific for the intact SARS N protein.

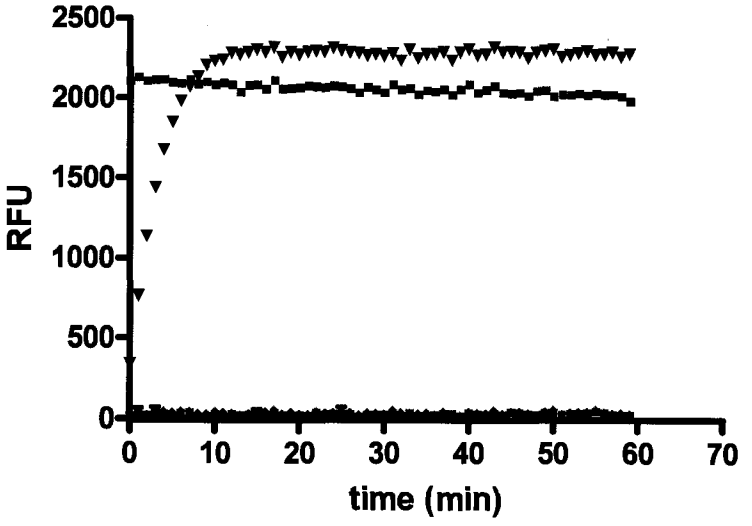
7.3.3. Protease detection using ovalbumin or RNAse A mixing

The lack of protease activity towards casein and the FRET-labelled -SR- peptide substrates led to a final experiment. To test for SARS N protein specific proteolysis, the SARS N protein was refolded in the presence of either ovalbumin or RNAse A. Ovalbumin possesses an -SR- at a.a. 104-105 and cleavage would result in 12 and 32 kDa fragments. RNAse A possesses an -SR- at a.a. 36-37 and cleavage would result in 9 and 8 kDa fragments. The experiment was performed as described in Methods (section 6.1.11).

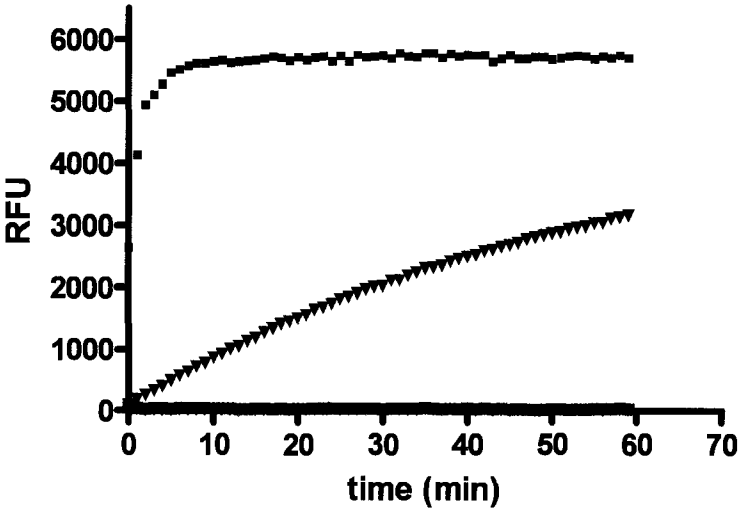
Upon the completion of the refolding protocol, samples of the protein solutions were separated on a 20% Tricine gel. The result in Figure 35 showed that while the SARS N protein still exhibited a distinctive cleavage pattern upon refolding (lane 1), both ovalbumin (lane 4) and RNAse A (lane 7) were unaffected. Samples of the protein mixtures, stored at 4°C, were obtained at various times following the protein refolding. These samples were also separated on Tricine gel and the result showed that A and B band intensities increased with

Figure 34. EDANS/DABCYL-conjugated SARS N protein peptide assay. EDANS/DABCYL-conjugated peptides were generated as artificial protease substrates. Assays were performed as described in Methods (section 6.1.10). Increased fluorescence (measured in relative florescence units, RFU) indicated proteolysis. 10 μg and 100 ng trypsin were used as a positive controls along with SARS N protein samples from either M15 or BL21(ΔE3) cells. SARS N protein was tested at a concentration of 2.5 mg/mL. Each SARS N protein sample is shown as an average of three trials. **A.** EDANS/DABCYL-conjugated - ^{181}SQASSRSSSRSRGNSRNSTP^{200}- peptide. **B.** EDANS/DABCY-conjugated - ^{168}LPKGFYAEGSRGGSQASS^{185}- peptide.

A.

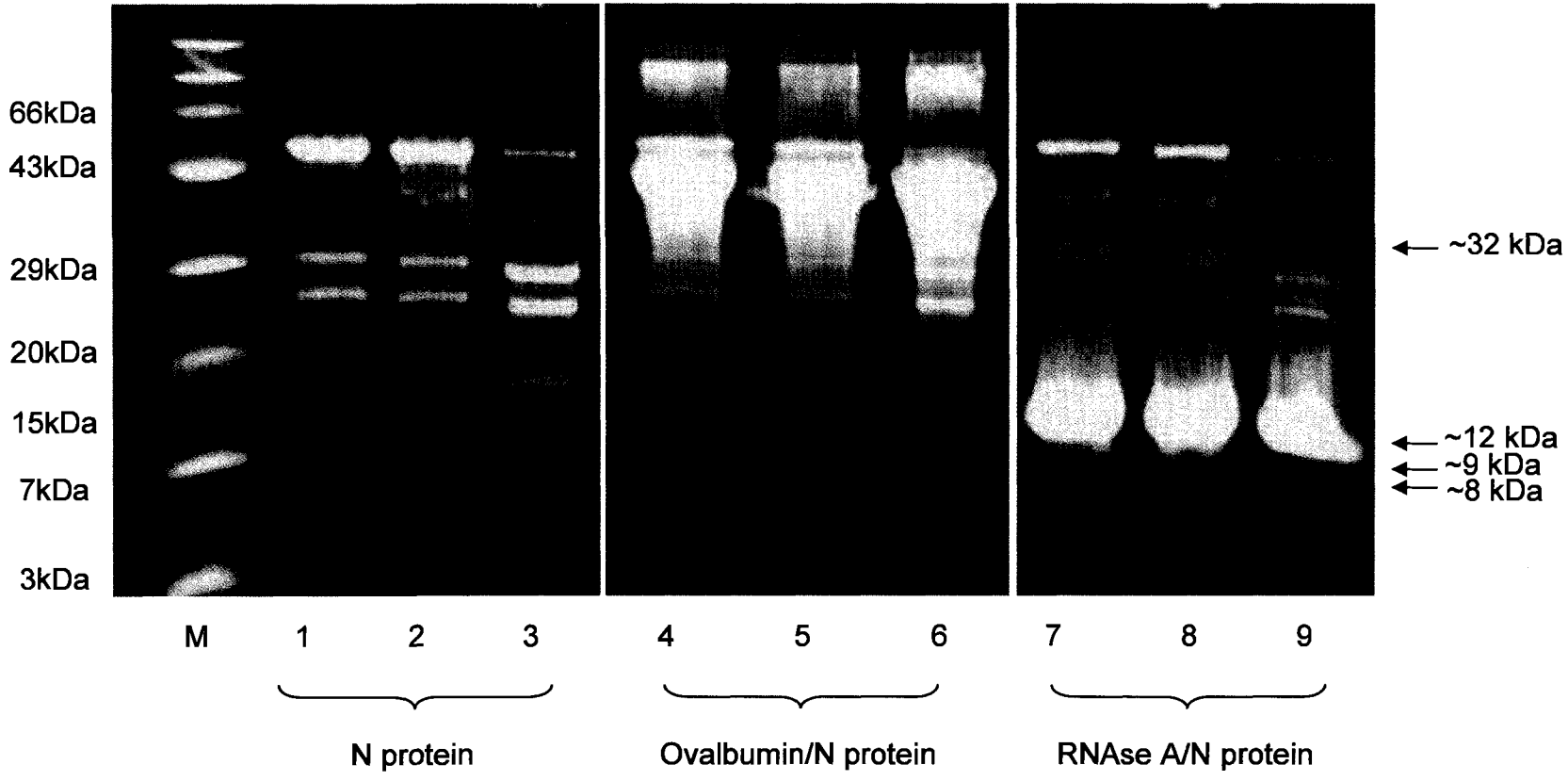


B.



▪ 10 μg trypsin
▾ 100 ng trypsin
▪ SNP (M15)
◆ SNP (BL21(ΔE3))

Figure 35. SARS N protein mixed with ovalbumin or RNAse A. Mixing experiments were performed as described in Methods (section 6.1.11). Aliquots of the reaction mixture were obtained at the indicated times were separated on a 20% Tricine gel. Staining was by Sypro Ruby Red stain. Lanes 1, 4, and 7 show protein samples obtained after overnight dialysis into a final refolding buffer (0 M urea). Lanes 2, 5, and 8 show samples obtained following further 2-hour incubation at 4°C. Lanes 3, 6, and 9 show a samples obtained following a 120-hour incubation at 4°C. Lanes 1-3 contain N protein refolded in the absence of secondary protein. Lanes 4-6 contain ovalbumin mixed with SARS N protein, Lanes 7-9 contain RNAse A mixed with the SARS N protein. Lane M shows molecular weight markers. Arrows indicate the predicted molecular weights of predicted ovalbumin (32 and 12 kDa) and RNAse A (9 and 8 kDa) cleavage products, following cleavage at an SR-motif.


66kDa
43kDa
29kDa
20kDa
15kDa
7kDa
3kDa
M
1
2
3
4
5
6
7
8
9
~32 kDa
~12 kDa
~9 kDa
~8 kDa
N protein
Ovalbumin/N protein
RNAse A/N protein

time, which indicated that the SARS N protein was being cleaved as a function of time. The additional proteolysis of the SARS N protein (Figure 35, lanes 2, 3) occurred in the presence of ovalbumin (Figure 35, lanes 4, 6) or RNAse A (Figure 35, lanes 7, 9). There was no detectable appearance of the predicted ovalbumin or RNAse A cleavage products. This result and the negative detection of proteolysis in the FITC-casein (section 7.3.1) and EDANS/DABCY-conjugated peptide assays (section 7.3.2), confirmed that the observed proteolysis was specific for the intact SARS N protein.

8.0 Discussion:

When SARS N protein cleavage was detected upon protein refolding, it was originally thought to be due to non-specific proteolysis. This cleavage has now been shown to be specific for SR-motifs and it appears to occur in the absence of protease. Therefore, it appears that the recombinant SARS N protein possesses an ability to undergo autocatalysis into a specific series of smaller molecular weight proteins.

The possibility that a contaminating bacterial protease was responsible for the proteolysis was excluded, the first step being *pQE2:snp* plasmid transformation into an *E. coli* cell line, which had been previously used to produce intact MKP3 proteins (Part 1 of the thesis). The BL21(ΔE3)-produced SARS N protein was capable of fragmenting into the same pattern of lower molecular weight A and B bands as in the original M15 cell line (SDS-PAGE in Figures 27 and 28). The result of the SDS-PAGE analysis was subsequently confirmed using the more accurate SELDI-TOF/MS method, which showed that the bands indicated on SDS-PAGE were identical in both M15 and BL21(ΔE3) productions of the N protein (Figure 29). The data implied that that either the cleavage was due to a shared protease in both M15 and BL21(ΔE3) cells, or that it was due to a property of the SARS N protein.

When N protein samples were analysed by an in-gel digest that was paired with MALDI-TOF/MS mass analysis, the data showed that the A band was composed of the C-terminal portion of the SARS N protein, while the B band was identified as the N terminal portion. Knowing the protein identities allowed LC/MS to be used to identify the cleavage motif. Molecular weights of all proteins in solution were determined by LC/MS by matching the detected masses to SARS N protein amino acid sequences that possess the same

molecular weights. This method ultimately identified cleavage sites in the 195-207 a.a. region that corresponded to SR-motifs (shown in Figure 33). While other potential cleavage sites were observed to exist, particularly if the SARS N protein was stored for an extended period, these protein fragments could not be identified by mass spectrometry. It may be that the fragments resulted from additional cleavages after arginine; the subsequent loss of a number of arginines can impair MS detection due to the reduced number of ionisable residues in the amino acid sequence. This may have resulted in the failure to detect the fragments using LC/MS.

Once the major cleavage products (leading to the A and B bands) were identified and the cleavage motif determined, a database of proteases was searched to find an enzyme that could result in the observed cleavages. While the search identified several proteases that can cleave following positively charged residues, these proteases will cleave either after any arginine (Arg-C and clostripain) or any arginine or lysine (trypsin). There are no known proteases showing specificity for an SR-motif. This finding made protease contamination more unlikely.

As a further test for proteases, two fluorescent cleavage experiments were performed using FITC-labelled casein or EDANS/DABCYL-labelled SARS N protein cleavage site. The results of the experiments were both negative. While casein does not possess an -SR- site, it does possess arginine and lysine residues that can be cleaved. The result of this test argued against the presence of a general protease that can cleave following any arginine or lysine residue. Next, to test for specific SR-motif cleavage, a peptide was produced that containing the entire SR-cleavage region (^{181}SQASSRSSSRSRGNSRNSTP^{200}). To test if cleavage might occur at alternate SR-motifs, a second peptide was produced from the region preceding SARS, SR-cleavage region. This peptide contained one SR-motif that showed no

evidence of cleavage (^{168}LPKGFYAEGSRGGSQASS^{185}-) by mass spectrometry. Neither substrate showed cleavage when tested with the SARS N protein.

Characterization of the EDANS/DABCYL-peptide assays using trypsin indicated that both assays were capable of detecting ~$2.5x10^{-6}$ mg/mL trypsin (500 pg in a 200 µl assay) following a 1 hour incubation at 25°C (sensitivity could be increased using a longer incubation time or higher incubation temperature). The assay was even more sensitive than the commercial FITC-casein-based Quanticleave™ assay (capable of detecting ~1 ng trypsin/50 µl assay; (*131*)). Therefore, it is unlikely that a protease escaped the limit of detection. The result showed that non-specific proteases were absent, but it also indicated that a SARS N protein, SR-region-specific protease was unable to cleave the artificial substrate. A possible explanation may be that the binding and consensus sequences on an autocatalytic SARS N protein are located on different regions. While the SR-peptides contained potential cleavage sites, if they did not possess the binding site, they would not be cleaved.

Finally, simultaneous refolding of the SARS N protein in the presence of ovalbumin and RNAse A was tested. Both albumin and RNAse A possess a single SR-motif, but neither showed any cleavage. This was in spite of being present at higher concentrations than the SARS N protein (10:1 or 30:1 molar excesses). The greater concentration should have led to enhanced proteolysis of the more plentiful substrate if a protease was present. However, the results in Figures 35 showed that neither ovalbumin nor RNAse exhibited proteolysis, but the SARS N protein was cleaved. This result was further evidence that proteolysis was specific to the intact SARS N protein.

The detection of SARS N protein proteolysis detailed in the current thesis is not the first observation of N protein cleavage. As discussed in the Introduction (section 5.1.1) a similar pattern of N protein cleavage was detected in a study using SARS coronavirus-infected Vero (monkey) cells (*127*). The data from that study showed that SARS N proteins were cleaved into a series of bands with four bands in the ~40-46 kDa range, three at ~27-31 kDa and four bands at ~16-23 kDa (all values estimated from SDS-PAGE and molecular weight markers). The authors also showed a second result using purified SARS N protein that was treated with caspase 3. The experiment resulted in the identification of three bands in the 27-31 kDa range and three bands in the 16-23 kDa range. In the current study, SDS-PAGE results from *E. coli* derived SARS N protein showed similar N protein bands at 29 and 25 kDa. This cleavage occurred in the absence of caspase 3 (which *E. coli* does not produce). While there are differences in the absolute molecular weights between the published results and the experimentally determined weights detailed in the current thesis, the mass differences are within the margin of error for estimation from SDS-PAGE due to the use of different molecular weight standards.

In the previous reporting of SARS N protein cleavage (*127*), MALDI and LC/MS data was not able to identify peptides from the SR-rich region, the site of cleavage which was identified using the same methods employed in this study. This may be significant. The authors of the study used peptide analysis to identify proteins, much like the MALDI experiments performed in Results section 7.2.2. However, in that study, the authors used two enzymes for their in-gel digest, trypsin and Glu-C (*127*). While, peptides from the SR-rich region were identified following an in-gel trypsin digest, the subsequent experiment using a Glu-C digest did not identify any of the SR-region peptides. This may be due to the different specificities of the two enzymes. Trypsin cleaves at positively charged residues including

arginine. Therefore, peptides with a pre-existing cleavage following an SR motif would still be identified on MALDI analysis. Since there is no net change to the peptide fingerprint whether trypsin digestion or a combined trypsin digestion with SARS N protein autocatalysis occurs, the cleaved SR-region peptides are detected and identified. However, peptide identification following Glu-C digestion uses a database of peptides that are predicted to be cleaved only after glutamyl or aspartyl (E or D) residues. If cleavage at the SR region had occurred prior to analysis, then these additional, unexpected cleavages would result in the peptides being truncated and therefore misidentified. The data published by Ying's group indicates that the sequence between amino acid positions 174-217 (186-229 in my recombinant SARS N protein) was not identified after Glu-C cleavage (*127*). While this is not conclusive, it can be speculated that the cleavages that were observed in the previous study and my work occurred at the same cleavage sites.

Unfortunately, the group that identified the SARS N protein cleavage did not attempt to identify the site of their observed cleavage or the general region of the SARS N protein to which each observed band corresponded. Instead, the authors attributed the resulting bands to cleavage by cellular caspase following a cursory test of the SARS N protein that was incubated with caspase 6 and 3 (*127*). While caspase could theoretically digest viral nucleocapsid proteins, the SARS N protein sequence lacks the canonical -DXXD- caspase 3 cleavage motif. In light of the experiments contained in this thesis, it is unlikely that the SARS N protein is being cleaved by caspase.

The SARS N protein and specifically its SR-rich region appears to be readily cleaved in both *E. coli* and Vero cells and attempts to identify a contaminating protease responsible for that cleavage, though not exhaustive, were unsuccessful. Guruprasad and co-workers (*132*) report that certain dipeptide sequences are statistically more likely to be found in

proteins with shorter *in vivo* half-lives than in those with longer in vivo half-lives. Both SR and RS (arginine-serine) dipeptides (and the N-protein SR-rich region contains both) are among those found more frequently in "unstable" proteins. In vivo instability is usually linked to digestion by circulating or cellular proteases. Therefore, a database of proteases was searched to find an enzyme or class of enzymes capable of the observed cleavages. Not only is the SR motif cleavage site is not a recognition site for any known *E. coli* proteases (consistent with our inability to detect evidence of protease contamination in our N-protein preparations), it is not recognized by any other protease in the database (*86*). Several proteases do cleave following positively charged residues, including arginine (Arg-C and clostripain) or any arginine or lysine (trypsin); however, this cleavage is otherwise non-specific and does not require S in the P2 position.

The putative biological role(s) of this SR-rich region is not completely known, butt a BLAST similarity search performed on the SR-rich region indicated that the amino acid region was not only found in corona virus obtained from infected humans, but also civet cats, and bats. Interestingly, SR-region analogues were also detected in the early intermediate protein 2 from Human herpesvirus 6. In the SARS N protein, it has been proposed that the SR-region allows the SARS N-protein to form the dimers or tetramers (*114;133*), which are thought to be essential in formation of the mature viral nucleocapsid and aiding in compacting the RNA/nucleocapsid complex (*133-136*). Deletion of the SR-rich region not only completely abolished the oligimerization of the N-protein but also resulted in a disordered distribution of N-proteins in mammalian cells (*114*). Considering that the region responsible for self-association identified by these previous studies was identical to the cleavage region identified in this study, it is reasonable to speculate that the cleavage of this sequence may aid in the unpacking of viral RNA, necessary to allow this RNA to serve as a

template for viral genome replication. And if binding and subsequent lysis of the SR-rich amino acid region are inherent properties of the SARS N-protein, the SR-rich amino acid sequence presents a potential target for a SARS therapeutic. If cleavage of this SR-rich region can be shown to be essential to allow the unpackaging and exposure of viral RNA for replication, an agent that prevents its cleavage (perhaps a competitive inhibitor such as a mimetic of the SR-region) may prove effective in interrupting viral replication in infected individuals.

9.0. General conclusion

The unifying theme of this thesis has been biochemical analysis. *In vitro* biochemical techniques were used to analyse the function of two proteins, showing that:

1) An interdomain binding site in MKP3 was capable of shielding the catalytic site of the protein, leading to a highly specific catalysis of ERK, but inappropriate expression of the protein was found to lead to deregulation in the binding function and shift the interdomain binding to interprotein oligomerization. This effect led to the deregulation of protein function.

2) In the SARS N protein, biochemical analysis determined that a region of the nucleocapsid exhibited proteolysis and characterization of the cleavage suggested autocatalytic activity.

What these finding show, apart from the very relevant conclusions in the respective discussion sections, is that *in vitro* biochemical analysis plays an important role in modern biochemical analysis. It complements well with both older *in vivo* and newer *in silico* analysis methods. The main advantage of *in vitro* work is the ability to isolate elements in a system in order to study a particular phenomenon. This approach allows one to tease apart the layers of function that are observed in *in vivo* studies (as in deciphering the role of MKP3 oligomers on pancreatic adenocarcinoma) as well as prove or give relevance to theoretical *in silico* data (shown by SARS N protein peptide identifications using theoretical and observed protease digestion). These finding demonstrate the advantages that a holistic approach to biochemistry, combining old and newer methods to analyse proteins, possesses.

10. References

1. Richard W.Hill and Gordon A.Wyse (1989) Muscle and other effectors, in *Animal Physiology* 2 ed., pp 507-548, Harper Collins, New York.

2. Kozasa, T. (2004) The structure of GRK2-G beta gamma complex: intimate association of G-protein signaling modules, *Trends Pharmacol. Sci. 25*, 61-63.

3. Metaye, T., Gibelin, H., Perdrisot, R., and Kraimps, J. L. (2005) Pathophysiological roles of G-protein-coupled receptor kinases, *Cell Signal. 17*, 917-928.

4. Saez, J. C., Retamal, M. A., Basilio, D., Bukauskas, F. F., and Bennett, M. V. (2005) Connexin-based gap junction hemichannels: gating mechanisms, *Biochim. Biophys. Acta 1711*, 215-224.

5. Solan, J. L. and Lampe, P. D. (2005) Connexin phosphorylation as a regulatory event linked to gap junction channel assembly, *Biochim. Biophys. Acta 1711*, 154-163.

6. Cuevas, B. D., Abell, A. N., and Johnson, G. L. (2007) Role of mitogen-activated protein kinase kinase kinases in signal integration, *Oncogene 26*, 3159-3171.

7. Goldsmith, Z. G. and Dhanasekaran, D. N. (2007) G protein regulation of MAPK networks, *Oncogene 26*, 3122-3142.

8. Roberts, P. J. and Der, C. J. (2007) Targeting the Raf-MEK-ERK mitogen-activated protein kinase cascade for the treatment of cancer, *Oncogene 26*, 3291-3310.

9. Jeffrey, K. L., Camps, M., Rommel, C., and Mackay, C. R. (2007) Targeting dual-specificity phosphatases: manipulating MAP kinase signalling and immune responses, *Nat. Rev. Drug Discov. 6*, 391-403.

10. Owens, D. M. and Keyse, S. M. (2007) Differential regulation of MAP kinase signalling by dual-specificity protein phosphatases, *Oncogene 26*, 3203-3213.

11. Hafen, E. (1998) Kinases and phosphatases--a marriage is consummated, *Science 280*, 1212-1213.

12. Camps, M., Nichols, A., Gillieron, C., Antonsson, B., Muda, M., Chabert, C., Boschert, U., and Arkinstall, S. (1998) Catalytic activation of the phosphatase MKP-3 by ERK2 mitogen-activated protein kinase, *Science 280*, 1262-1265.

13. Groom, L. A., Sneddon, A. A., Alessi, D. R., Dowd, S., and Keyse, S. M. (1996) Differential regulation of the MAP, SAP and RK/p38 kinases by Pyst1, a novel cytosolic dual-specificity phosphatase, *EMBO J. 15*, 3621-3632.

14. Muda, M., Theodosiou, A., Gillieron, C., Smith, A., Chabert, C., Camps, M., Boschert, U., Rodrigues, N., Davies, K., Ashworth, A., and Arkinstall, S. (1998) The mitogen-activated protein kinase phosphatase-3 N-terminal noncatalytic region is

responsible for tight substrate binding and enzymatic specificity, *J. Biol. Chem. 273*, 9323-9329.

15. Zhou, B. and Zhang, Z. Y. (1999) Mechanism of mitogen-activated protein kinase phosphatase-3 activation by ERK2, *J. Biol. Chem. 274*, 35526-35534.

16. Muda, M., Theodosiou, A., Rodrigues, N., Boschert, U., Camps, M., Gillieron, C., Davies, K., Ashworth, A., and Arkinstall, S. (1996) The dual specificity phosphatases M3/6 and MKP-3 are highly selective for inactivation of distinct mitogen-activated protein kinases, *J. Biol. Chem. 271*, 27205-27208.

17. Muda, M., Boschert, U., Dickinson, R., Martinou, J. C., Martinou, I., Camps, M., Schlegel, W., and Arkinstall, S. (1996) MKP-3, a novel cytosolic protein-tyrosine phosphatase that exemplifies a new class of mitogen-activated protein kinase phosphatase, *J. Biol. Chem. 271*, 4319-4326.

18. Furukawa, T., Fujisaki, R., Yoshida, Y., Kanai, N., Sunamura, M., Abe, T., Takeda, K., Matsuno, S., and Horii, A. (2005) Distinct progression pathways involving the dysfunction of DUSP6/MKP-3 in pancreatic intraepithelial neoplasia and intraductal papillary-mucinous neoplasms of the pancreas, *Mod. Pathol.*

19. Furukawa, T., Yatsuoka, T., Youssef, E. M., Abe, T., Yokoyama, T., Fukushige, S., Soeda, E., Hoshi, M., Hayashi, Y., Sunamura, M., Kobari, M., and Horii, A. (1998) Genomic analysis of DUSP6, a dual specificity MAP kinase phosphatase, in pancreatic cancer, *Cytogenet. Cell Genet. 82*, 156-159.

20. Furukawa, T., Sunamura, M., and Horii, A. (2006) Molecular mechanisms of pancreatic carcinogenesis, *Cancer Sci. 97*, 1-7.

21. Furukawa, T. and Horii, A. (2004) Molecular pathology of pancreatic cancer: in quest of tumor suppressor genes, *Pancreas 28*, 253-256.

22. Furukawa, T., Sunamura, M., Motoi, F., Matsuno, S., and Horii, A. (2003) Potential tumor suppressive pathway involving DUSP6/MKP-3 in pancreatic cancer, *Am. J. Pathol. 162*, 1807-1815.

23. Xu, S., Furukawa, T., Kanai, N., Sunamura, M., and Horii, A. (2005) Abrogation of DUSP6 by hypermethylation in human pancreatic cancer, *J. Hum. Genet. 50*, 159-167.

24. Toyota, T., Watanabe, A., Shibuya, H., Nankai, M., Hattori, E., Yamada, K., Kurumaji, A., Karkera, J. D., tera-Wadleigh, S. D., and Yoshikawa, T. (2000) Association study on the DUSP6 gene, an affective disorder candidate gene on 12q23, performed by using fluorescence resonance energy transfer-based melting curve analysis on the LightCycler, *Mol. Psychiatry 5*, 489-494.

25. Zhao, Y. and Zhang, Z. Y. (2001) The mechanism of dephosphorylation of extracellular signal-regulated kinase 2 by mitogen-activated protein kinase phosphatase 3, *J. Biol. Chem. 276*, 32382-32391.

26. Zhou, B., Wang, Z. X., Zhao, Y., Brautigan, D. L., and Zhang, Z. Y. (2002) The specificity of extracellular signal-regulated kinase 2 dephosphorylation by protein phosphatases, *J. Biol. Chem. 277*, 31818-31825.

27. Eblaghie, M. C., Lunn, J. S., Dickinson, R. J., Munsterberg, A. E., Sanz-Ezquerro, J. J., Farrell, E. R., Mathers, J., Keyse, S. M., Storey, K., and Tickle, C. (2003) Negative feedback regulation of FGF signaling levels by Pyst1/MKP3 in chick embryos, *Curr. Biol. 13*, 1009-1018.

28. Echevarria, D., Martinez, S., Marques, S., Lucas-Teixeira, V., and Belo, J. A. (2005) Mkp3 is a negative feedback modulator of Fgf8 signaling in the mammalian isthmic organizer, *Dev. Biol. 277*, 114-128.

29. Kim, M., Cha, G. H., Kim, S., Lee, J. H., Park, J., Koh, H., Choi, K. Y., and Chung, J. (2004) MKP-3 has essential roles as a negative regulator of the Ras/mitogen-activated protein kinase pathway during Drosophila development, *Mol. Cell Biol. 24*, 573-583.

30. Li, C., Scott, D. A., Hatch, E., Tian, X., and Mansour, S. L. (2007) Dusp6 (Mkp3) is a negative feedback regulator of FGF-stimulated ERK signaling during mouse development, *Development 134*, 167-176.

31. Smith, T. G., Sweetman, D., Patterson, M., Keyse, S. M., and Munsterberg, A. (2005) Feedback interactions between MKP3 and ERK MAP kinase control scleraxis expression and the specification of rib progenitors in the developing chick somite, *Development 132*, 1305-1314.

32. Castelli, M., Camps, M., Gillieron, C., Leroy, D., Arkinstall, S., Rommel, C., and Nichols, A. (2004) MAP kinase phosphatase 3 (MKP3) interacts with and is phosphorylated by protein kinase CK2alpha, *J. Biol. Chem. 279*, 44731-44739.

33. Turjanski, A. G., Vaque, J. P., and Gutkind, J. S. (2007) MAP kinases and the control of nuclear events, *Oncogene 26*, 3240-3253.

34. Raman, M., Chen, W., and Cobb, M. H. (2007) Differential regulation and properties of MAPKs, *Oncogene 26*, 3100-3112.

35. Dhillon, A. S., Hagan, S., Rath, O., and Kolch, W. (2007) MAP kinase signalling pathways in cancer, *Oncogene 26*, 3279-3290.

36. Dhanasekaran, D. N. and Johnson, G. L. (2007) MAPKs: function, regulation, role in cancer and therapeutic targeting, *Oncogene 26*, 3097-3099.

37. Meloche, S. and Pouyssegur, J. (2007) The ERK1/2 mitogen-activated protein kinase pathway as a master regulator of the G1- to S-phase transition, *Oncogene 26*, 3227-3239.

38. Dhanasekaran, D. N., Kashef, K., Lee, C. M., Xu, H., and Reddy, E. P. (2007) Scaffold proteins of MAP-kinase modules, *Oncogene 26*, 3185-3202.

39. Tanoue, T., Adachi, M., Moriguchi, T., and Nishida, E. (2000) A conserved docking motif in MAP kinases common to substrates, activators and regulators, *Nat. Cell Biol. 2*, 110-116.

40. Nichols, A., Camps, M., Gillieron, C., Chabert, C., Brunet, A., Wilsbacher, J., Cobb, M., Pouyssegur, J., Shaw, J. P., and Arkinstall, S. (2000) Substrate recognition domains within extracellular signal-regulated kinase mediate binding and catalytic activation of mitogen-activated protein kinase phosphatase-3, *J. Biol. Chem. 275*, 24613-24621.

41. Zhang, J., Zhou, B., Zheng, C. F., and Zhang, Z. Y. (2003) A bipartite mechanism for ERK2 recognition by its cognate regulators and substrates, *J. Biol. Chem. 278*, 29901-29912.

42. Zhou, B., Wu, L., Shen, K., Zhang, J., Lawrence, D. S., and Zhang, Z. Y. (2001) Multiple regions of MAP kinase phosphatase 3 are involved in its recognition and activation by ERK2, *J. Biol. Chem. 276*, 6506-6515.

43. Owens, D. M. and Keyse, S. M. (2007) Differential regulation of MAP kinase signalling by dual-specificity protein phosphatases, *Oncogene 26*, 3203-3213.

44. Dickinson, R. J. and Keyse, S. M. (2006) Diverse physiological functions for dual-specificity MAP kinase phosphatases, *J. Cell Sci. 119*, 4607-4615.

45. Kodama, M., Russell, D. S., and Duman, R. S. (2005) Electroconvulsive seizures increase the expression of MAP kinase phosphatases in limbic regions of rat brain, *Neuropsychopharmacology 30*, 360-371.

46. Pouyssegur, J., Volmat, V., and Lenormand, P. (2002) Fidelity and spatio-temporal control in MAP kinase (ERKs) signalling, *Biochem. Pharmacol. 64*, 755-763.

47. Hofken, T., Keller, N., Fleischer, F., Goke, B., and Wagner, A. C. (2000) Map kinase phosphatases (MKP's) are early responsive genes during induction of cerulein hyperstimulation pancreatitis, *Biochem. Biophys. Res. Commun. 276*, 680-685.

48. Sanchez-Perez, I., Martinez-Gomariz, M., Williams, D., Keyse, S. M., and Perona, R. (2000) CL100/MKP-1 modulates JNK activation and apoptosis in response to cisplatin, *Oncogene 19*, 5142-5152.

49. Timofeeva, A. V., Goryunova, L. E., Khaspekov, G. L., Kovalevskii, D. A., Scamrov, A. V., Bulkina, O. S., Karpov, Y. A., Talitskii, K. A., Buza, V. V., Britareva, V. V.,

and Beabealashvilli, R. S. (2006) Altered gene expression pattern in peripheral blood leukocytes from patients with arterial hypertension, *Ann. N. Y. Acad. Sci. 1091*, 319-335.

50. Tao, X. and Tong, L. (2007) Crystal structure of the MAP kinase binding domain and the catalytic domain of human MKP5, *Protein Sci. 16*, 880-886.

51. Christie, G. R., Williams, D. J., Macisaac, F., Dickinson, R. J., Rosewell, I., and Keyse, S. M. (2005) The dual-specificity protein phosphatase DUSP9/MKP-4 is essential for placental function but is not required for normal embryonic development, *Mol. Cell Biol. 25*, 8323-8333.

52. Hong, S. B., Lubben, T. H., Dolliver, C. M., Petrolonis, A. J., Roy, R. A., Li, Z., Parsons, T. F., Li, P., Xu, H., Reilly, R. M., Trevillyan, J. M., Nichols, A. J., Tummino, P. J., and Gant, T. G. (2005) Expression, purification, and enzymatic characterization of the dual specificity mitogen-activated protein kinase phosphatase, MKP-4, *Bioorg. Chem. 33*, 34-44.

53. Muda, M., Boschert, U., Smith, A., Antonsson, B., Gillieron, C., Chabert, C., Camps, M., Martinou, I., Ashworth, A., and Arkinstall, S. (1997) Molecular cloning and functional characterization of a novel mitogen-activated protein kinase phosphatase, MKP-4, *J. Biol. Chem. 272*, 5141-5151.

54. Cela, C. and Llimargas, M. (2006) Egfr is essential for maintaining epithelial integrity during tracheal remodelling in Drosophila, *Development 133*, 3115-3125.

55. Dickinson, R. J., Eblaghie, M. C., Keyse, S. M., and Morriss-Kay, G. M. (2002) Expression of the ERK-specific MAP kinase phosphatase PYST1/MKP3 in mouse embryos during morphogenesis and early organogenesis, *Mech. Dev. 113*, 193-196.

56. Eblaghie, M., Sanz-Ezquerro, J., Dickinson, R., Munsterberg, A., Keyse, S., and Tickle, C. (2002) 5 Pyst1/MKP3, a novel dual-specificity phosphatase, and FGF signalling in chick limb development, *J. Anat. 201*, 418.

57. Kawakami, Y., Rodriguez-Leon, J., Koth, C. M., Buscher, D., Itoh, T., Raya, A., Ng, J. K., Esteban, C. R., Takahashi, S., Henrique, D., Schwarz, M. F., Asahara, H., and Izpisua Belmonte, J. C. (2003) MKP3 mediates the cellular response to FGF8 signalling in the vertebrate limb, *Nat. Cell Biol. 5*, 513-519.

58. Tsang, M., Maegawa, S., Kiang, A., Habas, R., Weinberg, E., and Dawid, I. B. (2004) A role for MKP3 in axial patterning of the zebrafish embryo, *Development 131*, 2769-2779.

59. Vieira, C. and Martinez, S. (2005) Experimental study of MAP kinase phosphatase-3 (Mkp3) expression in the chick neural tube in relation to Fgf8 activity, *Brain Res. Brain Res. Rev. 49*, 158-166.

60. Zhang, Z. Y. (2002) Protein tyrosine phosphatases: structure and function, substrate specificity, and inhibitor development, *Annu. Rev. Pharmacol. Toxicol. 42*, 209-234.

61. Fjeld, C. C., Rice, A. E., Kim, Y., Gee, K. R., and Denu, J. M. (2000) Mechanistic basis for catalytic activation of mitogen-activated protein kinase phosphatase 3 by extracellular signal-regulated kinase, *J. Biol. Chem. 275*, 6749-6757.

62. Kim, Y., Rice, A. E., and Denu, J. M. (2003) Intramolecular dephosphorylation of ERK by MKP3, *Biochemistry 42*, 15197-15207.

63. Rigas, J. D., Hoff, R. H., Rice, A. E., Hengge, A. C., and Denu, J. M. (2001) Transition state analysis and requirement of Asp-262 general acid/base catalyst for full activation of dual-specificity phosphatase MKP3 by extracellular regulated kinase, *Biochemistry 40*, 4398-4406.

64. Karlsson, M., Mathers, J., Dickinson, R. J., Mandl, M., and Keyse, S. M. (2004) Both nuclear-cytoplasmic shuttling of the dual specificity phosphatase MKP-3 and its ability to anchor MAP kinase in the cytoplasm are mediated by a conserved nuclear export signal, *J. Biol. Chem. 279*, 41882-41891.

65. Colucci-D'Amato, G. L., D'Alessio, A., Califano, D., Cali, G., Rizzo, C., Nitsch, L., Santelli, G., and de, F., V (2000) Abrogation of nerve growth factor-induced terminal differentiation by ret oncogene involves perturbation of nuclear translocation of ERK, *J. Biol. Chem. 275*, 19306-19314.

66. Mourey, R. J., Vega, Q. C., Campbell, J. S., Wenderoth, M. P., Hauschka, S. D., Krebs, E. G., and Dixon, J. E. (1996) A novel cytoplasmic dual specificity protein tyrosine phosphatase implicated in muscle and neuronal differentiation, *J. Biol. Chem. 271*, 3795-3802.

67. Zhou, B. and Zhang, Z. Y. (2007) Application of hydrogen/deuterium exchange mass spectrometry to study protein tyrosine phosphatase dynamics, ligand binding, and substrate specificity, *Methods 42*, 227-233.

68. Butcher, G. Q., Lee, B., and Obrietan, K. (2003) Temporal regulation of light-induced extracellular signal-regulated kinase activation in the suprachiasmatic nucleus, *J. Neurophysiol. 90*, 3854-3863.

69. Fisher, D. L., Brassac, T., Galas, S., and Doree, M. (1999) Dissociation of MAP kinase activation and MPF activation in hormone-stimulated maturation of Xenopus oocytes, *Development 126*, 4537-4546.

70. Lunn, J. S., Fishwick, K. J., Halley, P. A., and Storey, K. G. (2007) A spatial and temporal map of FGF/Erk1/2 activity and response repertoires in the early chick embryo, *Dev. Biol. 302*, 536-552.

71. Barford, D., Flint, A. J., and Tonks, N. K. (1994) Crystal structure of human protein tyrosine phosphatase 1B, *Science 263*, 1397-1404.

72. Doman, T. N., McGovern, S. L., Witherbee, B. J., Kasten, T. P., Kurumbail, R., Stallings, W. C., Connolly, D. T., and Shoichet, B. K. (2002) Molecular docking and high-throughput screening for novel inhibitors of protein tyrosine phosphatase-1B, *J. Med. Chem. 45*, 2213-2221.

73. Mustelin, T., Feng, G. S., Bottini, N., Alonso, A., Kholod, N., Birle, D., Merlo, J., and Huynh, H. (2002) Protein tyrosine phosphatases, *Front Biosci. 7:d85-142.*, d85-142.

74. Stewart, A. E., Dowd, S., Keyse, S. M., and McDonald, N. Q. (1999) Crystal structure of the MAPK phosphatase Pystl catalytic domain and implications for regulated activation, *Nat. Struct. Biol. 6*, 174-181.

75. Farooq, A., Chaturvedi, G., Mujtaba, S., Plotnikova, O., Zeng, L., Dhalluin, C., Ashton, R., and Zhou, M. M. (2001) Solution structure of ERK2 binding domain of MAPK phosphatase MKP-3: structural insights into MKP-3 activation by ERK2, *Mol. Cell 7*, 387-399.

76. Farooq, A., Zeng, L., and Zhou, M. M. (2001) 1H, 13C and 15N resonance assignments of the ERK2 binding domain of the MAPK phosphatase MKP-3, *J. Biomol. NMR 19*, 195-196.

77. Zhou, B., Zhang, J., Liu, S., Reddy, S., Wang, F., and Zhang, Z. Y. (2006) Mapping ERK2-MKP3 binding interfaces by hydrogen/deuterium exchange mass spectrometry, *J. Biol. Chem. 281*, 38834-38844.

78. Yoshida, E., Atkinson, T. G., and Chakravarthy, B. (2004) Neuroprotective gene expression profiles in ischemic cortical cultures preconditioned with IGF-1 or bFGF, *Brain Res. Mol. Brain Res. 131*, 33-50.

79. Itoh, M., Tsukada, S., Orita, T., Nishiu, J., Tomoike, H., Nakamura, Y., and Tanaka, T. (1998) Identification by differential display of eight known genes induced during in vivo intimal hyperplasia, *J. Hum. Genet. 43*, 9-13.

80. Smith, A., Price, C., Cullen, M., Muda, M., King, A., Ozanne, B., Arkinstall, S., and Ashworth, A. (1997) Chromosomal localization of three human dual specificity phosphatase genes (DUSP4, DUSP6, and DUSP7), *Genomics 42*, 524-527.

81. Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature 227*, 680-685.

82. Lobley, A., Whitmore, L., and Wallace, B. A. (2002) DICHROWEB: an interactive website for the analysis of protein secondary structure from circular dichroism spectra, *Bioinformatics. 18*, 211-212.

83. Whitmore, L. and Wallace, B. A. (2004) DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data, *Nucleic Acids Res. 32*, W668-W673.

84. Whitmore, L. and Wallace, B. A. (2007) Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases, *Biopolymers*.

85. Lovrien, R. and Daumantas, M. (1995) Assays for Total Protein, in *Current protocols in protein science* (Taylor, G. P., Ed.) pp 3.4.1-3.4.24, John Wiley and Sons, Inc., Hoboken.

86. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., and Bairoch, A. (2005) Protein Identification and Analysis Tools on the ExPASy Server, in *The Proteomics Protocols Handbook* (Walker, J. M., Ed.) pp 571-607, Humana Press, Totowa.

87. Cheng, Y. and Prusoff, W. H. (1973) Relationship between the inhibition constant (K1) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction, *Biochem. Pharmacol. 22*, 3099-3108.

88. Mark, J. K., Smith, S., and Hefford, M. A. (2007) Over-expression and refolding of MAP kinase phosphatase 3, *Protein Expr. Purif. 54*, 253-260.

89. Scopes, R. K. (1995) Strategies for Protein Purification, in *Current protocols in protein science* (Taylor, G. P., Ed.) pp 1.2.1-1.2.4, John Wiley and Sons, Inc., Hoboken.

90. Scopes, R. K. (1995) Protein Purification Flow Charts, in *Current protocols in protein science* (Taylor, G. P., Ed.) pp 1.3.1-1.3.7, John Wiley and Sons, Inc., Hoboken.

91. Bennett, M. J., Schlunegger, M. P., and Eisenberg, D. (1995) 3D domain swapping: a mechanism for oligomer assembly, *Protein Sci. 4*, 2455-2468.

92. Liu, Y. and Eisenberg, D. (2002) 3D domain swapping: as domains continue to swap, *Protein Sci. 11*, 1285-1299.

93. Rousseau, F., Schymkowitz, J. W., and Itzhaki, L. S. (2003) The unfolding story of three-dimensional domain swapping, *Structure. 11*, 243-251.

94. Muda, M., Boschert, U., Dickinson, R., Martinou, J. C., Martinou, I., Camps, M., Schlegel, W., and Arkinstall, S. (1996) MKP-3, a novel cytosolic protein-tyrosine phosphatase that exemplifies a new class of mitogen-activated protein kinase phosphatase, *J. Biol. Chem. 271*, 4319-4326.

95. Wiland, A. M., Denu, J. M., Mourey, R. J., and Dixon, J. E. (1996) Purification and kinetic characterization of the mitogen-activated protein kinase phosphatase rVH6, *J. Biol. Chem. 271*, 33486-33492.

96. Bennett, M. J., Schlunegger, M. P., and Eisenberg, D. (1995) 3D domain swapping: a mechanism for oligomer assembly, *Protein Sci. 4*, 2455-2468.

97. Liu, Y. and Eisenberg, D. (2002) 3D domain swapping: as domains continue to swap, *Protein Sci. 11*, 1285-1299.

98. Rousseau, F., Schymkowitz, J. W., and Itzhaki, L. S. (2003) The unfolding story of three-dimensional domain swapping, *Structure. 11*, 243-251.

99. Pasha, M. K., Dimmock, J. R., Hollenberg, M. D., and Sharma, R. K. (2002) Enhanced activity of human N-myristoyltransferase by dimethyl sulfoxide and related solvents in the presence of serine/threonine-containing peptide substrates, *Biochem. Pharmacol. 64*, 1461-1467.

100. Wehbi, H., Feng, J., and Roberts, M. F. (2003) Water-miscible organic cosolvents enhance phosphatidylinositol-specific phospholipase C phosphotransferase as well as phosphodiesterase activity, *Biochim. Biophys. Acta 1613*, 15-27.

101. Bennett, M. J., Schlunegger, M. P., and Eisenberg, D. (1995) 3D domain swapping: a mechanism for oligomer assembly, *Protein Sci. 4*, 2455-2468.

102. Liu, Y. and Eisenberg, D. (2002) 3D domain swapping: as domains continue to swap, *Protein Sci. 11*, 1285-1299.

103. Rousseau, F., Schymkowitz, J. W., and Itzhaki, L. S. (2003) The unfolding story of three-dimensional domain swapping, *Structure. 11*, 243-251.

104. Wang, J. and Walsh, C. T. (1997) Mechanistic studies on full length and the catalytic domain of the tandem SH2 domain-containing protein tyrosine phosphatase: analysis of phosphoenzyme levels and Vmax stimulatory effects of glycerol and of a phosphotyrosyl peptide ligand, *Biochemistry 36*, 2993-2999.

105. Jiang, G., den Hertog, J., and Hunter, T. (2000) Receptor-like protein tyrosine phosphatase alpha homodimerizes on the cell surface, *Mol. Cell Biol. 20*, 5917-5929.

106. Tonks, N. K. (2006) Protein tyrosine phosphatases: from genes, to function, to disease, *Nat. Rev. Mol. Cell Biol. 7*, 833-846.

107. Marra, M. A., Jones, S. J., Astell, C. R., Holt, R. A., Brooks-Wilson, A., Butterfield, Y. S., Khattra, J., Asano, J. K., Barber, S. A., Chan, S. Y., Cloutier, A., Coughlin, S. M., Freeman, D., Girn, N., Griffith, O. L., Leach, S. R., Mayo, M., McDonald, H., Montgomery, S. B., Pandoh, P. K., Petrescu, A. S., Robertson, A. G., Schein, J. E., Siddiqui, A., Smailus, D. E., Stott, J. M., Yang, G. S., Plummer, F., Andonov, A., Artsob, H., Bastien, N., Bernard, K., Booth, T. F., Bowness, D., Czub, M., Drebot, M., Fernando, L., Flick, R., Garbutt, M., Gray, M., Grolla, A., Jones, S., Feldmann, H., Meyers, A., Kabani, A., Li, Y., Normand, S., Stroher, U., Tipples, G. A., Tyler, S., Vogrig, R., Ward, D., Watson, B., Brunham, R. C., Krajden, M., Petric, M., Skowronski, D. M., Upton, C., and Roper, R. L. (2003) The Genome sequence of the SARS-associated coronavirus, *Science 300*, 1399-1404.

108. Drosten, C., Gunther, S., Preiser, W., van der, W. S., Brodt, H. R., Becker, S., Rabenau, H., Panning, M., Kolesnikova, L., Fouchier, R. A., Berger, A., Burguiere, A. M., Cinatl, J., Eickmann, M., Escriou, N., Grywna, K., Kramme, S., Manuguerra, J. C., Muller, S., Rickerts, V., Sturmer, M., Vieth, S., Klenk, H. D., Osterhaus, A. D., Schmitz, H., and Doerr, H. W. (2003) Identification of a novel coronavirus in patients with severe acute respiratory syndrome, *N. Engl. J. Med. 348*, 1967-1976.

109. Navas-Martin, S. and Weiss, S. R. (2003) SARS: lessons learned from other coronaviruses, *Viral Immunol. 16*, 461-474.

110. Jackwood, M. W. (2006) The relationship of severe acute respiratory syndrome coronavirus with avian and other coronaviruses, *Avian Dis. 50*, 315-320.

111. Holmes, E. C. and Rambaut, A. (2004) Viral evolution and the emergence of SARS coronavirus, *Philos. Trans. R. Soc. Lond B Biol. Sci. 359*, 1059-1065.

112. Shortridge, K. F. (2003) Severe acute respiratory syndrome and influenza: virus incursions from southern China, *Am. J. Respir. Crit Care Med. 168*, 1416-1420.

113. He, R., Leeson, A., Ballantine, M., Andonov, A., Baker, L., Dobie, F., Li, Y., Bastien, N., Feldmann, H., Strocher, U., Theriault, S., Cutts, T., Cao, J., Booth, T. F., Plummer, F. A., Tyler, S., and Li, X. (2004) Characterization of protein-protein interactions between the nucleocapsid protein and membrane protein of the SARS coronavirus, *Virus Res. 105*, 121-125.

114. He, R., Dobie, F., Ballantine, M., Leeson, A., Li, Y., Bastien, N., Cutts, T., Andonov, A., Cao, J., Booth, T. F., Plummer, F. A., Tyler, S., Baker, L., and Li, X. (2004) Analysis of multimerization of the SARS coronavirus nucleocapsid protein, *Biochem. Biophys. Res. Commun. 316*, 476-483.

115. Yu, I. M., Gustafson, C. L., Diao, J., Burgner, J. W., Li, Z., Zhang, J., and Chen, J. (2005) Recombinant severe acute respiratory syndrome (SARS) coronavirus nucleocapsid protein forms a dimer through its C-terminal domain, *J. Biol. Chem. 280*, 23280-23286.

116. Wang, Y., Wu, X., Wang, Y., Li, B., Zhou, H., Yuan, G., Fu, Y., and Luo, Y. (2004) Low stability of nucleocapsid protein in SARS virus, *Biochemistry 43*, 11103-11108.

117. Wang, J., Ji, J., Ye, J., Zhao, X., Wen, J., Li, W., Hu, J., Li, D., Sun, M., Zeng, H., Hu, Y., Tian, X., Tan, X., Xu, N., Zeng, C., Wang, J., Bi, S., and Yang, H. (2003) The structure analysis and antigenicity study of the N protein of SARS-CoV, *Genomics Proteomics. Bioinformatics. 1*, 145-154.

118. Luo, H., Ye, F., Sun, T., Yue, L., Peng, S., Chen, J., Li, G., Du, Y., Xie, Y., Yang, Y., Shen, J., Wang, Y., Shen, X., and Jiang, H. (2004) In vitro biochemical and thermodynamic characterization of nucleocapsid protein of SARS, *Biophys. Chem. 112*, 15-25.

119. Duan, S. M., Zhao, X. S., Wen, R. F., Huang, J. J., Pi, G. H., Zhang, S. X., Han, J., Bi, S. L., Ruan, L., and Dong, X. P. (2003) Stability of SARS coronavirus in human specimens and environment and its sensitivity to heating and UV irradiation, *Biomed. Environ. Sci. 16*, 246-255.

120. Garzon, M. T., Lidon-Moya, M. C., Barrera, F. N., Prieto, A., Gomez, J., Mateu, M. G., and Neira, J. L. (2004) The dimerization domain of the HIV-1 capsid protein binds a capsid protein-derived peptide: a biophysical characterization, *Protein Sci. 13*, 1512-1523.

121. Ihnat, P. M., Vellekamp, G., Obenauer-Kutner, L. J., Duan, J., Han, M. A., Witchey-Lakshmanan, L. C., and Grace, M. J. (2005) Comparative thermal stabilities of recombinant adenoviruses and hexon protein, *Biochim. Biophys. Acta 1726*, 138-151.

122. Lidon-Moya, M. C., Barrera, F. N., Bueno, M., Perez-Jimenez, R., Sancho, J., Mateu, M. G., and Neira, J. L. (2005) An extensive thermodynamic characterization of the dimerization domain of the HIV-1 capsid protein, *Protein Sci. 14*, 2387-2404.

123. Yelamos, B., Nunez, E., Gomez-Gutierrez, J., Datta, M., Pacheco, B., Peterson, D. L., and Gavilanes, F. (1999) Circular dichroism and fluorescence spectroscopic properties of the major core protein of feline immunodeficiency virus and its tryptophan mutants. Assignment of the individual contribution of the aromatic sidechains, *Eur. J. Biochem. 266*, 1081-1089.

124. Huang, I. C., Bosch, B. J., Li, F., Li, W., Lee, K. H., Ghiran, S., Vasilieva, N., Dermody, T. S., Harrison, S. C., Dormitzer, P. R., Farzan, M., Rottier, P. J., and Choe, H. (2006) SARS coronavirus, but not human coronavirus NL63, utilizes cathepsin L to infect ACE2-expressing cells, *J. Biol. Chem. 281*, 3198-3203.

125. Yu, I. M., Oldham, M. L., Zhang, J., and Chen, J. (2006) Crystal structure of the severe acute respiratory syndrome (SARS) coronavirus nucleocapsid protein dimerization domain reveals evolutionary linkage between corona- and arteriviridae, *J. Biol. Chem. 281*, 17134-17139.

126. Saikatendu, K. S., Joseph, J. S., Subramanian, V., Neuman, B. W., Buchmeier, M. J., Stevens, R. C., and Kuhn, P. (2007) Ribonucleocapsid formation of severe acute respiratory syndrome coronavirus through molecular action of the N-terminal domain of N protein, *J. Virol. 81*, 3913-3921.

127. Ying, W., Hao, Y., Zhang, Y., Peng, W., Qin, E., Cai, Y., Wei, K., Wang, J., Chang, G., Sun, W., Dai, S., Li, X., Zhu, Y., Li, J., Wu, S., Guo, L., Dai, J., Wang, J., Wan, P., Chen, T., Du, C., Li, D., Wan, J., Kuai, X., Li, W., Shi, R., Wei, H., Cao, C., Yu, M., Liu, H., Dong, F., Wang, D., Zhang, X., Qian, X., Zhu, Q., and He, F. (2004) Proteomic analysis on structural proteins of Severe Acute Respiratory Syndrome coronavirus, *Proteomics. 4*, 492-504.

128. Fang, X., Ye, L. B., Zhang, Y., Li, B., Li, S., Kong, L., Wang, Y., Zheng, H., Wang, W., and Wu, Z. (2006) Nucleocapsid amino acids 211 to 254, in particular, tetrad

glutamines, are essential for the interaction between the nucleocapsid and membrane proteins of SARS-associated coronavirus, *J. Microbiol. 44*, 577-580.

129. You, J. H., Reed, M. L., and Hiscox, J. A. (2007) Trafficking motifs in the SARS-coronavirus nucleocapsid protein, *Biochem. Biophys. Res. Commun. 358*, 1015-1020.

130. Schagger, H. and von, J. G. (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa, *Anal. Biochem. 166*, 368-379.

131. Twining, S. S. (1984) Fluorescein Isothiocyanate-labeled Casein Assay for Proteolytic Enzymes, *Anal. Biochem. 143*, 30-34.

132. Guruprasad, K., Reddy, B. V., and Pandit, M. W. (1990) Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence, *Protein Eng 4*, 155-161.

133. He, R., Leeson, A., Ballantine, M., Andonov, A., Baker, L., Dobie, F., Li, Y., Bastien, N., Feldmann, H., Strocher, U., Theriault, S., Cutts, T., Cao, J., Booth, T. F., Plummer, F. A., Tyler, S., and Li, X. (2004) Characterization of protein-protein interactions between the nucleocapsid protein and membrane protein of the SARS coronavirus, *Virus Res. 105*, 121-125.

134. Chen, C. Y., Chang, C. K., Chang, Y. W., Sue, S. C., Bai, H. I., Riang, L., Hsiao, C. D., and Huang, T. H. (2007) Structure of the SARS coronavirus nucleocapsid protein RNA-binding dimerization domain suggests a mechanism for helical packaging of viral RNA, *J. Mol. Biol. 368*, 1075-1086.

135. Neuman, B. W., Adair, B. D., Yoshioka, C., Quispe, J. D., Orca, G., Kuhn, P., Milligan, R. A., Yeager, M., and Buchmeier, M. J. (2006) Supramolecular architecture of severe acute respiratory syndrome coronavirus revealed by electron cryomicroscopy, *J. Virol. 80*, 7918-7928.

136. Tan, Y. W., Fang, S., Fan, H., Lescar, J., and Liu, D. X. (2006) Amino acid residues critical for RNA-binding in the N-terminal domain of the nucleocapsid protein are essential determinants for the infectivity of coronavirus in cultured cells, *Nucleic Acids Res. 34*, 4816-4825.

11.0. Contributions of collaborators

I would like to thank all the individuals who had a hand in the completion of this thesis, either by lending their time or assistance to help resolve the many problems that I encountered over the years. Specific contributions, in order of thesis appearance, include:

Section 3.1. MKP3 Production. Dr. Remy Aubin supplied the original *Mkp3* cDNA that was used to produce recombinant MKP3 proteins. Sophie Smith assisted with MKP3 cloning, mutagenesis and DNA sequencing. Dr. Terry Cry performed mass spectrometric analysis to confirm the correct production of MKP3 and mutants.

Section 3.3.4. Immunodetection of MKP3 in pancreatic adenocarcinoma cells. Dr. Remy Aubin supplied human pancreatic and pancreatic adenocarcinoma cell lines and provided the experimental results contained in Figure 15.

Section 7.1. SARS N protein expression and purification. Dr. Sean Li supplied a SARS N protein cDNA and M15 *E. coli* cells. Both Bozena Jaentschke and Sylvie Fournier assisted equally with SARS N protein production and SDS-PAGE separation presented in this section.

Section 7.2. Mass spectrometric characterization of protein fragments. Sophie Smith assisted with the MALDI-TOF/MS and Dr. Terry Cry assisted with the LC/MS data acquisitions presented in this section.

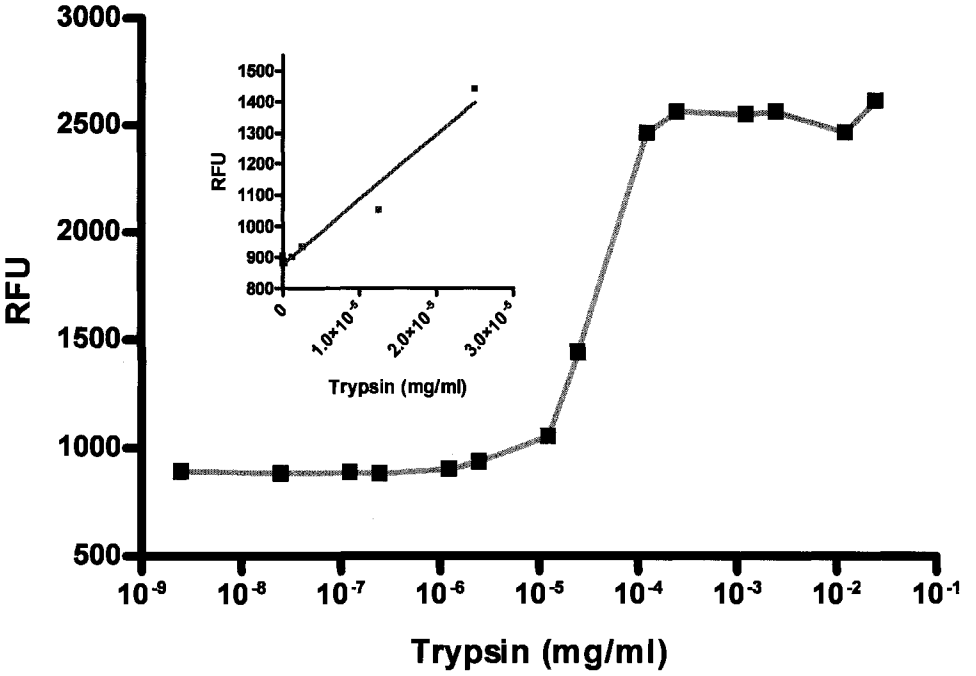
Section 7.3.3. Protease detection using ovalbumin or RNAse A mixing. Sylvie Fournier assisted by performing SDS-PAGE separation in this section.

12.0. Appendix A.

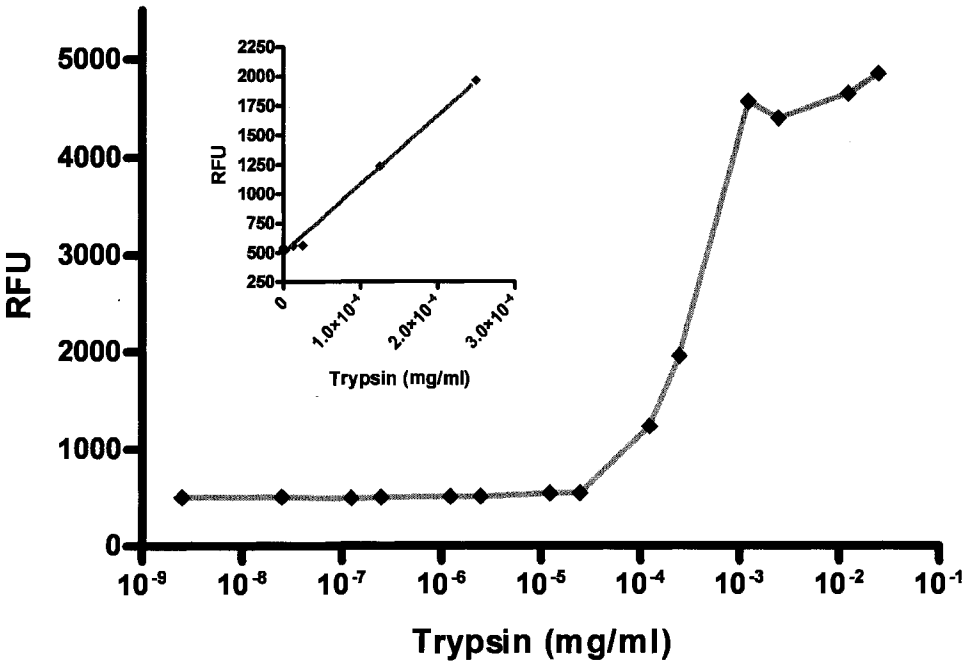
The characteristics of the EDANS/DABCYL-conjugated peptide cleavage assays were determined using trypsin. Trypsin concentration was increased as detailed in Methods and relative increase in fluorescence was detected (section 6.1.10). Peptide cleavage led to in increase in fluorescence and data was plotted as a function of trypsin concentration (log scale) to determine the limit of detection (Figure 36). Results show the detection limits were ~2.5 ng/mL trypsin for the ^{181}SQASSRSSSRSRGNSRNSTP^{200}- peptide (Fgure 36A) and ~12.5 ng/mL for the ^{168}LPKGFYAEGSRGGSQASS^{185}- peptide (Figure 36B). The inset graphs show the linear ranges for the assays. Results from linear regression analysis showed the ^{181}SQASSRSSSRSRGNSRNSTP^{200}- peptide assay was linear between 2.5-25 ng/mL trypsin, with a linear correlation coefficient (r^2) of 0.96. The ^{168}LPKGFYAEGSRGGSQASS^{185}- peptide linear range was 12.5-250 ng/mL with an r^2 value of 0.99.

Figure 36. Trypsin cleavage of EDANS/DABCYL-conjugated peptides. Assays were performed as described in section 6.1.10, using a 60-minute incubation time. The relative increase in fluorescence (in relative fluorescence units, RFU), from cleavage of SARS N protein peptides, is shown as a function of trypsin concentration (on a log scale). The inset graph plots the linear region of the assay on a linear scale.**A.** Cleavage of the ^{181}SQASSRSSSRSRGNSRNSTP^{200}- peptide. **B.** Cleavage of the ^{168}LPKGFYAEGSRGGSQASS^{185}- peptide.

A.


RFU
3000
2500
2000
1500
1000
500
10^{-9} 10^{-8} 10^{-7} 10^{-6} 10^{-5} 10^{-4} 10^{-3} 10^{-2} 10^{-1}
Trypsin (mg/ml)
RFU
1500
1400
1300
1200
1100
1000
900
800
0
1.0×10⁻⁵
2.0×10⁻⁵
3.0×10⁻⁵
Trypsin (mg/ml)

B.


RFU
5000
4000
3000
2000
1000
0
10^{-9} 10^{-8} 10^{-7} 10^{-6} 10^{-5} 10^{-4} 10^{-3} 10^{-2} 10^{-1}
Trypsin (mg/ml)
RFU
2250
2000
1750
1500
1250
1000
750
500
250
0
1.0×10⁻⁴
2.0×10⁻⁴
3.0×10⁻⁴
Trypsin (mg/ml)

Curriculum Vitae

John Mark

Education:

PhD. (Biochemistry); Expected Graduation 12/07 Years Attended: 2001 - 07
University of Ottawa Ottawa, ON
Advisor: Dr. Mary Alice Hefford

M.Sc. (Physiology); April 2001
University of Manitoba Winnipeg, MB
Thesis: A novel magnetic resonance imaging technique to assess tissue viability following acute myocardial infarction: Simultaneous T1 and T2* signal intensity monitoring following bolus injected contrast agent.
Advisor: Dr. Roxanne Deslauriers, Dr Ganghong Tian

B.Sc. Major (Zoology); April 1997
University of Manitoba Winnipeg, MB

Employment:

President
Kram Holdings Winnipeg, MB
Years Serving: 1998 – Present

Director
Vietnamese Folkdance Association of Manitoba Winnipeg, MB
Years Serving: 2001 – Present

Director
Canadians Helping Kids in Vietnam Winnipeg, MB
Years Serving: 1999 – 2005

Professional Memberships:

- Member Biophysics Society of Canada
- Member International Society for Magnetic Resonance in Medicine
- Member Society for Cardiovascular Magnetic Resonance
- Member Canadian Society of Biochemistry and Molecular & Cellular Biology

Awards:

- Health Canada Science Forum Best Poster Award (2008)
- University of Ottawa Graduate Admissions Scholarship (2004-2005)
- University of Ottawa Excellence Award (2001-2003)
- Biophysical Society of Canada poster competition prize (Canadian Federation of Biological Sciences meeting, 1999)
- Canadian Institutes for Health Research Fellowship (1998-2004)
- Graduated with Distinction (B.Sc. Zoology program, University of Manitoba)

Manuscripts:

Mark, J., Li, X., Cyr, T., Fournier, S., Jaentschke, B., and Hefford, M. A. SARS coronavirus: Unusual lability of the nucleocapsid protein, Biochem. Biophys. Res. Commun. 2008 [Epub ahead of print]

Mark, J. K., Aubin, R. A., Smith, S., and Hefford, M. A. Inhibition of Mitogen-activated Protein Kinase Phosphatase 3 Activity by Interdomain Binding, J. Biol. Chem. 2008 Oct; 283(42), 28574-28583.

Mark, JK. Smith, S., Hefford, MA. Over-expression and refolding of MAP kinase phosphatase 3. Protein Expr Purif. 2007 Aug; 54(2):253-260

Li G, Dai G, Xiang B, **Mark J.**, Tomanek B, Liu H, Deslauriers R, Tian G. Mapping myocardial viability using interleaved T1-T2* weighted imaging. Int J Cardiovasc Imaging. 2004 Apr; 20(2):135-143.

Tian, G., **Mark, J.**, Dai, G., Xiang, B., Li, G., Sun, J., Lazarow, N., Corne, R. and Deslauriers, R. Assessment of myocardial injury using magnetic resonance imaging. In Progress in Experimental Cardiology , 10: Pathophysiology of Cardiovascular Disease (Dhalla, N.S., Rupp, H., Angel, A. and Pierce, G.N., Eds.) Chapter 40, Boston: Kluwer Academic Publishers, 2003, 577-594.

Mark J, Dai G, Xiang B, Sun J, Lazarow N, Lindsay WG, Corne R, Deslauriers R, Tian G. Simultaneously Monitoring Both T1 and T2* Signal Intensities on a Bolus Injection of Gd-DTPA May Distinguish Infarcted Myocardium. J Magn Res Imag 2002 May; 15:532-540.

Mark J, Dai G, Xiang B, Sun J, Lindsay WG, Deslauriers R, Tian G. Contrast agent distribution in microvascular damage of infarcted pig myocardium. Acta Radiol. 2001 Sep; 42(5):515-520.

Selected oral presentations:

Mark, J., Dai, G., Xiang, B., Sun, J., Lazarow, N., Deslauriers, R., Tian, G."Is Contrast Induced T1 Shortening Necessarily Correlated With Myocardial Injury?". 42nd Annual Meeting, Canadian Federation of Biological Societies. Program and Proceedings, 1999.

Tian, G., **Mark, J.**, Dai, G., Xiang, B., Sun, J., Lazarow, N., Deslauriers, R."Is Contrast Induced T1 Shortening Necessarily Correlated With Myocardial Injury? Assessment of myocardial infarction using contrast-enhanced MR Imaging and localized 31P spectroscopy on isolated pig hearts". Proceedings of the International Society of Magnetic Resonance in Medicine, Seventh Annual Meeting and Exhibition, Philadelphia, USA, May22-28, Vol. 7: 1490, 1999.

Extracurricular activities:

- University of Manitoba, Canadian Chinese Student's Association
- Member: University of Manitoba, Winnipeg WASPS rugby club
- Volunteer: University of Manitoba, International Students' Center (Tutor for Volunteer English Practice Program)
- Volunteer: Big Sisters and Big Brothers Organization, Winnipeg, MB (Mentor)
- Volunteer: Mulvey Elementary School, Winnipeg, MB (Tutor)
- Volunteer: Canadians Helping Kids in Vietnam (Sponsorship coordinator)