

**IDENTIFICATION, CHARACTERIZATION AND
EXPRESSION ANALYSIS OF A NOVEL TPA (12-O-
TETRADECANOYLPHORBOL-13-ACETATE) INDUCED
GENE**

CHAN CHUNG YIP

(M.B.,B.S. (NUS), M.Med (Surg), MRCS (Edin))

A THESIS SUBMITTED FOR THE DEGREE OF DOCTOR OF MEDICINE

DEPARTMENT OF BIOCHEMISTRY

NATIONAL UNIVERSITY OF SINGAPORE

2006

ACKNOWLEDGMENTS

I would like to thank my thesis supervisor, Dr Caroline Lee, for her constant support and guidance, and Dr Thomas Adrian at the Northwestern University in Chicago, who has kindly allowed me to conduct my experiments leading to this thesis in his laboratory. I especially owe my gratitude to Dr Xianzhong Ding, research assistant professor at the same laboratory, for his mentorship and faith that he placed in my work.

I would like to thank the National Medical Research Council and Tan Tock Seng Hospital for sponsoring me in this endeavour. My sincere appreciation to colleagues in the Department of General Surgery, Tan Tock Seng Hospital, for their friendship and encouragement.

Lastly, and certainly not in the least, I would like to dedicate this thesis to my wife, Rachel, and my family, who have made the completion of this possible.

TABLE OF CONTENTS

<u>CHAPTER</u>		<u>PAGE</u>
1.	INTRODUCTION	1
	1.1 Introduction to Pancreatic Cancer	1
	1.1.1 The Pancreas	1
	1.1.2 Cancer of the Pancreas	2
	1.1.3 Epidemiology of Pancreatic Cancer	3
	1.1.4 Molecular Genetics of Pancreatic Adenocarcinoma.....	4
	1.2 Analyzing Differential Gene Expression in Cancer.....	14
	1.2.1 Protein Gel Electrophoresis and Modern Day Proteomics...	15
	1.2.2 Differential Hybridization.....	16
	1.2.3 Subtractive Hybridization	17
	1.2.4 Differential Display.....	17
	1.2.5 Microarrays	18
	1.2.6 Expressed Sequence Tags (ESTs) and SAGE	19
	1.3 Biology of PKC and TPA.....	20
	1.3.1 Cell Growth and Tumour Promotion	21
	1.3.2 PKCs and Pancreatic Cancer	22
	1.4 Biology of Transmembrane/ ER Proteins	23
	1.4.1 Orientation and Conformation of the Transmembrane Protein.....	24
	1.4.2 Protein Glycosylation.....	25
	1.5 Transcriptional Regulation.....	25
	1.5.1 Organisation of the Promoter.....	25
	1.5.2 RNA Polymerase II Core Promoter Elements.....	27
	1.5.3 Sp1/KLF Family of Transcriptional Factors.....	30
	1.6 Future Directions.....	32
2.	HYPOTHESIS AND AIMS	34
3.	MATERIALS AND METHODS.....	35
	3.1 Microarray and Identification of Novel Gene	35
	3.1.1 Cell Culture	35
	3.1.2 RNA Extraction	36
	3.1.3 Oligonucleotide Array Gene Expression Analysis.....	36
	3.1.4 Reverse Transcription and Real-Time Quantitative PCR...	38
	3.1.5 Rapid Amplification of cDNA Ends (RACE).....	39
	3.1.6 Construction of Plasmid for Promoter Analysis.....	39
	3.1.7 Transient Transfection.....	40
	3.1.8 Reporter Gene Assay.....	40
	3.2 Expression, Structural and Functional Characterization	40
	3.2.1 Cell Culture and Transfection Protocol.....	40
	3.2.2 Real-Time RT-PCR Analysis of mRNA Expression in Human Tissues and Cancer Cells	41
	3.2.3 Plasmids Construction	42
	3.2.4 Western Blotting	43
	3.2.5 Deglycosylation Assay.....	43
	3.2.6 Immunofluorescence	44
	3.2.7 Cell Proliferation Assay by Cell Counting.....	44

TABLE OF CONTENTS (continued)

<u>CHAPTER</u>	<u>PAGE</u>
3.2.8	45
3.2.9	45
3.2.10	46
3.3	46
3.3.1	46
3.3.2	47
3.3.3	48
3.3.4	49
3.3.5	49
3.4	50
3.4.1	50
3.4.2	50
4.	51
4.1	51
4.1.1	52
4.1.2	55
4.1.3	60
4.1.4	61
4.1.5	63
4.1.6	64
4.2	65
4.2.1	65
4.2.2	67
4.2.3	70
4.2.4	73
4.2.5	74
4.2.6	78
4.2.7	80
4.2.8	80
4.2.9	82
4.3	82
4.3.1	82
4.3.2	83
4.3.3	86
4.3.4	88
4.3.5	90

TABLE OF CONTENTS (continued)

<u>CHAPTER</u>		<u>PAGE</u>
5.	DISCUSSION AND CONCLUSIONS	92
	REFERENCES	98

SUMMARY

Pancreatic cancer is a deadly disease with very poor prognosis. The phorbol ester TPA has been found to have opposite effects on pancreatic cancer cell growth and proliferation. Hence we hypothesized that previously undescribed phorbol ester regulated genes are involved in the growth-dynamics of pancreatic cancer. Using oligonucleotide microarray, we generated a list of genes that are differentially expressed following treatment in pancreatic cancer cells with the phorbol ester TPA. We focused our attention on hypothetical genes that hitherto have not been functionally characterized, in the hope of finding novel proteins that might be useful as a diagnostic or prognostic marker, or as a target for intervention. Using transient transfection as a screening tool, we observed differential growth dynamics of cells transfected with one of these hypothetical genes, and subsequently focused on the structural and functional characterization of this gene, which we have named TPA-induced Trans Membrane Protein (TTMP).

Realtime-PCR analysis using the same samples sets was performed to confirm up-regulation of TTMP with TPA stimulation seen on microarray. Induction of the gene was also noted on realtime-PCR to be fairly rapid following TPA treatment and was concentration dependent. Full length transcript of the gene was cloned and the sequence has been deposited in NCBI Genbank (AY830714). Using computational analysis, the amino acid sequence conformed to a single-pass transmembrane topology, and comparison to its orthologues in mouse and chicken was made. We then investigated the mechanism of induction of this gene following exposure to TPA. Pretreatment with actinomycin D did not change degradation kinetics of the message upon induction with TPA. Using a reporter gene luciferase assay, the mode of induction was seen to be at the promoter level.

TTMP is widely expressed and has a high level of expression in normal pancreas but is minimally expressed in the cancer cell lines HeLa and CD18. Deglycosylation assays showed that the protein undergoes post-translational modification by N-glycosylation and addition of sialic acid moieties. Confocal immunofluorescence microscopy demonstrated that TTMP is localized to the endoplasmic reticulum and that this localization process is dependent on the transmembrane domain. TTMP inhibited CD18 pancreatic cancer cell proliferation. siRNA duplexes knocked-down TTMP expression and this led to an increase in cell proliferation, as did clones stably expressing an in-frame N-terminal truncation of TTMP. Cell cycle analysis showed that forced expression of TTMP induced a G1 phase arrest in CD18 pancreatic cancer cells. Forced expression of TTMP was also noted to inhibit proliferation in HeLa cervical cancer cells.

Lastly, basal activity of the promoter region of this gene was characterized. Using deletion constructs of the promoter cloned into the luciferase reporter vector, the core promoter region was identified. Further mutational analysis of the core promoter region showed that 2 putative Sp1 binding sites were responsible for basal activity of the gene. Physical interaction of Sp1 proteins to these sites was demonstrated using gel-shift assays.

In conclusion, we have identified and characterized a novel gene that potentially plays a role in pancreatic tumorigenesis.

LIST OF TABLES

<u>TABLE</u>		<u>PAGE</u>
I.	Genes differentially expressed after 8 hours of TPA treatment.....	54
II.	Exon-intron structure of TTMP.....	60

LIST OF FIGURES

<u>FIGURE</u>		<u>PAGE</u>
1.	Progression model for pancreatic cancer	8
2.	Core promoter elements	26
3.	Time course of H ³ -Thymidine incorporation assay in CD18 cells following treatment with TPA	51
4.	Differential growth dynamics at 72 hours following transient transfection with AK026829-ORFpcDNA3.1.....	53
5.	Concentration response and time course following TPA treatment for CD18 and HeLa cells	55
6.	The transcription start sites of TTMP	56
7.	Nucleotide sequence and deduced amino acid sequence of TTMP	59
8.	The deduced membrane topology of TTMP	61
9.	Alignment of the amino acid sequences of human TTMP with mouse and chicken orthologues.....	62
10.	Induction of TTMP mRNA expression in CD18 cells	64
11.	Expression profile of TTMP in different normal tissues and cancer cells ...	66
12.	Genomic organization and open reading frame of TTMP, and TTMP expression constructs	69
13.	Molecular size of TTMP	70
14.	Prediction of N-glycosylation of TTMP	72
15.	Glycosylation pattern of the TTMP protein	73
16.	Immunofluorescence localization of TTMP in HeLa cells	76
17.	Effect of forced expression of TTMP on cell proliferation in CD18 pancreatic cancer cells	77
18.	Effect of forced expression of TTMP on cell proliferation of CD18 cells in three-dimensional collagen 1 gels	77
19.	Effects of forced expression of TTMP and siRNA duplexes targeted to TTMP on cell proliferation in CD18 pancreatic cancer cells	78
20.	Effect of forced expression of the C-terminal fragment of TTMP on cell proliferation in CD18 pancreatic cancer cells	79

LIST OF FIGURES (continued)

<u>FIGURE</u>		<u>PAGE</u>
21.	Forced expression of TTMP causes G0/G1 phase cell cycle arrest in CD18 pancreatic cancer cells	81
22.	Effect of TTMP on HeLa cell proliferation	81
23.	Sequence of the 5' flanking region of the hTTMP gene	85
24.	Deletion analysis of the 5' flanking region of the hTTMP gene	86
25.	Mutational analysis of the proximal promoter region of the hTTMP gene..	87
26.	Electrophoretic mobility shift analysis of nuclear protein interactions with DNA fragments derived from the hTTMP proximal promoter.....	89

1. BACKGROUND

1.1 INTRODUCTION TO PANCREATIC CANCER

1.1.1 The Pancreas

The human pancreas measures 15-25 cm in length and weighs 70-150 grams. It is connected to the duodenum by the ampulla of Vater, where the main pancreatic duct joins with the common bile duct. The pancreas has its embryological origin as two buds developing on the dorsal and the ventral side of the duodenum. The ventral and dorsal buds fuse together to form the single organ. The terms head, body and tail are used to designate regions of the organ from proximal to distal. The pancreas is an organ with two physiological functions. The acinar and ductal portions of the organ contribute to the exocrine function whereas the islets of Langerhans provide the endocrine function of the pancreas. The acinar and ductal cells secrete enzymes and sodium bicarbonate into the digestive tract respectively. Acinar cells are pyramidal in shape with basal nuclei, regular arrays of rough endoplasmic reticulum, a prominent Golgi complex and numerous zymogen granules containing the digestive enzymes. Proteases and phospholipase originating from the acinar cells are secreted as inactive precursors whereas amylases and nucleases are secreted as active enzymes. The inactive precursors become active only in the duodenum. The pancreatic ducts, which secrete a fluid rich in bicarbonate, are lined with columnar epithelial cells. Secretion of the pancreatic juice is regulated by hormonal stimulation, principally by secretin and CCK, although neural input is also involved. The islets of Langerhans are compact spheroidal clusters embedded in the exocrine tissue. These islets are responsible for secretion of insulin, glucagon, somatostatin and other peptide hormones. In addition to glandular components, the pancreas has a rich blood supply. The arterial blood passes through each lobule, first to the islets and then to the adjacent acini. Various growth factors expressed in the

developing pancreas and its surrounding mesenchyme-derived cells are considered to be involved in the development of the endocrine and exocrine cells (1-5). Members of the EGF family of growth factors such as EGF, TGF α and betacellulin can bind to EGF receptors expressed on pancreatic islet cells, acinar cells, and ductal cells and exert various effects on cell differentiation and proliferation. Other growth factors such IGF-I and PDGF play a role in pancreatic development (1-5).

1.1.2 Cancer of the Pancreas

The vast majority of cases of pancreatic cancer are adenocarcinomas arising from the pancreatic ducts (6-8). The typical histomorphology of ductal adenocarcinoma is one of small neoplastic glands surrounded by an intense desmoplastic stromal reaction, together with inflammatory cells. Rare tumors arise from pancreatic acinar tissue or from neuroendocrine cells in the islets of Langerhans. These tumors tend to have a much different biologic behavior than usual ductal pancreatic adenocarcinoma and are not discussed further here. Cystic tumors of the pancreas (both mucinous and serous) also occur and in their pure form have a substantially better prognosis than ductal carcinomas. These cystic lesions are also excluded from this discussion.

On the scale of public attention, pancreatic cancer ranks far below breast cancer or prostate cancer. One reason is that it affects far fewer people - tens rather than hundreds of thousands in the United States, but no other cancer is as aggressive (7-10). About two-thirds of all pancreatic cancers have already metastasized by the time they are diagnosed, making curative surgery an option for just 1 in 6 patients. The success of this procedure, however, is about 1 in 500 (6-10). The five year survival rate is less than 1% and over 90 % of patients die within one year of diagnosis. In spite of considerable progress in understanding normal pancreatic physiology, the factors that regulate

pancreatic cancer cell proliferation and the reasons for the aggressiveness of this cancer are poorly understood. The only treatment shown to have any effect is surgical resection (7). Thus the impetus is to understand the molecular mechanisms of pancreatic cancer and for the search of novel molecular targets for cancer prevention, diagnosis and treatment.

1.1.3 Epidemiology of Pancreatic Cancer

Though not amongst the top 10 cancers in Singapore, the incidence of pancreatic cancer has nonetheless steadily risen in the last 35 years. In a 5 year period spanning 1998 to 2002, there is an increase of 9% in males and 16% in females as compared to the preceding 5 years (11). The aetiology of pancreatic adenocarcinoma remains poorly defined, although important clues of disease pathogenesis have emerged from epidemiological and genetic studies. Pancreatic adenocarcinoma is a disease that is associated with advancing age (12). It is rare before the age of 40, and culminates in a 40 fold increased risk by the age of 80. Environmental factors, in particular smoking, might modulate pancreatic adenocarcinoma risk (12). On the genetic level, numerous studies have documented an increased risk in relatives of pancreatic adenocarcinoma patients (approximately threefold), and it is estimated that 10% of pancreatic cancers are due to an inherited predisposition (13). As with most cancer types, important insights have emerged from the study of rare kindreds that show an increased incidence of pancreatic adenocarcinoma. However, unlike familial cancer syndromes for breast, colon and melanoma, pancreatic adenocarcinoma that is linked to a familial setting has a lower penetrance (<10%) and maintains a comparable age of onset to sporadic cases in the general population. Among the genetic lesions that are linked to familial pancreatic adenocarcinoma are germline mutations in CDKN2A/p16 (which encodes the tumour suppressors INK4A and ARF), BRCA2, LKB1 and MLH1 (14). The low penetrance of

pancreatic adenocarcinoma that is associated with these germline mutations might point to a role in the malignant progression of precursor lesions rather than in the limiting events that control initiation of neoplastic growth from normal pancreatic cells. With respect to CDKN2A and BRCA2, this notion gains experimental support from the observation that inactivation of these genes is not detected in premalignant ductal lesions that are thought to represent early stages of pancreatic tumorigenesis. Beyond the classical tumour-suppressor mutations, additional genetic defects seem to be operative in rare families in which pancreatic cancer is inherited as an autosomal-dominant trait with very high penetrance (13). A pancreatic cancer syndrome that has so far been identified in a single family has been linked to chromosome 4q32-34 and is associated with diabetes, pancreatic exocrine insufficiency and pancreatic adenocarcinoma, with a penetrance approaching 100% (15). Patients with hereditary pancreatitis, which is associated with germline mutation in the cationic trypsinogen gene PRSS1, experience a 53-fold increased incidence of pancreatic adenocarcinoma (16,17). Mutations in PRSS1 cause the encoded enzyme either to be more effectively autoactivated or to resist inactivation and, consequently, to display deregulated proteolytic activity. It is assumed that the resulting inflammation promotes tumorigenesis, in part by producing growth factors, cytokines and reactive oxygen species (ROS), thereby inducing cell proliferation, disrupted cell differentiation and selecting for oncogenic mutations.

1.1.4 Molecular Genetics of Pancreatic Adenocarcinoma

Pancreatic carcinogenesis is a multistep process accompanied by accumulation of many genetic alterations (18). Irreversible genetic changes occur in the initiation and progression stages of carcinogenesis, while aberrant expression of other genes accompanies the promotional stage of tumor formation (18). Genetic alterations will be

selected in a carcinoma only if these mutations provide the tumor with a selective growth advantage over its neighboring cells, allows a particular cell to evolve into a separate clonal population of tumor cells (18,19). This growth advantage is the phenotypic reflection of changes in the biological pathways in which the protein products of the mutated genes normally participate (18,20). The activation of K-ras appears to be a virtual prerequisite for the development of pancreatic carcinoma and alterations in both K-ras and p16 is an extremely uncommon combination among other human tumor types (21). Severe chromosomal alterations such as deletions, translocations and gene amplification are found in pancreatic cancers (22). This has resulted in a large number of aberrant genes including mutated K-ras and p53 (20-22). A molecular and pathological analysis of evolving pancreatic adenocarcinoma has revealed a characteristic pattern of genetic lesions. The challenge now is to understand how these signature genetic lesions – mutations of KRAS, CDKN2A, TP53, BRCA2 and SMAD4/DPC4 – contribute to the biological characteristics and evolution of the disease. The progression model for colorectal cancer has served as a template for relating sequential, defined mutations to increasingly atypical growth states (23). Whether pancreatic adenocarcinoma behaves in such a progression series has become an active area of research and certainly answers will follow in time.

The pancreatic-duct cell is generally believed to be the progenitor of pancreatic adenocarcinoma. As defined in the landmark study by Cubilla and Fitzgerald (24), the increased incidence of abnormal ductal structures (now designated pancreatic intraepithelial neoplasia, PanIN) (25) in patients with pancreatic adenocarcinoma, and the similar spatial distribution of such lesions to malignant tumours, are consistent with the hypothesis that such lesions might represent incipient pancreatic adenocarcinoma. Histologically, PanINs show a spectrum of divergent morphological alterations relative to

normal ducts that seem to represent graded stages of increasingly dysplastic growth (26). Cell proliferation rates increase with advancing PanIN stages, which is consistent with the idea that these are progressive lesions (25). A growing number of studies have identified common mutational profiles in simultaneous lesions, providing supportive evidence of the relationship between PanINs and the pathogenesis of pancreatic adenocarcinoma. Specifically, common mutation patterns in PanIN and associated adenocarcinomas have been reported for KRAS and for CDKN2A (27). In addition, similar patterns of loss of heterozygosity (LOH) at chromosomes 9q, 17p and 18q (harbouring CDKN2A, TP53 and SMAD4 respectively) have been detected in coincident lesions. Furthermore, studies have consistently shown an increasing number of gene alterations in higher-grade PanINs (28-31).

KRAS. The earliest ductal lesions do not usually display genetic alterations. Activating KRAS mutations are the first genetic changes that are detected in the progression series, occurring occasionally in histologically normal pancreas and in about 30% of lesions that show the earliest stages of histological disturbance (32). KRAS mutations increase in frequency with disease progression, and are found in nearly 100% of pancreatic adenocarcinomas; they seem to be a virtual rite of passage for this malignancy (33). WAF1 (p21/CIP1) seems to be coordinately induced with the onset of KRAS mutations, perhaps due to activation of the mitogen-activated protein kinase (MAPK) pathway (34).

Activating mutations of RAS-family oncogenes produce a remarkable array of cellular effects, including the induction of proliferation, survival and invasion through the stimulation of several effector pathways (35). Although the roles of specific KRAS effector pathways in pancreatic cancer pathogenesis have not been resolved, there is

evidence for an important contribution of autocrine epidermal growth-factor (EGF) signaling (36-40). This autocrine loop and the resulting stimulation of the phosphatidylinositol 3-kinase (PI3K) pathway is required for transformation of several cell lineages by RAS-family oncogenes (41). Consistent with the existence of such an autocrine loop, pancreatic adenocarcinomas overexpress EGF-family ligands (such as transforming growth factor- α (TGF- α) and EGF) and receptors (EGFR, ERBB2 or Her2/neu, and ERBB3) (36,38,41). EGFR and ERBB2 induction occurs in low-grade PanINs, indicating that autocrine EGF-family signaling might be operative at the earliest stages of pancreatic neoplasia (42). The functional importance of this pathway is illustrated by the growth inhibition of pancreatic adenocarcinoma cell lines in vitro and in xenografts following attenuation of EGFR signaling by blocking antibodies or expression of dominant-negative EGFR alleles (39,40,43).

CDKN2A/p16. Germline mutations in the CDKN2A tumour-suppressor gene are associated with the familial atypical mole-malignant melanoma syndrome. In addition to a very high incidence of melanoma, the inheritance of mutant CDKN2A alleles confers a 13 fold increased risk of pancreatic cancer (44). Although pancreatic adenocarcinoma arises in some but not all kindreds with CDKN2A mutations, there are no clear genotype-phenotype associations, indicating a modulating role for environmental factors in disease penetrance (45,46). FAMMM kindreds that harbour mutant loci other than CDKN2A such as cyclin dependent kinase 4 (CDK4) alleles that abrogate INK4A binding or other uncharacterized loci, do not have increased incidence of pancreatic adenocarcinoma (47,48).

Loss of CDKN2A function brought about by mutation, deletion or promoter hypermethylation also occurs in 80-95% of sporadic pancreatic adenocarcinomas (33).

CDKN2A loss is generally seen in moderately advanced lesions that show features of dysplasia (Fig 1). The dissection of the role of CDKN2A has been a fascinating story as this tumour-suppressor locus at 9q21 encodes two tumour suppressors – INK4A and ARF – via distinct first exons and alternative reading frames in shared downstream exons (49). Given this physical juxtaposition and frequent homozygous deletion of 9p21 (in ~ 40% of tumours), many pancreatic cancers sustain loss of both the INK4A and ARF transcripts, thereby disrupting both the retinoblastoma (RB) and p53 tumour suppression pathways. INK4A inhibits CDK4/ CDK6-mediated phosphorylation of RB, thereby blocking entry into the S (DNA synthesis) phase of the cell cycle. ARF stabilizes p53 by inhibiting its MDM2-dependent proteolysis. INK4A seems to be the more important pancreatic cancer suppressor at this locus, as germline and sporadic mutations have been identified that target INK4A, but spare ARF (33,50,51).

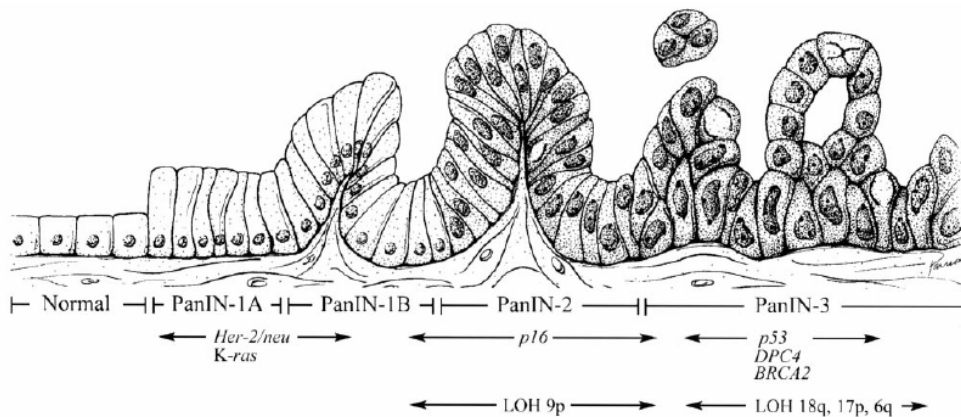


Figure 1. Progression model for pancreatic cancer. The progression from histologically normal epithelium to low-grade PanIN to high grade PanIN is associated with the accumulation of specific genetic alterations. (Adapted from Cancer Res 2000, 60:2002-2006)

TP53. The TP53 tumour suppressor gene is mutated, generally by missense alterations of the DNA binding domain, in more than 50% of pancreatic adenocarcinomas (33). TP53 mutations arise in later stage PanINs that have acquired

significant features of dysplasia, reflecting the function of TP53 in preventing malignant progression. In contrast to many other cancer types, there does not seem to be a reciprocal relationship in the loss of CDKN2A and TP53 (34,52), which points to non-overlapping functions of ARF and p53 in pancreatic cancer suppression. TP53 loss probably facilitates the rampant genetic instability that characterizes this malignancy. These tumours have profound aneuploidy and complex cytogenetic rearrangements, as well as intratumoural heterogeneity, which is consistent with the ongoing genomic rearrangements (53,54).

Cytogenetic studies have provided evidence that telomere dynamics might contribute to this genomic instability. Although reactivation of telomerase is crucial to the emergence of immortal cancer cells, a preceding and transient period of telomere shortening and dysfunction might also contribute to carcinogenesis by leading to the formation of chromosomal rearrangements through breakage-fusion-bridge cycles (55,56). The survival of cells with critically short telomeres (crisis), which continue to go through breakage-fusion-bridge events, is enhanced by the inactivation of p53-dependent DNA damage response (57), allowing the acquisition of oncogenic chromosomal alterations (58). Studies in the telomerase-knockout mouse support this model, as telomere dysfunction and p53 loss cooperate to promote the development of carcinomas in multiple tissues (56). An analysis of a large series of human pancreatic cancer cell lines revealed that telomeres were frequently lost from chromosome ends and that anaphase bridging occurred, indicating that persistent genomic instability is associated with critically short telomeres (58). As these features were observed in both low and high grade tumours, the authors conclude that telomere dysfunction was an early step in the pathogenic process. Moreover, studies of pancreatic adenocarcinomas

revealed that tumours have shortened telomere length and that the activation of telomerase is a late event (58-60).

BRCA2. Inherited BRCA2 mutations are typically associated with familial breast and ovarian cancer syndrome, but also carry a significant risk for the development of pancreatic cancer. Approximately 17% of pancreatic cancers that occur in a familial setting harbour mutations in this gene (61). As is the case for those individuals with germline CDKN2A mutations, the penetrance of pancreatic adenocarcinoma in BRCA2 mutation carriers is relatively low (~7%) and the age of onset is similar to that of patients with the sporadic form of the disease. Familial breast cancer alleles other than BRCA2 do not seem to predispose to pancreatic adenocarcinoma. Loss of wild-type BRCA2 seems to be a late event in those individuals who inherit germline heterozygous mutations of BRCA2, which is restricted to severely dysplastic PanINs and adenocarcinomas (61). Although the numbers are small, these patients do not show an elevated incidence of PanINs. These data are consistent with the possibility that BRCA2 loss promotes the malignant progression of existing lesions in pancreatic neoplasia. BRCA2 is necessary for the maintenance of genomic stability by regulating the homologous recombination based DNA repair processes. Consequently, BRCA2 deficiency in normal cells results in the accumulation of lethal chromosomal aberrations (62). The fact that BRCA2 is selectively mutated late in tumorigenesis probably reflects the need for DNA damage response pathways to be inactivated first, for example by TP53 mutation, so that the damage occurred can be tolerated.

Chromosomal instability. Defects in the mitotic spindle apparatus conferred by centrosome abnormalities might also contribute to the aneuploidy and genomic instability of pancreatic adenocarcinomas. Centrosome abnormalities are detected in 85% of

pancreatic adenocarcinomas, and there is a correlation between levels of such abnormalities and the degree of chromosomal aberrations (63). Overall, the loss of TP53 and BRCA2, and the detection of abnormal mitosis and severe nuclear abnormalities in PanIN-3 lesions indicate that genomic instability is initiated at this stage of tumour progression.

These observations have several implications. First, the detection of clonal genetic alterations in PanINs and the synchronous adenocarcinomas is consistent with the concept that PanINs are indeed neoplastic growths that are precursors to adenocarcinomas. Although KRAS mutations are early, and probably necessary event in the development of pancreatic adenocarcinoma, their absence in the earliest lesions indicates that KRAS activation is not responsible for neoplastic initiation. This notion is supported by the observation of different KRAS mutation between PanINs of the same individual. One possibility is that the earliest lesions might be non-clonal areas of aberrant proliferation and altered states of differentiation that are associated with the replacement of damaged cells and with inflammatory processes. These disruptions in tissue architecture and induction of cell proliferation could produce a field defect in which there is significant selection for cells that sustain activating KRAS mutations. Along these lines, inflammatory stimuli promote the expression of both TGF- α and EGFR in the pancreatic ducts, providing a pathway that could synergize with activated KRAS (37).

In addition to the extreme aneuploidy of pancreatic adenocarcinomas, there is a high degree of genetic heterogeneity within these tumours. For instance, different KRAS mutations and 9q, 17p and 18q LOH patterns have been observed in adjacent PanINs, and several KRAS mutations have been detected in the same adenocarcinomas (27-29). Importantly, it seems that there is spatial distribution of genetic heterogeneity (28).

Neoplastic foci from adjacent regions tend to show similar mutation patterns, whereas increasing genetic divergence has been documented in more geographically distant foci. It seems likely that adenocarcinomas can develop from the clonal progression of one of several related but divergent lesions. These features might indicate that a key event beyond the initiation of PanINs is the acquisition of a mutated state that allows initiated cells to acquire progression associated genetic lesions. It is tempting to speculate that this tremendous degree of heterogeneity and ongoing instability lies at the heart of the resistance of pancreatic tumours to chemotherapy and radiotherapy.

The marked chromosomal abnormalities and the disruptions in DNA-repair processes in pancreatic adenocarcinoma might reflect the existence of additional loci, the genomic alterations of which contribute to the malignant progression. This is supported by the detection of recurrent chromosomal amplifications and deletions by comparative genomic hybridization (CGH) and other cytogenetic methods (54,64). In addition to the signature losses of 17p, 9p and 18q, deletions of chromosomes 8p, 6q and 4q, and amplifications of chromosomes 8q, 3q, 20q and 7p have been consistently reported.

Microsatellite instability. Microsatellite instability is a second mode of genomic instability that, in contrast to the large scale alterations that are associated with chromosomal instability, is characterized by very high mutation rates at small DNA repeat sequences. This phenotype is caused by mutations in DNA mismatch repair genes, including MLH1, MSH2 and MSH6 and is associated with hereditary non-polyposis colon cancer (HNPCC) syndrome (65). There seems to be an elevated risk of pancreatic cancer in HNPCC families (66,67). The pancreatic adenocarcinomas in HNPCC patients show distinct molecular genetic profiles, such as a lower rate of KRAS

and TP53 mutation, frameshift mutations in BAX and TGF β II, characteristic histopathology and a less-aggressive clinical course compared with pancreatic adenocarcinomas that occur outside of this syndrome (68-70).

SMAD4/DPC4. Another frequent alteration in pancreatic adenocarcinoma is the loss of SMAD4/DPC4 (71), which encodes a transcriptional regulator that is a keystone component in the transforming growth factor- β (TGF- β) family signaling cascade (72). This gene maps to 18q21, a region that sustains deletion in approximately 30% of pancreatic cancers (71). The pathogenic role of SMAD4 inactivation is strongly supported by the identification of inactivating intragenic lesions of SMAD4 in a subset of tumours. SMAD4 seems to be a progression allele for pancreatic adenocarcinoma, as its loss occurs only in later stage PanINs (29,30). Moreover, there does not seem to be a strong predisposition to pancreatic adenocarcinoma in patients that inherit a germline SMAD4 mutation (that is, in juvenile polyposis syndrome patients). Loss of SMAD4 is a predictor of decreased survival in pancreatic adenocarcinoma (31), which is consistent with a role for it in disease progression. The mechanism by which SMAD4 loss contributes to tumorigenesis is likely to involve its role in TGF- β mediated growth inhibition. TGF- β inhibits the growth of most normal epithelial cells by either blocking the G1-S cell cycle transition or by promoting apoptosis (72). The cellular responses to TGF- β are partially, but not exclusively, SMAD4-dependent (73) and correspondingly, pancreatic adenocarcinomas show a degree of TGF- β resistance. The roles of TGF β signaling in pancreatic adenocarcinoma pathogenesis are not well defined. Studies have shown inconsistent effects of this cytokine on cultured cell lines with respect to cell proliferation rates and dependency on SMAD4 status for TGF β responsiveness (74-77).

These ambiguous results probably stem from the heterogeneity that is associated with cancer cell lines and the non-physiological conditions that are encountered in vitro.

LKB1/STK11. The Peutz-Jeghers syndrome (PJS), which is caused by LKB1/STK11 mutations is another familial cancer syndrome that is associated with an increased incidence of pancreatic adenocarcinoma (78). PJS patients are primarily affected with benign intestinal polyposis at a young age, although advancing age carries an increased risk of developing gastrointestinal malignancies, including a more than 40 fold increase in pancreatic adenocarcinoma (79).

1.2 ANALYZING DIFFERENTIAL GENE EXPRESSION IN CANCER

Genomics research has transformed molecular biology from a data-poor to a data-rich science. The draft sequences of the human genome were published in 2001 (80,81), with further refinement addressing the shortcomings of these drafts and leading towards the goal of a complete human sequence reported just recently published (82). Current estimates from gene-prediction programs suggest that there are 24,500 or fewer protein-coding genes. Researchers at the International Human Genome Sequencing Consortium have confirmed the existence of 19,599 protein-coding genes in the human genome, and identified another 2,188 DNA segments that are predicted to be protein-coding genes (Human Genome Project Information, http://www.ornl.gov/sci/techresources/human_genome/faq/genenumber.shtml). Of these, only about a third to half has been functionally characterized. Although classical genetics has been a powerful tool for dissecting molecular disease that are affected by gain or loss of function of a protein encoded by a single gene, such a strategy has proved to be less fruitful for understanding diseases such as cancer that are controlled by many genes. Adding to the complexity is the fact that many of the so-called oncogenes or

tumour suppressor genes are signaling molecules themselves, each of which functions to control the expression of a subset of downstream genes. So, the analysis of differential gene expression, known as expression genetics or functional genomics, has become one of the most widely used strategies for the discovery and understanding of the molecular circuitry underlying cancer.

Over the past two decades, several methods have been developed to allow comparative studies of gene expression between normal and cancer cells. Starting with simple approaches that used gel electrophoresis to compare protein expression, methods that focused on mRNA analysis have evolved and become increasingly sophisticated, as a result of the inventions of recombinant DNA, DNA sequencing and PCR technologies. The principles behind some of the main methods can be grouped as follows.

1.2.1 Protein Gel Electrophoresis and Modern Day Proteomics

Perhaps the earliest and arguably the most successful example of studying differential gene expression in cancer was the discovery of the p53 tumour suppressor protein in the late 1970s. The protein was found to be overexpressed on a one-dimensional protein gel when normal cells were compared with those that were infected with simian virus 40 (SV40) DNA tumour virus (83). The later development of two-dimensional (2D) protein gel electrophoresis, which separates proteins by both size and charge, allowed a more complete visualization of cellular protein expression (84). The main shortcoming of these methods is the inability to recover sufficient amounts of the differentially expressed protein species for further molecular characterization leading to identification of just 2000 of the estimated 10000 or more different proteins expressed in a cell. Newer techniques for the analysis of protein expression, collectively known as

proteomics, have been developed in recent years. This involves mainly the use of mass spectrometry to greatly improve the sensitivity and allow characterization of small quantities of protein (85), as well as the use of protein biochips to analyze differential profiling of proteins in a fashion analogous to the array-based format of DNA microarrays (86).

1.2.2 Differential Hybridization

Due to the advent of recombinant DNA technology in the late 1970s, studies of comparative gene expression has shifted from looking at proteins to the analysis of mRNA expression using complementary DNA. The earliest approach was differential hybridization, in which the pair of mRNA samples to be compared were radioactively labeled as cDNA probes with ^{32}P by reverse transcription with oligo-dT primers that anneal to the polyadenylic chains (polyA tails) present at the 3' termini of all eukaryotic mRNAs. The resulting two cDNA probes were then differentially hybridized to duplicate filters, which had on them tens of thousands of plaques from a phage cDNA library (87). Comparison of the hybridization pattern to cDNA-containing phage plaques between two mRNA probes allowed the identification of genes that were uniquely expressed in one but not the other RNA sample. This strategy has implicated several differentially expressed genes that are involved in the hormone responsiveness of human breast cancer cells (88) and that are overexpressed during infection by human T-cell leukaemia/ lymphoma virus (89). However, it was soon realized that such a 'reverse northern' approach of using complex cDNA probes would not be able to detect most genes which are expressed at a low level (87). As a result, differential screening quickly gave way to hybridization methods that use cDNA probes with reduced complexity after a 'subtraction process'.

1.2.3 Subtractive Hybridization

In the early 1980s, an ingenious approach known as subtractive hybridization was devised to enrich for cDNA probes which represent mRNAs that are uniquely expressed in one cell but not the other (90). This method removes most of the cDNAs that represented the genes that are commonly expressed in both cells being compared, and left behind only single-stranded cDNAs that represented a few differentially expressed genes. Well known and important examples utilizing this technique includes the discovery of T-cell receptors (91) and the identification of cyclin dependent kinase inhibitor WAF1 (also known as p21) as a target gene of p53 by Bert Vogelstein and colleagues (92). Since then, several PCR-based subtractive hybridization strategies have been developed, including representational difference analysis (RDA) and suppression PCR, which allow a smaller amount of mRNA samples to be analyzed.

1.2.4 Differential Display

A sensitive method was required so that it could be applied to systems in which scarce biological samples are available, and by which all mRNAs whether scarce or abundant, can be represented. Also the method needed to be systematic, so that a complete search of all the expressed genes in a cell was possible. Based on these crucial criteria, differential display was developed by integrating PCR and DNA sequencing by gel electrophoresis, two of the most simple, powerful and commonly used molecular biological methods (93). Differential display works by systematically amplifying the 3' termini of eukaryotic mRNA by reverse transcription-PCR using one of the three anchored oligo-dT primers (that is the run of Ts ending with a C,G or A) in combination with a set of short primers of arbitrary sequences. Based on the finding that each arbitrary primer would recognize its corresponding mRNA targets with a minimum of seven matching bases, mathematical models have been proposed to predict the relation

between the number of arbitrary primers and the coverage of expressed genes in any given eukaryotic cell (94). Unlike microarray, DD does not require any previous knowledge of mRNA or gene sequences, making it an 'open' system that is applicable to any eukaryotic organism. Genes that have been identified by DD include regulated targets of oncogenes such as RAS, v-REL and ERBB (95-98).

1.2.5 Microarrays

cDNA microarrays (99) and oligoarrays (100) are based on the differential hybridization strategy, in which cDNA plaques are replaced with spotted cDNAs or oligos, and radioactive labels are replaced with fluorescent ones. The immense potential of these methods are based on their ability to simultaneously analyze the expression of mRNAs from tens of thousands of genes, which can then be further analyzed using computers, in the hope that gene-expression patterns can be transformed into more easily interpretable biological pathways for the understanding and classification of cancer (101). DNA microarrays have been used to profile gene-expression patterns of almost all of the main cancers, including leukaemia (102), lymphoma (103), adenocarcinoma of the lung (104), breast (105), prostate (106), with the promise to change the way cancer is diagnosed, classified and treated. However, the realization of these potentials will be a considerable challenge, as the different tumour types can often be more striking than their similarities (107,108). One of the greatest advantages of microarrays over other methods is that each spot on a microarray represent a known sequence. So once a signal is detected, the nature of the gene is known. However, the down side of such a benefit is that it also makes array-based methods 'closed' systems that are only able to cover known gene sequences. The inherent complexity of the cDNA probes that are used in differential hybridization strategies remains the root cause of the lack of signal sensitivity and specificity for most low abundance mRNAs (109,110).

Without a doubt, all human genes can eventually be condensed on to a single array, but uncertainty remains as to whether each of these tens of thousands of cDNA probes will hybridize to only their corresponding target template and to nothing else on the chip. Researchers are thus cautious about the accuracy of microarray data, but most studies place the blame on inadequate bioinformatical and statistical tools for 'data mining' (111,112), rather than on the fundamental problem of the complexity of cDNA probes. As with any other method for the analysis of differential gene expression, data from microarray experiments should be considered with caution, unless each time point can be verified by an independent method such as northern-blot analysis.

1.2.6 Expressed Sequence Tags (ESTs) and SAGE

Expressed sequence tags (ESTs) is based on the strategy of a single run of sequencing of the 3' ends of randomly picked cDNA clones from a cDNA library (113). EST sequencing not only resulted in the discovery of many novel genes, but also provided information on the number of times a corresponding cDNA sequence was represented in a cDNA library from either normal or tumour cells. This strategy has resulted in cataloging and banking of cDNA clones by the National Institutes of Health Cancer Genome Anatomy Program, which provides a convenient source of these clones for functional studies of genes that have been identified by methods for comparative analysis of gene expression. However, because of the high cost and labour intensive nature of comprehensive EST sequencing, the method itself is rarely used to directly identify differentially expressed genes.

Unlike EST sequencing whereby cDNA clones were randomly picked from cDNA libraries, SAGE technology measures the level of gene expression based on the frequency of occurrence of the 3' signature SAGE tags of 10-14 bases in length that

might be unique to each transcript (114). Like differential display, SAGE is an 'open' system based gene discovery tool. Due to the minimal sequence information that is required to define an expressed gene or mRNA, a dozen or more SAGE tags from different genes can be obtained and sequenced at one time, thus speeding up the EST counting process. However, because of the same limited sequence requirement, adequate gene assignment using SAGE methods requires an extensive bioinformatics support for meaningful analysis of the expression pattern for a gene of interest. A recent development is a beads based EST sequencing method known as massively parallel signature sequencing (MPSS) which combines signature sequencing with in vitro cloning of millions of templates on separate 5 micron diameter microbeads (115). Individual mRNAs are identified through the generation of a 17-20 base signature sequence which is immediately adjacent to the 3' end of the 3' most Sau3A restriction site in cDNA sequences. MPSS then captures, identifies and analyses expression levels of genes in a sample by counting the number of these individual mRNA molecules that represent each gene.

1.3 BIOLOGY OF PKC AND TPA

The discovery of Protein kinase C (PKC) in 1977 by Nishizuka and co-workers represented a major breakthrough in the signal transduction field (116). PKC has been identified as the cellular receptor for the lipid second messenger diacylglycerol (DAG), and is therefore a key enzyme in the signalling mechanisms by activation of receptors coupled to phospholipase C, which leads to a transient elevation in DAG levels. The phorbol esters and related diterpenes are natural products that have attracted great interest because of their high potency as tumour promoters in the mouse skin. The phorbol esters exert a variety of effects in cells, which include changes in proliferation, malignant transformation, differentiation and cell death. These natural compounds have

proved to be important tools to delineate the signal transduction pathways involved in growth factor actions and oncogene function. It is well established that the phorbol esters activate protein kinase C (PKC). The marked potency of phorbol esters and their stability compared with the second messenger diacylglycerol (DAG) makes these agents the preferred activators of PKC in cell culture and in vivo models (117).

1.3.1 Cell Growth and Tumour Promotion

The complexity of phorbol ester actions is probably related to the presence of multiple phorbol ester/DAG receptors, which include not only PKC isozymes but also other classes of receptors. In most cases, at least five or more isozymes are present in a single cell and have overlapping or opposite functions. The overlap in function may result from a relatively poor selectivity of individual isozymes towards cellular substrates. An example of opposite roles for PKC isozymes in cell growth is illustrated in fibroblast cell lines, in which PKC δ inhibits cell growth and PKC ϵ is growth stimulatory (118). Inoculation of nude mice with cells overexpressing PKC ϵ results in the formation of tumours, suggesting that this nPKC may function as an oncogene (119). Altered patterns of growth signalling by overexpression of other PKC isozymes has also been reported (120).

While it was initially established that phorbol esters are mitogenic through PKC activation, these compounds may also inhibit cell growth or induce apoptosis in several cell types (121-123). PKC δ mediates apoptosis in numerous cell systems in response to phorbol esters or external stimuli (124-126). PKC isozymes operate as regulators of the cell-cycle both during G₁/S progression and G₂/M transition (127). Activation of PKCs by phorbol esters may promote early phases of mitogenesis, as suggested by the

involvement of PKCs in growth factor actions, mitogen-activated protein kinase (MAPK) activation, and expression of early response genes (128). A bimodal regulation of G1 progression has been observed in some cell lines (129). In fact, overexpression of active forms of PKCs (eg PKC η) blocks the normal phosphorylation of the Rb protein in quiescent cultures of NIH 3T3 cells restimulated to enter the cell cycle, and delay progression of cyclin-dependent inhibitors p21^{WAF} and p27^{KIP1} and/or a reduced expression of cyclin E or cyclin A (130-132). Several studies suggested that Cdc2, the kinase involved in G2/M transition, as well as Cdc25 phosphatase are also PKC targets. In fact, phorbol ester treatment of HeLa, melanoma, and U937 myeloid leukaemia cells results in cell cycle arrest in G2/M (133-135). The PKC-mediated signalling pathways regulating cell growth and cell death are under active investigation and appear to be cell-type dependent.

1.3.2 PKCs and Pancreatic Cancer

The effects of TPA and the role of PKC in pancreatic cancer are mixed and contrasting. Screening of expression pattern of PKC isoforms indicated that the expression of PKC μ correlates with the resistance to Fas-mediated apoptosis in different pancreatic cancer cell lines (136). Both classical PKC and novel PKC signalling pathways enhance anchorage independent growth in MiaPaCa-2 pancreatic cancer cells (137). In CCK-responsive pancreatic cancer cells, PKCs could also mediate invasiveness and the production of MMP-9 (138). Overexpression of PKC α in HPAC human pancreatic cancers results in enhanced tumorigenicity and increased proliferation (139,140). In contrast, another study demonstrated inhibition of pancreatic cancer cell growth through p21-mediated G₁ arrest following activation by TPA/PKC α (141). This opposite effect compared to that in the other two studies could be due to the different

duration of stimulation of pancreatic cancer cells by TPA/PKC α . Our own observations concurred with inhibition of pancreatic cancer cell growth following stimulation with TPA and subsequent activation of PKC (142,143). While we have found p21 to be similarly upregulated, it had resulted in a phase G₂/M arrest and not G₁ arrest as noted in the earlier study (144). Evidence also points to a role of members of the protein kinase C family as mediators of resistance towards apoptosis induced by CD95 and TRAIL-receptors in ductal pancreatic adenocarcinoma cells (145). In a follow-up study, forced expression of PKC μ led to a strongly reduced CD95-mediated apoptosis, enhanced cell growth and to a significant increase of telomerase activity (146). The anti-apoptotic proteins c-FLIPL and Survivin were found to be upregulated in conjunction with PKC μ overexpression. Endogenous overexpression of PKC μ was also noted when comparing immunohistochemical data of pancreatic cancer tissue with normal tissue. Another study showed increased expression of the pro-apoptotic protein Bad and TRAIL receptors following activation of conventional PKC isoforms (147). PKC ζ appears to play a role in maintaining motility of pancreatic cancer cells (148). PKC ζ has also been shown to be involved in directing Sp1-dependent VPF (Vascular permeability factor)/ VEGF (Vascular endothelial growth factor) expression in pancreatic cancer cells (149).

1.4 BIOLOGY OF TRANSMEMBRANE/ ER PROTEINS

Although the basic structure of biological membranes is provided by the lipid bilayer, membrane proteins perform most of the specific functions of membranes. It is the proteins that give each type of membrane in the cell its characteristic functional properties. Different membrane proteins are associated with the membranes in different ways. Many extend through the lipid bilayer with part of their mass on either side. Like their lipid neighbours, these transmembrane proteins are amphipathic, having regions

that are hydrophobic and regions that are hydrophilic. Their hydrophobic regions pass through the membrane and interact with the hydrophobic tails of the lipid molecules in the interior of the bilayer, where they are sequestered away from water. Their hydrophilic regions are exposed to water on either side of the membrane. Other membrane proteins are associated with the cytosolic monolayer of the lipid bilayer either by an amphipathic α helix exposed on the surface of the protein, or by one or more covalently attached lipid chains, which can be fatty acid chains or prenyl groups. Yet other membrane proteins are entirely exposed at the external cell surface, being attached to the lipid bilayer only by a covalent linkage (via a specific oligosaccharide) to phosphatidylinositol in the outer lipid monolayer of the plasma membrane.

1.4.1 Orientation and Conformation of the Transmembrane Protein

A transmembrane protein always has a unique orientation in the membrane. This reflects both the asymmetric manner in which it is synthesized and inserted into the lipid bilayer in the ER and the different functions of its cytosolic and noncytosolic domains. These domains are separated by the membrane-spanning segments of the polypeptide chain, which contact the hydrophobic environment of the lipid bilayer and are composed largely of amino acid residues with nonpolar side chains. Because the peptide bonds themselves are polar and because water is absent, all peptide bonds in the bilayer are driven to form hydrogen bonds with one another. The hydrogen bonding between peptide bonds is maximized if the polypeptide chains forms a regular α helix as it crosses the bilayer, and this is how the great majority of the membrane-spanning segments of polypeptide chains are thought to traverse the bilayer.

Because transmembrane proteins are notoriously difficult to crystallize, relatively few have been studied in their entirety by x-ray crystallography. The DNA cloning and sequencing techniques however, have revealed the amino acid sequence of large numbers of transmembrane proteins and it is often possible to predict from an analysis of the protein's sequence which parts of the polypeptide chain extend across the lipid bilayer. Segments containing about 20-30 amino acids with the high degree of hydrophobicity are long enough to span a lipid bilayer as an α helix, and they can often be identified by means of a hydropathy plot.

1.4.2 Protein Glycosylation

The great majority of transmembrane proteins in animal cells are glycosylated. As in glycolipids, the sugar residues are added in the lumen of the ER and edited in the Golgi apparatus. For this reason, the oligosaccharide chains are always present on the non-cytosolic side of the membrane. Another difference between proteins (or parts of proteins) on the two sides of the membrane results from the reducing environment of the cytosol. This environment decreases the likelihood that intrachain or interchain disulfide (S-S) bonds will form between cysteine residues on the cytosolic side of membranes. These intrachain and interchain bonds do form on the non-cytosolic side of membranes, where they can have an important role in stabilizing either the folded structure of the polypeptide chain or its association with other polypeptide chains respectively.

1.5 TRANSCRIPTIONAL REGULATION

1.5.1 Organisation of the Promoter

The core promoter is the minimal stretch of contiguous DNA sequence that is sufficient to direct accurate initiation of transcription by the RNA polymerase machinery (150). The core promoter is the site of action of the RNA polymerase II transcriptional

machinery. Typically, the core promoter encompasses the site of transcription initiation and extends either upstream or down stream for an additional ~35 nt. There are several sequence motifs, including the TATA box, initiator (Inr), TFIIB recognition element (BRE), and downstream core promoter element (DPE), that are commonly found in core promoters (Fig 3). In addition to the core promoter, other cis-acting DNA sequences that regulate RNA polymerase II transcription include the proximal promoter, enhancers, silencers, and boundary/insulator elements (151-153). These elements contain recognition sites for a variety of sequence-specific DNA binding factors that are involved in transcriptional regulation. The proximal promoter is the region in the immediate vicinity of the transcription start site (roughly -250 to +250 nt). Enhancers and silencers can be located many kbp from the transcription start site and act either to activate or to repress transcription. Boundary/ insulator elements appear to prevent the spreading of the activating effects of enhancers or the repressive effects of silencers or heterochromatin.

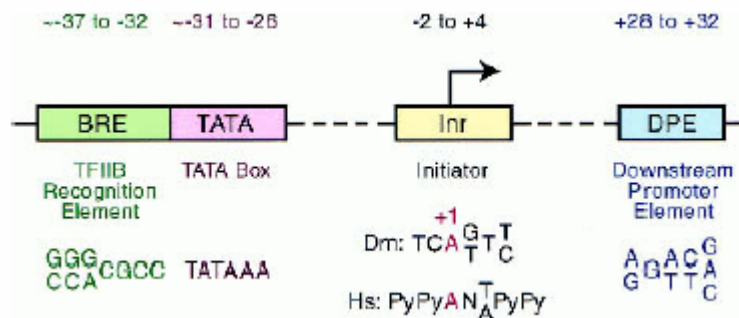


Figure 2. Core promoter elements. Some core promoter motifs that can participate in transcription by RNA polymerase II are depicted. Each of these elements is found in only a subset of core promoters. Any specific core promoter may contain some, all, or none of these motifs. The BRE is an upstream extension of a subset of TATA boxes. The DPE requires an Inr, and is located precisely at +28 to +32 relative to the A₊₁ nucleotide in the Inr. The Inr consensus sequence is shown for both *Drosophila* (DM) and humans (Hs). (Adapted from *Genes and Development* 2002, 16:2583-2592)

1.5.2 RNA Polymerase II Core Promoter Elements

TATA box. The TATA box was the first eukaryotic core promoter motif to be identified (154). In metazoans, the TATA box is typically located about 25-30 nt upstream of the transcription start site, and the consensus sequence for the TATA box is TATAAA. It has been observed, however, that a wide range of sequences function as a TATA box in vivo (155). In humans, it was found that 32% of 1031 potential promoter regions contain a putative TATA box motif (156). TATA box-binding protein (TBP) is the predominant TATA box binding protein. In addition, there are TBP-related factors (TRFs) that are closely related to TBP which also binds the TATA box (157). Transcription factor IID (TFIID) is a multi subunit protein that consists of TBP and approximately 13 TBP-associated factors (TAFs) (158). Accurate and efficient transcription from the core promoter requires the RNA polymerase II along with auxiliary factors that are commonly termed the “basal” or “general” transcription factors, which include transcription factor (TF) IIA, TFIIB, TFIID, TFIIE, TFIIIF and TFIIH. With TATA box dependent core promoters, these factors then assemble into a transcription pre initiation complex (PIC), which guides RNA polymerase II onto the promoter DNA. TFIIH contains a DNA helicase which aids RNA polymerase II to gain access to the template strand at the transcription start point. RNA polymerase II remains at the promoter, synthesizing short lengths of RNA until it undergoes a conformational change and is released to begin transcribing another gene

Initiator (Inr) element. The Inr element encompasses the transcription start site, and was identified in a variety of eukaryotes (159,160). Inr elements are found in both TATA-containing as well as TATA-less core promoters. The consensus for the Inr in mammalian cells is Py-Py(C)-A₊₁-N-T/A-Py-Py (161). The A₊₁ position is designated the

+1 start site because transcription commonly initiates at this nucleotide. More generally, however, transcription initiates at a single site or in a cluster of multiple sites in the vicinity of the Inr (and not necessarily at the A₊₁ position). A variety of factors have been found to interact with the Inr element. There is considerable evidence that TFIID binds to the Inr in a sequence specific manner (162,163). More specifically, it appears that TAF2 and TAF1 are the key subunits of TFIID that interact with the Inr (164,165). Aside from TFIID binding to the Inr, it has been observed that purified RNA polymerase II (or RNA polymerase II along with TBP, TFIIB, TFIIF) is able to recognize the Inr and to mediate transcription in an Inr-dependent manner in the absence of TAFs (166,167). These results suggest that TFIID and RNA polymerase II may recognize and interact with the Inr at different steps in the transcription process.

Downstream core promoter element (DPE). The DPE was identified as a downstream core promoter binding site for purified *Drosophila* TFIID (163). TFIID binds cooperatively to the Inr and DPE motifs, as mutation of either the Inr or the DPE results in loss of TFIID binding to the core promoter. The DPE is found most commonly in TATA-less promoters. With naturally occurring TATA-less core promoters, mutation of the DPE motif results in a 10 to 50 fold reduction in basal transcription activity (163,168,169). Although the DPE has been studied mainly in *Drosophila*, it is also present in humans (168,170). The DPE is located precisely at +28 to +32 relative to the A₊₁ position in the Inr. All of the known DPE containing promoters possess identical spacing between the Inr and DPE motifs, and the alteration of the spacing between the Inr and DPE by a single nucleotide causes a several fold reduction in TFIID binding and basal transcriptional activity (168,169). The consensus sequence for the DPE is estimated to be A/G₊₂₈-G-A/T-C/T-G/A/C. There is also a minor preference for G at +24 (169). Although the DPE consensus sequence is somewhat degenerate, it should be

considered that both DPE and Inr motifs are required in DPE dependent promoters and that the spacing between the DPE and Inr is invariant (which enables the cooperative binding of TFIID to the two motifs). Thus, the functional consensus for DPE-dependent core promoters consists of the Inr and DPE motifs with the DPE located at +28 to +32 relative to A₊₁.

TFIIB recognition element (BRE). The BRE is a TFIIB binding site that is located immediately upstream of some TATA boxes (171). TFIIB is able to bind directly to the BRE in a sequence specific manner. The BRE consensus is G/C-G/C-G/A-C-G-C-C (where the 3'C of the BRE is followed by the 5'T of the TATA box), and at least a 5 out of 7 match with the BRE consensus was found in 12% of a collection of 315 TATA-containing promoters. In vitro transcription experiments with purified basal transcription factors revealed that the BRE facilitates the incorporation of TFIIB into productive transcription initiation complexes. On the other hand, the BRE was observed to have a negative effect on basal transcription by in vitro transcription with a crude extract or by transient transfection analysis (172).

CpG islands. CpG islands, which generally range in size from 0.5 to 2kpb, contain promoters for a wide variety of genes. CpG islands typically lack TATA or DPE core promoter elements, but contain multiple GC box motifs that are bound by Sp1 related transcription factors (173,174). In addition, transcription from CpG islands initiates from multiple weak start sites that are often distributed over a region of about 100nt, which is in contrast to transcription from TATA or DPE dependent core promoters that occurs from a single site or localized cluster (of less than 10nt) of sites. The analysis of 1031 human genes revealed that about half of the potential promoter regions are located in CpG islands (156). From the core promoter perspective, CpG island may

contain multiple weak core promoters rather than a single strong promoter. The presence of Sp1 binding sites in CpG islands is particularly notable. Not only does Sp1 contribute to the maintenance of the hypomethylated state of CpG islands (173,174), but it may also function in concert with the basal transcription factors to mediate transcription initiation. It has been found, for example, that Sp1 binding sites in conjunction with an Inr motif can activate transcription in the absence of a TATA box (157,175). Hence it is possible that CpG island promoters consist of multiple Sp1+Inr pairs that collectively generate the array of start sites that are observed.

1.5.3 Sp1/KLF Family of Transcriptional Factors

Sp1 was the first mammalian transcription factor to be cloned (176). It binds to GC-rich sequences including GC boxes, CACCC boxes (also called GT-boxes) and basic transcription elements collectively termed Sp1 sites. Early studies indicated that Sp1 was responsible for recruiting TATA-binding protein (177,178) and fixing the transcriptional start site at TATA-less promoters(179). These findings together with the fact that 'Sp1-sites' are found in the promoters of many housekeeping genes, led to the widely held notion that Sp1 acts as a basal transcription factor and that Sp1 sites represent constitutive promoter elements that support basal transcription at these promoters. However, early studies also showed that Sp1 was subject to extensive post-translational modification by both glycosylation and phosphorylation (180,181), indicating that its activity was likely to be regulated. The identification of several transcription factors with high homology to Sp1 (182,183) together with the recognition that Sp1 is a part of a large multigene family, further indicated that transcription from Sp1 sites may be more complex than first envisioned. In keeping with this idea, Sp1 sites have been found to be involved in tissue specific gene expression (184,185) and in control of transcription following a number of different stimuli, for example in response to oncogenes (186),

antimetabolites (187), growth stimulation (188,189) and differentiation (190). This demonstrates that Sp1 site dependent transcription can be highly regulated in cells.

Sp1 belongs to a family of transcription factors containing at least twenty identified members in mammals. The family includes the Sp proteins and the Kruppel like factors (KLFs). The diversity of this family is further increased by transcriptional and post-transcriptional mechanisms, such as alternate splicing, promoter usage, or initiation codon usage (191,192). Members of this family are characterized by a highly conserved C-terminal DNA-binding domain containing three zinc fingers. Although conservation within their DNA binding domains means that Sp1 related proteins can interact with the same DNA sequences, different family members do have preferences for different sequences. For example, Sp1, Sp3 and Sp4 bind with higher affinity to GC boxes than to CACCC boxes (183,193,194), whereas many of the KLFs bind preferentially to CACCC boxes over GC boxes. The basic transcription element sequence on the other hand, may represent a promiscuous sequence which binds many family members with more similar affinities (195). Since all Sp/KLF factors appear to bind with varying affinities to GC boxes, CACCC boxes and basic transcription elements, all family members have the potential to affect Sp1 site dependent transcription in cells in which they are expressed.

In contrast to the Sp proteins, which share homology with Sp1 throughout their sequence, the KLFs are only homologous to Sp1 in their C-terminal DNA binding domains. While subsets of the Sp1 related family can be grouped based on sequence conservation (196), their N-termini are not conserved throughout the family and can contain a variety of domain types. As might be expected from this diversity, the family members also differ in their transcriptional activity, with some being activating and some repressive. Additionally, many of these factors have been found to activate or repress

transcription dependent on cell line or promoter examined (194,197). Thus, the transcriptional properties of these proteins can be context-dependent. In keeping with this idea, both activating and repressive domains have been identified in Sp3 and GKLF (198,199).

Although some members of the Sp/KLF family are highly restricted in their tissue distribution (eg EKLF), others are more widespread or ubiquitously expressed (eg Sp1, Sp3 and UKLF, see Table 1). As a result, individual cells appear to express multiple members of the family. This realization, together with data showing that Sp/KLF factors are able to modulate each other's activity at multiple levels (i.e. at the level of expression, promoter interaction and protein-protein interaction), indicates that the Sp/KLF family is likely to make up a network through which transcription can be fine-tuned by changes in the expression profiles of its members (200).

1.6 FUTURE DIRECTIONS

Novel genomic and proteomic technologies for global expression analysis have shown promise in providing a molecular taxonomy. Signature profiles have allowed the improved classification of tumour types and the elucidation of prognostic markers. These methods have recently been used to study pancreatic adenocarcinoma and have revealed potential new diagnostic markers and therapeutic targets (201-203). The identification of recurrent chromosomal amplifications and deletions in pancreatic adenocarcinomas indicates that there are numerous loci involved in the pathogenesis of this malignancy. High resolution gene discovery technologies, coupled with the validation potential of inducible mouse models, should expand the list of essential targets for more productive drug development initiatives. Another important avenue for pancreatic adenocarcinoma gene discovery might be from genetic mapping studies of

pancreatic adenocarcinoma prone kindreds. The genetic lesions in most of these families have yet to be identified. Segregation analysis of a large number of kindreds has indicated that susceptibility might be due to autosomal dominant inheritance of a rare allele(s) (14). The identification of such an allele would be of great potential use for the early identification of patients at risk and in understanding the biology of disease.

2. HYPOTHESIS AND AIMS

The overall hypothesis tested is that previously undescribed PKC or phorbol ester regulated genes are involved in the growth-dynamics and multi-step pathogenesis of pancreatic cancer. The following aims were proposed to test this hypothesis:

- 1) Identify and sequence a previously uncharacterized gene that is differentially expressed in TPA treated pancreatic cancer cells.
- 2) Characterize the expression pattern, sub-cellular localization and function of this gene.
- 3) Clone the promoter of this gene and understand the basal transcriptional regulation of its expression.

3. MATERIALS AND METHODS

3.1 MICROARRAY AND IDENTIFICATION OF NOVEL GENE

3.1.1 Cell Culture

The human pancreatic adenocarcinoma cell line, CD18, human cervical carcinoma cell line, HeLa, and human embryonic kidney cell subclone, TSA201 were cultured in Dulbecco's Modified Eagle Medium (DMEM) (Sigma Chemicals) that was supplemented with 10% fetal bovine serum (Gibco), penicillin (100U/ml), streptomycin (100µg/ml) and amphotericin B (0.25µg/ml) (Cellgro). Cells were grown as monolayers at 37°C in a humidified environment with 95%O₂/5%CO₂. TPA (Sigma) was reconstituted in dimethyl sulfoxide (DMSO) to a final concentration of 1mM stock solution. For preparation of samples for microarray, CD18 cells were cultured in T75 flasks to 70% confluence and washed twice with phosphate buffered saline (PBS) before being grown in serum free media for 16 hours. The cells were then stimulated with 0.1µM TPA in serum-free media for 8 hours. For time-course and concentration-response experiments, both CD18 and HeLa cells were grown in T25 flasks to 70% confluence and subjected to the same media changes as above. Cells were then treated with TPA in various concentrations and for various lengths of time before being harvested for RNA extraction. Paired non-stimulated controls for the microarray as well as the time-course and concentration-response experiments were prepared in parallel and were exposed to equivalent volumes of DMSO as the stimulated samples. For mRNA degradation assays, CD18 cells were pre-treated with 0.1µM TPA for 4 hours before actinomycin D (Calbiochem) was added to a final concentration of 5µg/ml. Subsequently, cells were harvested at various time points for RNA extraction.

3.1.2 RNA Extraction

The cells were washed twice with PBS and RNA was extracted using TRIzol reagent (Invitrogen). To further purify the samples, the lysates were processed with the total RNA miniprep kit (Sigma). The eluted purified RNA samples were then subjected to DNAase I treatment (Ambion), according to the manufacturer's protocol. RNA was quantified by absorption spectroscopy and the quality of the samples was monitored by gel electrophoresis before subsequent use.

3.1.3 Oligonucleotide Array Gene Expression Analysis

Human oligonucleotide probes (60mers) were designed for each target gene (Compugen, manufactured by Sigma-Genosys, Inc). Oligonucleotides (9,861) were resuspended (30 μ M) in 3X saline sodium citrate (SSC) buffer and spotted onto poly-L-lysine coated slides using a MagnaSpotter robot (BioAutomation) with a 12-pin print head (Telechem) configuration in a humidified, HEPA-filtered hood. After spotting, the DNA was cross-linked to the slides by UV irradiation (350mJ/cm²) with a Stratalinker UV Crosslinker (Stratagene), blocked by succinic anhydride treatment, and rinsed in ethanol. The printed arrays were boxed and stored desiccated at room temperature.

Total RNA was isolated as described above. To generate fluorescently-labeled single-stranded cDNA target, 40 μ g of total RNA was incubated with 2 μ g of anchored oligodeoxythymidylate primer in a total volume of 30 μ l at 70°C for 10 min and chilled on ice. To each sample was added 10X first strand buffer (6 μ l), 0.1M DTT (6 μ l), 20X aminoallyl-dNTP mix (3 μ l; 10mM dATP, 10mM dCTP, 10mM dGTP, 6mM dTTP, and 4mM aminoallyl-dUTP), RNase inhibitor (0.5 μ l) (RNasin, Promega), and StrataScript reverse transcriptase (3 μ l; 50U/ μ l). After incubation at 48°C for 60 min, an additional 50 units (1 μ l) of StrataScript was added to the samples and incubated for an additional 60

min. The reaction was stopped by adding 12 μ l of 0.5M EDTA (pH 8). Residual RNA was hydrolyzed by adding 12 μ l of 1M NaOH to the mixture followed by incubation at 65°C for 15 min and cooled to room temperature. The reaction was neutralized with 16.8 μ l of 1N HCl. First-strand cDNA was purified from unincorporated amino allyl-dUTPs on QIAquick PCR purification columns (Qiagen) according to manufacturer's protocols, except that QIAquick wash buffer was replaced with 5 mM potassium phosphate buffer (pH 8.5) containing 80% ethanol, and cDNA was eluted with 4 mM potassium phosphate buffer (pH 8.5) and vacuum-dried. cDNA was resuspended in 10 μ l 0.05 M Na₂CO₃ buffer (pH 9), mixed with either Cy3 or Cy5 monofunctional NHS-ester (Amersham Pharmacia), and incubated for 90 min in the dark at room temperature. Cy3- and Cy5-conjugated cDNA targets were then purified by QIAquick PCR purification columns, combined, vacuum-concentrated, and diluted to 55 μ l with hybridization solution containing final concentrations of 50% formamide, 4.1x Denhardtts, and 4.4x SSC. To reduce nonspecific hybridization, the hybridization solution also contained final concentrations of 15 μ g human Cot1 DNA (Invitrogen), 12 μ g poly-deoxyadenylate, and 5 μ g of yeast tRNA (Sigma). After clarification by centrifugation, the cDNA/hybridization solution was applied to DNA microarrays and incubated at 42°C for 16-20 h. After hybridization, microarray slides were washed by immersion in 2X SSC, 0.1%SDS for 5 min at room temperature, 1X SSC, 0.01% SDS for 5 min at room temperature, and 0.2X SSC for 2 min at room temperature. The microarrays were dried by centrifugation for 5 min at 1000 rpm and scanned immediately with a ScanArray 4000 confocal laser system (Perkin-Elmer). Fluorescent intensities were background subtracted, and normalization and filtering of the data were performed using the QuantArray software package (Perkin-Elmer). After normalization, expression ratios were calculated for each feature.

3.1.4 Reverse Transcription and Real-Time Quantitative PCR

Real-time PCR was performed to validate the microarray results, as well as to quantify transcript levels for the mRNA degradation studies. Reverse transcription was first carried out to generate cDNA using 1µg of RNA with 0.5µg/µl oligoDT primers (Invitrogen), 1mM of each dNTP and 1.25U AMV reverse transcriptase (Promega) in a 20µl reaction volume. Following which, real-time PCR was performed using the cDNA generated and quantitative assessment of DNA amplification was detected via the dye SYBR Green I using the ABI PRISM 7700 Sequence Detector (Applied Biosystems). The following primers were used: TTMP forward: 5' GTCGCTTAGCTGGAGTGCG 3', reverse: 5' CTGCCTCTCCGTGCTCTACC 3'; GAPDH forward: 5' TGGGCTACTGAGCACCAG 3', reverse: 5' TGGGCTACTGAGCACCAG 3'. 1µl of the reverse transcription reaction was used to provide cDNA template for the PCR reaction. This was mixed together with SYBR Green PCR Mastermix (Applied Biosystems) and primers at a concentration of 500nM in a total volume of 20µl. The following cycling parameters were used for the PCR reaction: 50°C for 2 minutes, 95°C for 10 minutes, followed by 40 cycles of 95°C for 15 seconds and 60°C for 1 minute. A final heating step up to 95°C was performed to obtain melting curves of the final PCR products. The fluorescence threshold cycle value (Ct) is obtained for each curve and normalized to that obtained for the GAPDH housekeeping gene in the same sample to normalize for discrepancies in sample loading. The difference in Ct values between treated and control samples were then computed and exponentially multiplied to the base of 2 to obtain relative differences in expression levels. All experiments were carried out in duplicates and independently performed at least twice.

3.1.5 Rapid Amplification of cDNA Ends (RACE)

5' and 3' RACE was performed to identify full length transcripts containing the transcription start site, using colonic RACE ready cDNA (Ambion). 5' RACE nested PCR was performed with upstream RACE primers supplied with the kit and the following gene specific downstream primers: outside: 5' CGACTCAGAGAGGGCGATGT 3', inside: 5' CCTTTCGGTGAGCAGGTGAG 3'. Similarly, 3' RACE was performed using downstream RACE primers and the following set of nested PCR upstream primers: outside: 5' CCTCACAGCCAGTAGACGAGC 3', inside: 5' GCAGCCAGAAGAGAACACGC 3'. PCR conditions consisted of one cycle of 5 minutes at 95°C, followed by 35 cycles of denaturing at 95°C for 30 seconds, annealing at 60°C for 30 seconds and extension at 72°C for 3 minutes. A final extension cycle of 72°C for 6 minutes was performed to ensure complete synthesis of double stranded DNA. PCR products obtained after the second round of PCR was cloned into pGEM T-Easy vector (Promega) and isolates were screened and selected for sequencing analysis

3.1.6 Construction of Plasmid for Promoter Analysis

Human genomic DNA was isolated using a Wizard DNA Purification Kit (Promega). A 2kb fragment of the TTMP 5' flanking region was engineered by unidirectional cloning of PCR product using the following primers into the KpnI and SmaI sites of the promoterless and enhancerless firefly luciferase reporter vector pGL₃-Basic (Promega) (-1909/+95pGL₃). Forward primer (-1909, with KpnI restriction site in bold): 5' **GGGGTACCCCTTCAGAAC** TTATATTCCTTCCACTG 3', reverse primer (+95): 5' TGTTCTCTTCTGGCTGCC G 3'

3.1.7 Transient Transfection

For study of promoter activity of the luciferase constructs, 1.5×10^4 cells were seeded into 48 well plates 16 hours before transfection. Transfection was carried out using Metafectane (Biontex Laboratories GmbH, Munich, Germany) with modifications to the manufacturer's protocol. 250ng of luciferase reporter plasmid with 1 μ l of metafectane was transfected in serum-containing media. The cells were washed with phosphate buffered saline and fresh serum-containing media was changed 6 hours after the start of transfection. 30 hours after the start of transfection, the media was again changed, and cells were grown in serum-free conditions overnight. For basal promoter activity, the cells were then lysed 48 hours from the start of transfection. For the TPA induction studies, TPA was added to the media to a final concentration of 0.1 μ M, and cells were lysed 6 hours after TPA stimulation.

3.1.8 Reporter Gene Assay

Cells were lysed by one freeze and thaw cycle in reporter lysis buffer (Promega). Luciferase activity was measured using the Victor light luminescence counter (Wallac/Perkin-Elmer). All experiments were carried out in triplicates and independently performed at least three times. Data of luciferase activity are normalized to protein concentration, and are shown as mean \pm SEM of three independent experiments.

3.2 EXPRESSION, STRUCTURAL AND FUNCTIONAL CHARACTERIZATION

3.2.1 Cell Culture and Transfection Protocol

The human pancreatic adenocarcinoma cell line, CD18, human cervical carcinoma cell line, HeLa, and human embryonic kidney cell subclone, TSA201, were cultured in Dulbecco's modified Eagle's medium (DMEM) (Sigma Chemicals)

supplemented with 10% fetal bovine serum (Gibco), penicillin (100U/ml), streptomycin (100µg/ml), and amphotericin B (0.25µg/ml) (Cellgro). Cells were grown as monolayers at 37°C in a humidified environment with 95% O₂/5% CO₂. For transient transfection assays, cells were seeded into 6 well plates for 16 hours before transfection. Transfections were carried out using Metafectane (Biontex Laboratories GmbH, Munich, Germany) with modifications to the manufacturer's protocol. The cells were washed with phosphate-buffered saline and fresh serum containing media was changed 6 hours following transfection. Stable expression clones were selected by growing the transiently transfected cells in G418 selection media for 3 weeks.

3.2.2 Real-Time RT-PCR Analysis of mRNA Expression in Human Tissues and Cancer Cells

Multiple human tissue cDNA panels were purchased from BD Clontech (cat# K1420-1 and K-1424-1, Mountain View, CA). The cDNA content from different tissues has been normalized to expression levels of several different housekeeping genes including GAPDH and beta-actin. This normalization ensures an accurate assessment of tissue specificity and relative abundance of the target mRNA. To examine TTMP expression in different human tissues, 1 µl cDNA from each tissue was used for real-time PCR analysis. Quantification of DNA amplification was carried out using the SYBR Green I dye on the ABI PRISM 7700 Sequence Detector (Applied Biosystems, Foster City, CA). The following primers were used: TTMP forward: 5'-GTGCTTAGCTGGAGTGCG-3', reverse: 5'-CTGCCTCTCCGTGCTCTACC-3'; GAPDH forward: 5'-GGGCTACACTGAGCACCCAG-3', reverse: 5'-TGGGCTACACTGAGCACCCAG-3'. Similarly, TTMP mRNA expression in normal pancreas, HeLa and CD18 cells was compared using real-time PCR with GAPDH as an internal control. In brief, RNA was extracted from both HeLa and CD18 cells using Sigma

mRNA extraction kit. 1 µl of RNA was reverse-transcribed into cDNA using MLV reverse transcriptase (Promega, Madison, WI). Normal pancreas cDNA was purchased from BD Clontech as above. Expression of both GAPDH and TTMP was estimated by SYBR Green real-time PCR as described above. TTMP expression was normalized to that of the house keeping gene, GAPDH. The data was analyzed according to our previous publication (204).

3.2.3 Plasmid Construction

The TTMP expression vector (TTMPpcDNA3.1) was constructed based on the open reading frame of TTMP (GenBank Accession No. AY830714). To identify the protein sequences that determine the localization of TTMP, two constructs that express truncated TTMP proteins were also established, one containing the cDNA sequence coding for the C-terminal part (amino acid residues from 68 to 217) of TTMP with transmembrane domain (CT-TTMPpcDNA3.1) and another containing the cDNA sequence coding C-terminal part (amino acid residues from 80 to 217) of TTMP without transmembrane domain (CT^{TMminus}-TTMPpcDNA3.1) (Fig 12C). PCR amplifications of these sequences were performed using the following primers; TTMP forward primer: 5'-**CACCATGGACCTGGCCCAACC**-3', CT-TTMP forward primer: 5'-**CACCATGATCATCACCTCCATTTTCC**-3', CT^{TMminus}-TTMP forward primer: 5'-**CACCATGGTAACTTATGTTGATGAAGATG**-3', reverse primer for all 3 sequences: 5'-**TTCATAGAGCAAGAGGGATG**-3'. The forward primers contained the Kozak sequence (**CACC**) preceding the ATG start codon to facilitate directional cloning into the pcDNA3.1/D/V5-His-TOPO expression vector (Invitrogen, Carlsbad, CA). The reverse primer did not contain the stop codon to enable read-through and transcription of the V5/His sequence on the pcDNA3.1 expression vector, thus generating fusion proteins of TTMP, CT-TTMP or CT^{TMminus}-TTMP with the V5/His-tag. In addition, a

5'UTR/TTMPpcDNA3.1 construct was also generated using the following primers: forward primer: 5'-ATA**AAGCTT**AGACTTTCCCTGCCGGCAC-3' (with HindIII restriction site highlighted in bold), reverse primer: 5'-AT**GGGCCCTC**TTTCATAGAGCAAGAGGGATG-3' (with Apal restriction site highlighted in bold and two extra nucleotides highlighted in bold and italics to preserve reading frame with V5/His-tag). These PCR products with the incorporated restriction sites were cloned into pcDNA3.1D/V5-His-TOPO vector by standard restriction cloning techniques. For all PCR reactions, Platinum Pfx DNA polymerase (Invitrogen, Carlsbad, CA) was used for its high-fidelity proof-reading function, and conditions were as follow: 95°C for 5 minutes, followed by 35 cycles of 95°C for 15 seconds, 59°C for 15 seconds and 72°C for 1 minute, and a final extension step of 72°C for 2 minutes.

3.2.4 Western Blotting

Identification of protein expression as well as selection of stable cell lines was performed using standard western blotting techniques. Briefly, cells were lysed in lysis buffer and then clarified by microcentrifugation. The supernatants were recovered and protein concentrations measured using the Bio-Rad protein assay reagent. Equivalent amounts of cell lysate protein were resolved by SDS-PAGE 15% electrophoresis and transferred to nitrocellulose membranes by electroblotting. The membranes were then blocked (with 5% non-fat milk) before incubation with anti-V5 antibody (Invitrogen, 1:5000) and subsequently HRP conjugated secondary antibody. The membranes were then detected by chemiluminescence and light emission was captured on X-ray film.

3.2.5 Deglycosylation Assay

Deglycosylation was carried out on protein lysates extracted from TSA201 cells transiently transfected with TTMP expression vectors using the Glyko enzymatic

deglycosylation kit (ProZyme, San Leandro, CA). The manufacturer's denaturing protocol was followed. Following deglycosylation, cell lysates were analyzed using standard western blotting techniques as described above.

3.2.6 Immunofluorescence

HeLa cells were plated on coverslips at a density of 2.5×10^5 cells in 6 well plates. After overnight culture, cells were transiently transfected with the full-length TTMP expression plasmid (TTMPpcDNA3.1), the plasmid expressing the C-terminal end of TTMP with the transmembrane domain (CT-TTMPpcDNA3.1) or the plasmid expressing the C-terminal end of TTMP without the transmembrane domain (CT^{TMminus}-TTMPpcDNA3.1). The transfected cells were grown for 24 hours and subsequently washed once with PBS, and fixed with 4% formaldehyde for 20 minutes. They were then washed twice with PBS, permeabilized with 0.2% Triton in PBS for 10 minutes and washed once before blocking with 1% BSA/PBS for 1 hour. These cells were incubated overnight at 4°C with V5 antibody at 1:800 dilution as well as the endoplasmic reticulum-specific protein disulphide isomerase (PDI) antibody at 1:100 dilution (Santa-Cruz, CA). Immunofluorescent secondary antibodies Alexfluor 488-labelled goat anti-mouse IgG₁ and Alexfluor 546-labelled goat anti-rabbit IgG (Molecular Probes, Carlsbad, CA) were used at 1:2000 dilution at room temperature for 1 hour to identify V5 antibody and PDI antibody respectively. The coverslips were then mounted with antifade reagent and visualized using Laser Scanning Confocal Microscopy (Carl Zeiss LSM510, Tingen, Germany).

3.2.7 Cell Proliferation Assay by Cell Counting

Stably transfected CD18 cells and HeLa cells were seeded into 12-well microplates at a density of 2.5×10^4 cells per well. The cells were then grown in serum-

free conditions for a further 24, 48, 72 or 96 hours. At the end of each time period, the cells were trypsinized to produce single cell suspensions and the cells from each well were counted using an automated cell counter (Guava Technologies, Hayward, CA).

3.2.8 siRNA Gene Silencing Assay

A mixture of siRNA duplexes specific to TTMP mRNA was provided by Dharmacon (Lafayette, CO). A control scrambled siRNA duplex was also obtained. To transfect the siRNA duplex, CD18 cells were plated into a 6-well plate and grown in DMEM media containing 10% FBS overnight. The final concentration of siRNA used for transfection was 100 nM. The siRNA duplex was first mixed with Metafectene (Biontex, Munich, Germany) in 1:4 ratio in Opti-MEM media, then the siRNA-Metafectene complex was incubated with CD18 cells for 6 hours before changing to serum containing media. 48 hours following transfection, the cell number in each well was counted and TTMP mRNA expression was measured by real time RT-PCR.

3.2.9 Cell Proliferation in Collagen I Gel

Collagen I gel was purchased from BD Biosciences (rat tail collagen I gel) and stored at 4°C. To grow cells in 3-D collagen gel, 1M NaOH was mixed with collagen I stock solution on ice. Single cell suspensions were mixed with collagen I gel and transferred to 24 well plates. Following incubation for 30 minutes at 37°C, 200 µl of DMEM was added on top of the polymerized collagen I gel. Media was replaced with fresh media every 3 days and colonies were photographed at the end of a 2 week growth period.

3.2.10 Flow Cytometry

CD18 cells with stable expression of TTMPpcDNA3.1 or empty vector were cultured overnight in serum free condition. The cells were then digested with trypsin-EDTA solution to produce a single cell suspension and then fixed with ice-cold ethanol for at least 12 hours at 4°C. The cells were then centrifuged and the cell pellets washed with PBS. The final cell pellets were reconstituted with Telford reagent and shaken horizontally for an hour in the dark at room temperature. The red fluorescence of propidium iodide was recorded by flow cytometry at 488nm excitation λ and 610nm emission λ .

3.3 TRANSCRIPTIONAL REGULATION

3.3.1 Cell Culture and Transient Transfection

The human cervical carcinoma cell line, HeLa were cultured in Dulbecco's Modified Eagle Medium (DMEM) (Sigma Chemicals) that was supplemented with 10% fetal bovine serum (Gibco), penicillin (100U/ml), streptomycin (100 μ g/ml) and amphotericin B (0.25 μ g/ml) (Cellgro). For study of promoter activity of the luciferase constructs, 1.5×10^4 cells were seeded into 48 well plates 16 hours before transfection. Transfection was carried out using Metafectane (Biontex) with modifications to the manufacturer's protocol. 250ng of luciferase reporter plasmid with 1 μ l of metafectane was transfected in serum-containing media. The cells were washed with phosphate buffered saline and fresh serum-containing media was changed 6 hours after the start of transfection. 30 hours after the start of transfection, the media was again changed, and cells were grown in serum-free conditions overnight. For basal promoter activity, the cells were then lysed 48 hours from the start of transfection.

3.3.2 Construction of Plasmids for Promoter Analysis

Human genomic DNA was isolated using a Wizard DNA Purification Kit (Promega). Progressive deletion constructs of the TTMP 5' flanking region was engineered by unidirectional cloning of PCR product using the following primers into the KpnI and SmaI sites, or MluI and BglII sites of the promoterless and enhancerless firefly luciferase reporter vector pGL₃-basic (Promega). The TTMP specific sequences for these primers were taken from the human genomic sequence contig NT_005612.14.

Sense primers (with KpnI restriction site in bold).

(-1909) 5' GGG **GTA CCC** CTT CAG AAC TTA TAT TCC TTC CAC TG 3'

(-766) 5' GGG **GTA CCT** GCT TGC TGT GTC AAT CAC TGT T 3'

(-299) 5' GGG **GTA CCA** CTG TGG CAA ACC CCA ACA A 3'

(-284) 5' GGG **GTA CCC** CTA AAC GGA GAT GGT CTG CA 3'

(-219) 5' GGG **GTA CCA** TCA GAA AAA CAG CCT TCG 3'

Anti-sense primer for PCR amplification with the above primers.

(+95) 5' TGT TCT CTT CTG GCT GCC G 3'

Sense primers (with MluI restriction site in bold).

(-135) 5' ATT **ACG CGT** ACT CAC AGT TGC CCC TCC T 3'

(-69) 5' ATT **ACG CGT** CAG TGG GTG GAG TGT GAG G 3'

(+18) 5' ATT **ACG CGT** ACA TGG ACC TGG CCC AAC 3'

Anti-sense primer for PCR amplification with the above 3 primers (with BglII restriction site in bold).

(+95) 5' CAA **GAT CTT** GTT CTC TTC TGG CTG CCG 3'

Plasmid DNAs were purified with FastPlasmid Mini (Eppendorf) for screening purposes, and restriction digest analyzed. Screened plasmid DNAs were purified using the GenElute Midiprep Kit (Sigma) and sequenced.

3.3.3 Site-Directed Mutagenesis for Mutation of Transcription Factor Binding Sites

Site directed mutagenesis was performed using the Quikchange site-directed mutagenesis kit (Stratagene) with the following primers (mutated bases underlined):

Sp1 (nucleotides -77 to -68) mutant forward:

5' GGT CCG CGA CCA A TAC CAG TGG GTG GAG TG 3'

Sp1 mutant reverse:

5' CAC TCC ACC CAC TGG TAT GGT CGC GGA CC 3'

Mzf1/Sp1 (nucleotides -54 to -45) mutant forward:

5' GGG TGG AGT GTG AGG AAA GGA GGT CGC TCG ACT C 3'

Mzf1/Sp1 mutant reverse:

5' GAG TCG AGC GAC CTC CTT TCC TCA CAC TCC ACC C 3'

Following PCR generation of mutant constructs, the native wild type plasmids were digested with DpnI and the remaining reaction mixture was transformed into JM109 chemically competent cells (Promega). Colonies following transformation were isolated and screened by sequencing.

3.3.4 Reporter Gene Assay

Cells were lysed by one freeze and thaw cycle in reporter lysis buffer (Promega). Luciferase activity was measured using the Victor light luminescence counter (Wallac/Perkin-Elmer). All experiments were carried out in triplicates and independently performed at least three times. Data of luciferase activity are normalized to protein concentration, and are shown as mean \pm SEM of three independent experiments.

3.3.5 Electrophoretic Mobility Shift Assay (EMSA)

Nuclear extracts from HeLa cells were purchased from Upstate technologies. For gel retardation experiments, 87.5fmol of ³²P-labelled double stranded oligonucleotides containing the putative Sp1 or Mzf1/Sp1 binding site was incubated for 20 minutes at room temperature with 2ug of the HeLa nuclear extract, and 1ug of poly(dI-dC) in binding buffer (Pierce). For competition assays, cold unlabelled 25x or 100x molar excess of the double-stranded oligonucleotide containing the wild type putative binding sites, the oligonucleotide with the mutated binding sites that was used for site-directed mutagenesis, an oligonucleotide containing an Sp1 consensus site, or a nonspecific oligonucleotide were preincubated for 10 minutes with the nuclear extracts before addition of the labeled wild type oligonucleotide for a further 20 minutes. For supershift assay, 2ug of Sp1(1C6) antibody (sc-420) (Santa-Cruz) or normal mouse IgG (Santa-Cruz) was preincubated with nuclear extract for 10 minutes before addition of labeled WT oligonucleotide for a further 20 minutes. Reactions were mixed with loading buffer (Pierce) and loaded on a 0.5X TBE, 5% nondenaturing polyacrylamide gel and run for 1 hour at 10V/cm. Gels were dried and exposed to radiographic film for 12-24 hours.

Oligonucleotides used for EMSA:

Sp1 (nucleotides -77 to -68)

WT: 5' ACGCCCCGCCAGTG 3'

Mutant: 5' ACGCCCATACCAGTG 3'

Sp1 Consensus: 5' TCGACGGGGCGGGGCTTA 3'

Sp1/Mzf1 (nucleotides -54 to -45)

WT: 5' ACGCCCCGCCAGTG 3'

Mutant: 5' ACGCCCATACCAGTG 3'

Sp1 Consensus: 5' TCGACGGGGCGGGGCTTA 3'

Mzf1 Consensus: 5' AGTGGGGACGGGGAGGGGGAA 3'

3.4 MISCELLANEOUS

3.4.1 Sequencing

All cDNA sequences, expression vector as well as reporter vector constructs were confirmed by sequencing using the dye termination method with the ABI 3100 Genetic Analyzer (Applied Biosystems), in the Northwestern University Biotechnology Laboratory.

3.4.2 Statistical Analysis

Data were analyzed by ANOVA with Dunnett's or Bonferoni's corrections for multiple comparisons, as appropriate. This analysis was performed with the Prism software package (GraphPad, San Diego, CA). Data were expressed as mean \pm SEM.

4. RESULTS

4.1 IDENTIFICATION AND SEQUENCING OF A NOVEL GENE, TTMP

The aim of the first part of this study was to identify and sequence a previously uncharacterized gene that is differentially expressed in TPA treated pancreatic cancer cells. The strategy chosen was to perform oligonucleotide microarray expression profiling of CD18 pancreatic cancer cells treated with TPA as compared to untreated control cells in order to identify differentially expressed transcripts. The rationale for using TPA as a trigger for differential gene expression is due to its mixed and contrasting effect on cell proliferation and the cell cycle in pancreatic cancer cells (see background). Our own observations concurred with inhibition of pancreatic cancer cell growth following stimulation with TPA and subsequent activation of PKC (142,143) (Fig 3).

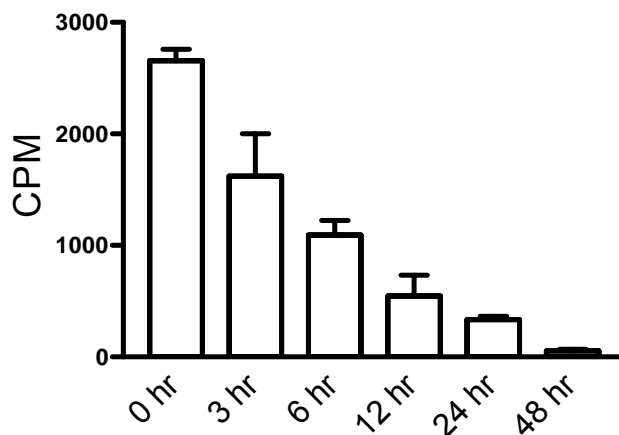


Figure 3. Time course of H³-Thymidine incorporation assay in CD18 cells following treatment with TPA at 10⁻⁷M concentration. Results shown are expressed as means of absolute counts per minute ± SEM from three separate experiments.

4.1.1 TPA Induction of TTMP

Microarray expression profiling was performed using RNA extracted from TPA-treated (0.1 μ M TPA for 8 hours) as well as untreated CD18 pancreatic cancer cells to identify differentially expressed transcripts. Genes identified as such, include several that are involved in pancreatic carcinogenesis, as well as other known genes that have yet to be studied for their role in pancreatic cancer. Genes that were differentially expressed in our study are shown in Table 1. In addition, there are a number of novel uncharacterized genes that are found to be differentially expressed following TPA treatment. The functions of these genes are unknown and myriad, and we have elected to use differential growth dynamics as an easily observed change in order to single out genes for further characterization. These hypothetical genes were transiently transfected into CD18 pancreatic cancer cells and observed for differences in cell proliferation over the subsequent days. Only hypothetical genes that contain full coding sequences with putative start and stop codons were chosen for transient transfection. One of the differentially expressed transcripts identified from the microarray study was found to be homologous to cDNA sequence AK026839 that represents a novel uncharacterized gene FLJ23186. This transcript was increased following TPA treatment in CD18 pancreatic cancer cells by a mean of 3.3-fold over untreated control cells in 3 separate experiments. Validation of the microarray data for FLJ23186 was obtained using real-time quantitative PCR on the original RNA isolates that were used for the microarray experiments. This showed a mean induction of 7.6-fold from the 3 samples. We have chosen to proceed with further structural and functional characterization of this gene based on the differential cell proliferation observed when the coding sequence was transiently transfected into CD18 cells (Fig 4). We have named this gene TPA induced Trans-Membrane Protein (TTMP) from its predicted membrane topology (see below).

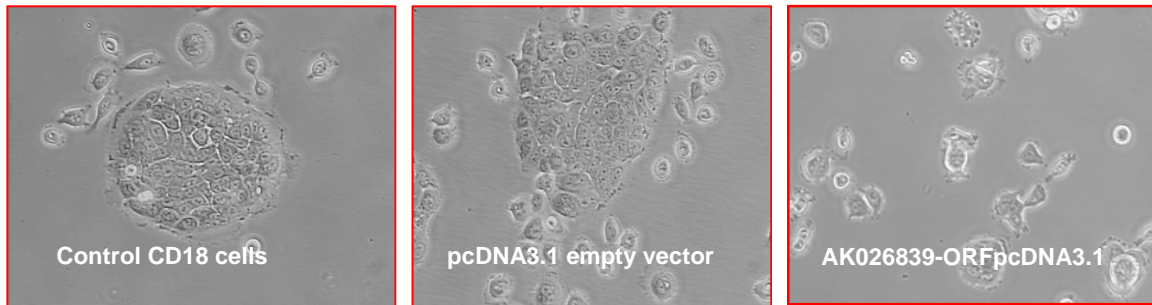


Figure 4. Differential growth dynamics at 72 hours following transient transfection of CD18 pancreatic cancer cells with AK026839-ORFpcDNA3.1 expression vector.

Concentration-response and time-course studies were performed to further characterize the pattern of induction of TTMP by TPA. The pancreatic cancer cell line, CD18 and the cervical carcinoma cancer cell line HeLa were used. The concentration-response experiments in CD18 cells showed a maximal induction of 8.2-fold with $1\mu\text{M}$ TPA for 24 hours. (Fig 5A) The time-course showed that induction of TTMP had an early onset, with a 3.4-fold increase at 2 hours and maximal induction of 10.5-fold at 6 hours with $0.1\mu\text{M}$ TPA. (Fig 5B and C) Similarly, in HeLa cells, induction of TTMP expression showed a concentration-dependent response with a maximal increase 6-fold with $0.01\mu\text{M}$ TPA for 24 hours. (Fig 5D) In time-course experiments with HeLa cells, maximal induction of 11-fold was seen at 6 hours after treatment with $0.1\mu\text{M}$ TPA. (Fig 5E and F) We have thus confirmed the induction of TTMP expression by TPA in two different cell lines.

Accession No.	Gene name and description	Relative mRNA levels \pm S.D.
NM_002421	Matrix metalloproteinase 1 (interstitial collagenase) (MMP1) mRNA	6.96 \pm 1.50
M34671	Lymphocytic antigen CD59/MEM43 mRNA, complete cds	5.27 \pm 0.55
U03106	Wild-type p53 activated fragment-1 (WAF1) mRNA, complete cds	5.23 \pm 1.92
M17017	Beta-thromboglobulin-like protein mRNA, complete cds	5.18 \pm 2.79
NM_005771	Retinol dehydrogenase homolog (RDHL) mRNA	5.13 \pm 2.28
NM_002203	Integrin, alpha 2 (CD49B, alpha 2 subunit of VLA-2 receptor) (ITGA2) mRNA	4.69 \pm 1.27
NM_002539	Ornithine decarboxylase 1 (ODC1) mRNA	4.50 \pm 0.63
NM_001109	A disintegrin and metalloprotease domain 8 (ADAM8) mRNA	4.01 \pm 1.05
NM_001218	Carbonic anhydrase XII (CA12), mRNA	4.01 \pm 1.65
NM_005114	Heparan sulfate (glucosamine) 3-O-sulfotransferase 1 (HS3ST1) mRNA	3.88 \pm 0.64
NM_004419	Dual specificity phosphatase 5 (DUSP5) mRNA	3.83 \pm 1.12
NM_005438	FOS-like antigen-1 (FOSL1), mRNA	3.80 \pm 0.42
NM_001432	Epiregulin (EREG) mRNA	3.34 \pm 1.35
AK026839	cDNA: FLJ23186 fis, clone LNG11945	3.30 \pm 0.44
NM_007231	Solute carrier family 6 (neurotransmitter transporter), member 14 (SLC6A14), mRNA	3.26 \pm 0.23
NM_001300	Core promoter element binding protein (COPEB), mRNA	3.12 \pm 0.89
NM_002639	Protease inhibitor 5 (maspin) (PI5) mRNA	3.08 \pm 0.23
NM_004995	Matrix metalloproteinase 14 (membrane-inserted) (MMP14) mRNA	2.87 \pm 0.85
NM_014302	Sec61 gamma (SEC61G), mRNA	2.82 \pm 0.06
NM_012425	Ras suppressor protein 1 (RSU1), mRNA	2.66 \pm 0.40
NM_016038	CGI-97 protein (LOC51119), mRNA	2.63 \pm 0.12
NM_001197	BCL2-interacting killer (apoptosis-inducing) (BIK), mRNA	2.59 \pm 0.44
AF320070	Hepatocellular carcinoma-associated protein HCA10 mRNA, complete cds	2.58 \pm 0.42
NM_006529	Glycine receptor, alpha 3 (GLRA3), mRNA	2.57 \pm 0.42
NM_001102	Actinin, alpha 1 (ACTN1) mRNA	2.52 \pm 0.20
NM_003560	Phospholipase A2, group VI (PLA2G6), mRNA	2.50 \pm 0.40
NM_003648	Diacylglycerol kinase, delta (130kD) (DGKD) mRNA	2.48 \pm 0.43
M59040	Cell adhesion molecule (CD44) mRNA, complete cds	2.37 \pm 0.43
NM_002149	Hippocalcin-like 1 (HPCAL1) mRNA	2.32 \pm 0.18
NM_004354	Cyclin G2 (CCNG2), mRNA	2.32 \pm 0.24
M30142	Decay-accelerating factor mRNA, complete cds	2.31 \pm 0.20
NM_001665	Ras homolog gene family, member G (rho G) (ARHG) mRNA	2.31 \pm 0.07
AF303888	Microtubule-associated proteins 1A/1B light chain 3 mRNA, complete cds	2.27 \pm 0.23
NM_004156	Protein phosphatase 2 (formerly 2A), catalytic subunit, beta isoform (PPP2CB) mRNA	2.25 \pm 0.26
AF118124	Myeloid cell leukemia sequence 1 (MCL1) mRNA, complete cds	2.24 \pm 0.11
NM_004616	Transmembrane 4 superfamily member 3 (TM4SF3) mRNA	0.49 \pm 0.02
NM_000387	Carnitine/acylcarnitine translocase (CACT) mRNA	0.45 \pm 0.07
NM_002083	Glutathione peroxidase 2 (gastrointestinal) (GPX2) mRNA	0.45 \pm 0.06
NM_006149	Lectin, galactoside-binding, soluble, 4 (galectin 4) (LGALS4) mRNA	0.44 \pm 0.03
AF190122	Putative superoxide-generating NADPH oxidase Mox2 mRNA, complete cds	0.42 \pm 0.02
NM_007127	Villin 1 (VIL1), mRNA	0.42 \pm 0.09
NM_016253	3-alpha hydroxysteroid dehydrogenase type IIb (LOC51708), mRNA	0.35 \pm 0.05
NM_002165	Inhibitor of DNA binding 1, dominant negative helix-loop-helix protein (ID1) mRNA	0.25 \pm 0.07

Table 1. Genes differentially expressed after 8 hours of TPA treatment. The cDNA sequence with homology to TTMP is in bold. Other novel uncharacterized hypothetical proteins have been omitted from the list.

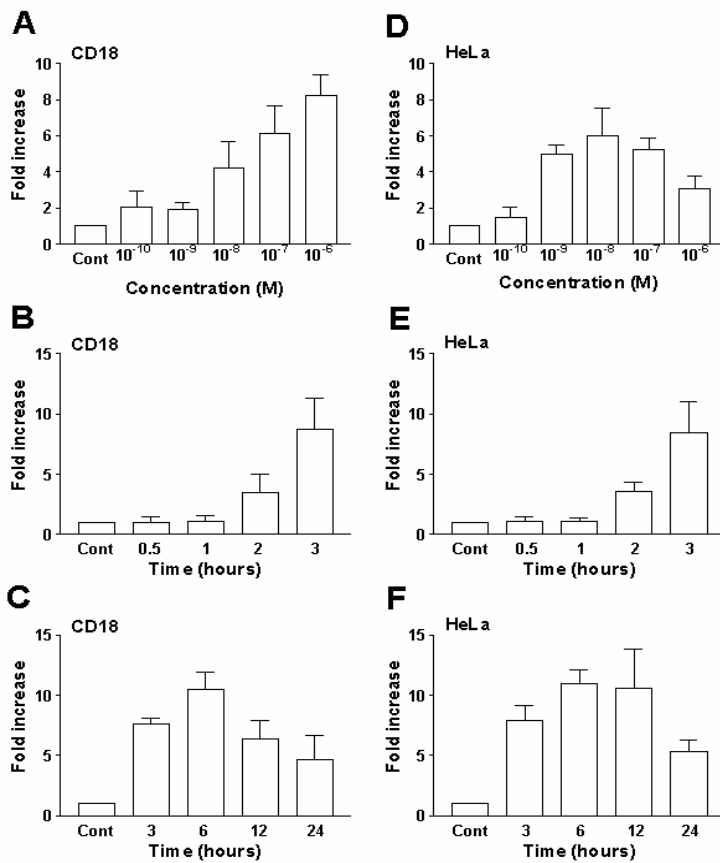


Figure 5. Concentration-response (A and D) and time-course following TPA treatment for CD18 (B and C) and HeLa (E and F) cells, respectively. Results are shown as means of fold increase \pm SEM over control cells on real-time quantitative PCR after normalization with GAPDH expression from two separate experiments.

4.1.2 Full Length Transcript(s) of TTMP

We used 5'-RACE to identify full length transcript(s) containing the transcription start site(s) of TTMP. 5'-RACE was performed using RACE-ready cDNA from human colon and a pair of TTMP-specific nested primers. A single strong band of about 400 to 500 bp and several weaker bands were seen on agarose gel electrophoresis after PCR amplification. These PCR products were cloned and sequenced. Isolates with inserts of

sizes other than the dominant band of around 400 to 500 bp did not contain sequences with homology to TTMP. Of the isolates with sequence homology to TTMP, the majority contains an adenine residue at the 5' end which is just 3 nucleotides upstream from the start of the deposited cDNA sequence (GenBank accession no. AK026839). The sequence around this predominant transcription start site completely matches the consensus initiator element sequence (Inr) commonly found in TATA-less promoters (161). Furthermore, the nucleotides at position +28 to +32 bear striking resemblance to the distal promoter element (DPE) which acts in concert with the Inr to initiate the start of transcription (205) in promoters lacking a TATA box. (Figure 6A and B)

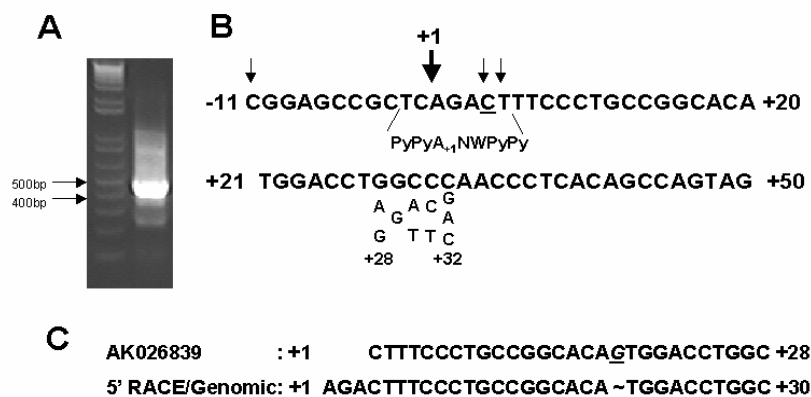


Figure 6. The transcription start sites of TTMP. (A) Agarose gel electrophoresis of 5'-RACE products showing a dominant band between the 400 and 500bp mark of the DNA ladder. (B) Sequence of the 5' end of TTMP cDNA obtained by 5'-RACE. Arrows denote the transcription start sites identified. The major transcription start site is denoted as +1. The first nucleotide on the previously deposited sequence AK026839 is underlined. The consensus sequence for Inr and DPE are shown below their corresponding positions on the cDNA sequence. (C) Comparison between the previously deposited sequence AK026839 and the 5'-RACE/Genomic sequence that showing omission of a guanine residue. The resulting ATG is the putative start codon of the open reading frame from the full-length sequence AY830714.

We also performed 3'-RACE using primers which gave products that overlapped those obtained from 5'-RACE in order to completely sequence TTMP. The overlapping sequences have been assembled and deposited into GenBank under the accession no. AY830714. We compared this sequence with the deposited GenBank cDNA sequence AK026839. This revealed an extra guanine nucleotide at position 18 of the deposited sequence AK026839. However, bioinformatics analysis of our sequencing data showed complete homology at the 5' end with human genomic sequences (NT_005612 and NT_86640) as well as with most of the homologous human ESTs deposited in GenBank. The absence of this guanine residue in our deposited sequence brings into tandem the preceding adenine residue with the downstream thymine-guanine residues, thus forming a putative start codon. In-silico analysis (ORF Finder, NCBI: <http://www.ncbi.nlm.nih.gov/gorf/gorf.html>) predicts an open reading frame beginning at this start codon, which is an in-frame N-terminal extension of the putative coding sequence described in the deposited sequence AK026839. The deletion of the extra guanine residue from our sequencing data is shown in Figure 6C. This is significant, since the previously predicted amino acid sequence starts from an ATG site further downstream at nucleotide position 219 of AK026839. This encodes a product with a short hydrophilic N-terminal followed by the central hydrophobic domain. Although structurally this putative protein would have conformed to a membrane topology, as seen with the longer sequence, it could be lacking certain residues on the N-terminal end that may confer important and specific functions of this protein.

The full-length TTMP cDNA sequence AY830714 as well as the predicted amino acid sequence is shown in Fig 7. In addition to the previously mentioned discrepancy in the 5' untranslated region surrounding the putative start codon, there are two other single nucleotide differences in the open reading frame of this sequence when compared

to the cDNA sequence AK026839 or the genomic sequences NT_005612 and NT_086640. The nucleotide in position 386 of AY830714 as well as its equivalent position in the genomic sequences is a guanine, whereas an adenine is found in its equivalent position in the cDNA sequence AK026839. The resulting codon on AK026839 yields a lysine residue as compared to a glutamic acid residue at position 123 of the amino acid sequence predicted from the full-length cDNA AY830714. In position 449 as well as that in its homologous position on AK026839, an adenine is found. Whereas the same nucleotide is found in the alternate genomic assembly NT_086640, a guanine is noted in place on the equivalent position on human genomic sequence NT_005612. In addition, several other single nucleotide differences are also noted in the 3' untranslated region (Fig 7). The disparity in the nucleotide sequences may have arisen from errors in sequencing, or may represent single nucleotide polymorphisms in the gene. However, if this did represent polymorphism, the difference between a basic and an acidic amino acid residue at position 123 in the protein sequence could have marked functional consequences.

1	agactttcccctgccggc acATGGACCTGGCCCAAC CCTCACAGCCAGTAG	50
1	H D L A Q P S Q P V D	11
51	ACGAGCTGGAGCTCTCGGTCTCGAGCGGCAGC CAGAAAGAAACACGCCCT	100
12	E L E L S V L E R Q P E E N T P	27
101	CTCAATGGTGC CGA CAA GGTCTTCC CTTCCTTGGACCGAGGCGTCC CCC C	150
28	L N G A D K V F P S L D E E V P P	44
151	GGC CGAGGC TAACAAGGAAAGC CCCTGGAGCTC CTGTAATAAGAAATGTGG	200
45	A E A N K E S P W S S C N K N V V	61
201	TTG GAA GAT GC AAA CTG TG GAT GAT CAT C A CCT CC ATTTTC TAGGTTCC	250
62	G R C K L W H I I T S I F L G V	77
251	ATTACA GTG AT CAT CATAG CCTTATGTCTT GCT GC AGT AAC TTATGTTGA	300
78	I T V I I I G L C L A A V T Y V D	94
301	TGAAGATGAAAATGAAA TA CTT GAA TTATC ATC AAA CAAA CA TTCTTCA	350
95	E D E N E I L E L S S N K T F F I	111
351	TCATGC TGAAGATT CCA GAGGAGTGTGTTG CTGAA GAG GAA TTGCC TCA C	400
112	H L K I P E E C V A E E E L P H	127
401	CTG CTC ACC GAAAG GCT C A CAGATGTG TAC AGTAC ATC GCC CTCTCTGAG	450
128	L L T E R L T D V Y S T S P S L S	144
451	TCGTTATTTTA CTT CAGTT GAAA TA GT GGA CTT CA GTG GTG AAAATGCCA	500
145	R Y F T S V E I V D F S G E N A T	161
501	CAGTAA CGTATGAC CTG CAATTTGGGGTTC CAT CAGATGATGAAAATTTT	550
162	V T Y D L Q F G V P S D D E N F	177
551	ATGAACTATATGATGAGTGGAGGTTG GTG CTGGC CATTTT GCTACAGGA	600
178	H K Y H H S E E L V L G I L L Q D	194
601	TTTCCCTGATCAGAAATATA CCTGGTTG TGA GAGTCTGGGGCTTGATCCAA	650
195	F R D Q N I P G C E S L G L D P T	211
651	CAT CCC TCTTGCTCTATGAAATGAAgtgatggaggc tggctctctgtctgaa	700
212	S L L L Y E *	217
701	agc agt gct ct acc aagtc ctggagat gaa ggg aatcc act ct gttttgc	750
751	aga aaa gat tct tgt ggatt aat ac a ga agr acc agcaac ac ca gag ggg	800
801	gga gar tcc tttct ctc cc gat tct ac agt ctggc tct aagcc cagtaaa	850
851	ac a gct ccc ga gca actgcttca gct gggtc cagtc ttgaca aa ggc agga	900
901	agc cagcta gggg ggggc gat agggc cagc gggc atgtcc cactgttg	950
951	aggtc actgt att ctgtt tgt ttt tgt ttt tgt tct cgt ttt gt ttt tga	1000
1001	gar agggc tct gtt ctgtc gcttagct gga gtgc ggtggc gctgac atgg	1050
1051	cactgt tcttcttgaagc actc cacc acc tgggc tacc ttttcttttagt	1100
1101	gca gaggtgc actgtcttcttttaggt gggatc gc gta agc at gagctgg	1150
1151	tagagc argga gagga ggc agcc aggtta cga agact aagcc aatatt	1200
1201	cactga agt cactc tcttc tcc ccc acc cactc gat tggatc taccctc taa	1250
1251	gct aggtctgtgaagaaa aggaa ggc ac ttt aga agacc tca gc agt ggg	1300
1301	ttctgt gct tacc tcc atgacc tgc ac ctg agt at ctt agc ca ggc agc	1350
1351	tta gga ac acc acc aaggttac ttt ga aat ctatgtat ata gctagttac	1400
1401	agc gggagct atgtgtttctt catta ttttgc agctc ctc gttgttctt	1450
1451	gtgatt cct ga acac ctttttggaa at atgggtc ctctgtga gtttt ggc	1500
1501	ac actactac actttggatagttc actcra ttt at ttttt at aactc	1550
1551	cattagaga atcattaa ggtgtttt aat at ctgct ctggat at tar gcat	1600
1601	tggcttttt gttgc tca gctgt ac aaaa ccttctctgtgggac tct agt gct	1650
1651	tca ggc aat gatgtgt at cct gtt ac agatgggt ggtata ca aga aga	1700
1701	cca ctc ttt tttga aaata aac ttttggaa gcttttgc cagct atctggg	1750
1751	gggtaggag gata tta gc aac tgt at tgggtgtc tac agata cagaatt	1800
1801	gctgtgt ga ggg aar tgatt gtttt gtt gggaa aagaaa tttac cagg	1850
1851	aga aag agt tttgt gct gt att gtg ag agatct cgcct ctc agtta aatg	1900
1901	agc cct gggtt aaa gtc agtgt gaa ggg ca gct gt gtc gggc ac ggc c	1950
1951	aga gctgtc gc rcc agact agattt ga ctt gagtt ctttat gaccc agga	2000
2001	ctc tggata at gta aat ttt gcttct ct atttaact aga agata c at gta c	2050
2051	tat agatc att gtc tca tttta gct at tgt tctt aaaa cta gt gaa act a	2100
2101	gtggat tttc tct tct ct ttttat ttt ct gca t gtt a aat gt gaa ctt	2150
2151	agt gta ttt gt attttgta gaa aat aat ga aaa at ttt aat ggaga atga	2200
2201	ttt aaaa acc attta caata c attat ttttgc at ct gta aaaa aaaaaaaa	2250
2251	aa	

Figure 7. Nucleotide sequence and deduced amino acid sequence of human TPA induced Trans-Membrane Protein (TTMP) (AY830714). Uppercase letters represent the coding region, lowercase letters represent the 5' and 3' untranslated regions (UTR). Differences in the nucleotide sequence between TTMP and cDNA sequence AK026839 and/or genomic sequence are in italics and underlined. * represents the stop codon. The underlined sequence aataa in the 3'UTR is a consensus polyadenylation sequence.

4.1.3 In-Silico Analysis of TTMP

Using the full-length cDNA sequence, we performed a blast comparison with the human genomic sequence database at the NCBI. This revealed that the TTMP gene is located on the long arm of chromosome 3 (3q13.2) and lies between nucleotide positions 18300382 to 18332203 of genomic sequence NT_005612. There is also an alternate assembly found on chromosome 3 genomic sequence NT_086640 at positions 18338313 to 18370147. Both alignments are in the sense direction. Exon-intron boundaries were determined based on alignment of the cDNA sequence with the genomic sequence NT_005612, as well as on the “gt....ag’ rule of intron-exon splicing. The TTMP gene is about 33kb in length and is organized into 6 exons, with the coding region spanning from exon 2 to exon 6. (Table 2)

Exon	No of bases	Nucleotide sequence around exon-intron boundaries		Nucleotide position in NT_005612.14
1	156		AGACTT....GCCGAG	18300382 – 18300537
2	130	ttctag	GCTAAC....CTGCAG	18307353 – 18307482
3	128	ctttag	TAAGTT....GAAAGG	18316831 – 18316958
4	71	gtttag	CTCACA....CTTCAG	18323536 – 18323606
5	182	tttcag	TGGTGA....TCTATG	18326957 – 18327138
6	1569	ttttag	AATGAA....ATCTGT	18330635 – 18332203

Table 2. Exon-intron structure of TTMP. Nucleotide sequences of exons and introns are shown in upper and lower case respectively.

The predicted open reading frame is from nucleotides 20 to 673 and codes for a 217 amino acid product with a calculated molecular mass of 24.30kDa and a pI of 4.02. Protein topology analysis using PSORTII (<http://psort.nibb.ac.jp>) predicted a type II trans-membrane protein sited across the endoplasmic reticulum (ER) membrane with a single transmembrane domain from residues 73 – 89. Protein sequence analysis using a separate online software program, TMHMM (<http://www.cbs.dtu.dk/services/TMHMM>) also predicted a similar transmembrane protein with a single transmembrane helix from

residues 66-88. Both programs predicted a membrane topology with the C-terminal sited in the cytoplasm (Fig 8). Motif scanning analysis using Scansite (<http://scansite.mit.edu>) did not find sequence homology with any conserved functional domain. Therefore, human TTMP may represent the founding member of a hitherto uncharacterized protein family with novel functional domains. However, the possibility exists that the software analysis was unable to recognize degeneracy in the amino acid sequence that may contain similar functional domains to known protein motifs.

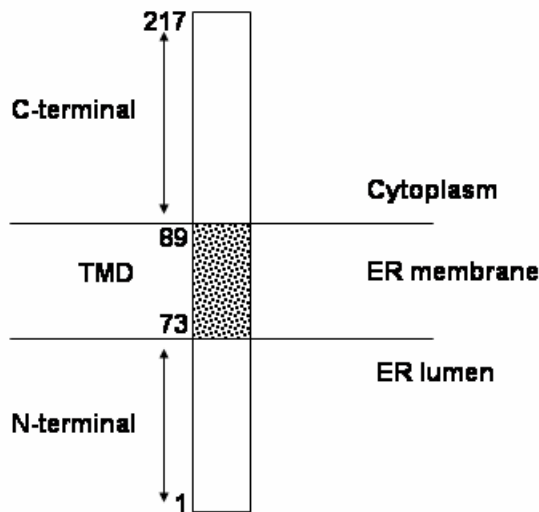


Figure 8. The deduced membrane topology of TTMP, ER=endoplasmic reticulum.

4.1.4 Conservation of Orthologous Gene Sequence in Mouse and Chicken

Comparison of the predicted protein sequence from murine orthologous cDNA sequence AK078878 with human TTMP revealed a 68% sequence homology. Notably, the central hydrophobic domain together with the C-terminal end of the predicted protein

from residues 61 to 128 shows greater conservation with the human sequence. However, the chicken orthologous cDNA XM_416636 displayed only 29% identity to human TTMP (Fig 9). Unlike murine TTMP, the chicken orthologue is equally divergent at both the N- and the C-terminal end. This probably represents divergence of the gene early on in evolution. The conservation of the C-terminal end of both human and mouse genes may underlie a common structure or function. While the function of this protein is not yet known, it is possible that the differences in the N-terminal sequences give rise to species specific activity of the translated protein.

A

```

H. sapiens      MDLAQPSQ-PVDELELSVLERQPEENTPLNGADKVFPSLDEEVPPAEANKE$PWS$SCNKN 59
M. musculus    MEEGSR$QS$PREELELSMLDGPQEELT$PLNNDLRIQ$ENSAEDP$SPAQVGKE$PWS$PCNKS 60
               *   ** *   ***** *   ** *****   *   *   **   ***** **

H. sapiens      VVGRCKLWMIITSIFLGVITVIIIGLCLAAVTVVDEDENEILELS$SNKTF$FIMLKIPEEC 119
M. musculus    VVGKCKLWMMVITIFLFCFIIIVIVISLCLVGVTYIDEDENEILELS$SNKTF$FIMLKIPEEC 120
               *** ***** *   ** *   ** *   ***   ** *****

H. sapiens      VAEEELPHLLTERLTDVYST$P$LSRYFT$SVEIVDF$SGENATVTYDLQF$GVPSDDENFMK 179
M. musculus    ANEEELHLLTERLTDTYRQ$PALSRRFFTSADILDF$SVENATVTYHLQF$GVPSDDDDFMK 180
               **** ***** *   ** *** ** *   ** ***** ***** *   ***

H. sapiens      YMMSEELVGLGILLQDFRDQNI$PGCESLGLDPT$SLLLYE 217
M. musculus    YMMSEELVGLGIMRQ$SFHDKNI$STCESLGLDPE$SLLLYE 218 68% identity
               ***** *   *   *   *   ***** *****

```

B

```

H. sapiens      MDLAQPSQP--VDELELSVLERQ-----PEENTPLNGADKVFPSLDEEVPPAEA 47
G. gallus      MSWLRACYRDRGPRTE$TEAMKRQ$SGGEHEVIELPEVNGEESTADHKKP-LNPQVPAT$K 59
               *           *   **           ** *   ** *   *   **

H. sapiens      NKES$PWS$SCNKNVGRCKLWMIITSIFLGVITVIIIGLCLAAVTVVDEDENEILELS$SNK 107
G. gallus      ERDQ-WK$SCRKI$E$WKCKLW$MVLTTIFVVLFLVILISLALYSNVYTD$DDYWYTD$ELLQN 118
               *   ** *   ***** *   **   ** *   *   *   **

H. sapiens      TFFIMLKIPEECVAEEELPHLLTERLTDVYST$P$LSRYFT$SVEIVDF$SGENATVTYDLQ 167
G. gallus      YHNF$GKENLLCGLPHVF$SEDI$IKRITDVYSS$PALGRYFR$SAKVDYF$SNE$SSTVFYQLE 178
               *   *           *   ***** *   ** *   ** *   *   **

H. sapiens      FGVPSDDENFMKYMMSEELVGLGILLQDFRDQNI-----IPGCESLGLDPT$SLLLYE 217
G. gallus      FFVPPSTEGFMENV$MNPDFIRNVLLQNIYDEEDT$SNPGT$SECTRLKLDPVSLT$ST- 233 29% identity
               *   ** *   ** *   *   ** *   *   *   ** *   **

```

Figure 9. Alignment of the amino acid sequences of human TTMP with (A) mouse and (B) chicken orthologues. The mouse and chicken sequences are 68% and 29% identical to the human sequence, respectively.

4.1.5 Mechanism of TTMP mRNA Induction by TPA

We investigated whether the up-regulation of TTMP mRNA by TPA was mediated by a decrease in the rate of mRNA degradation or by an increase in TTMP transcript levels by activation of its promoter. CD18 cells were pre-treated with or without 0.1 μ M TPA for 4 hours before both treated and control cells were exposed to 5 μ g/ml actinomycin D to inhibit further transcriptional activity. A time-course was then performed and real-time PCR was used to determine TTMP mRNA transcript levels following inhibition of transcription. The results show that TPA did not change TTMP mRNA degradation kinetics, suggesting that the up-regulation of TTMP by TPA is not through a change in mRNA stability. (Fig 10A and B)

We then investigated whether TPA increases TTMP promoter activity. PCR techniques was used to clone a 2kb fragment (-1909 to +95 relative to transcription start site) of the human genomic TTMP 5' flanking region into the promoterless luciferase reporter vector pGL₃-basic. Basal activity of the cloned fragment (-1909/+95pGL₃) was determined at 48 hours after transient transfection into HeLa cells and this was found to be increased by 20-fold over the empty pGL₃-basic vector. (Fig 10C) The construct was then used to test for promoter activity following treatment with 0.1 μ M TPA for 6 hours. This showed a 2-fold increase over non-treated controls. This suggests that TPA up-regulates TTMP gene expression at the promoter level (Fig 10D). The 2-fold increase in promoter activity following TPA induction is modest compared to the 11-fold increase seen in the mRNA levels at the same time point. This may be due to the synergistic requirement of other cis-acting transcriptional elements, which are not present in our 5' flanking construct of 2kb, to further enhance transcriptional activation.

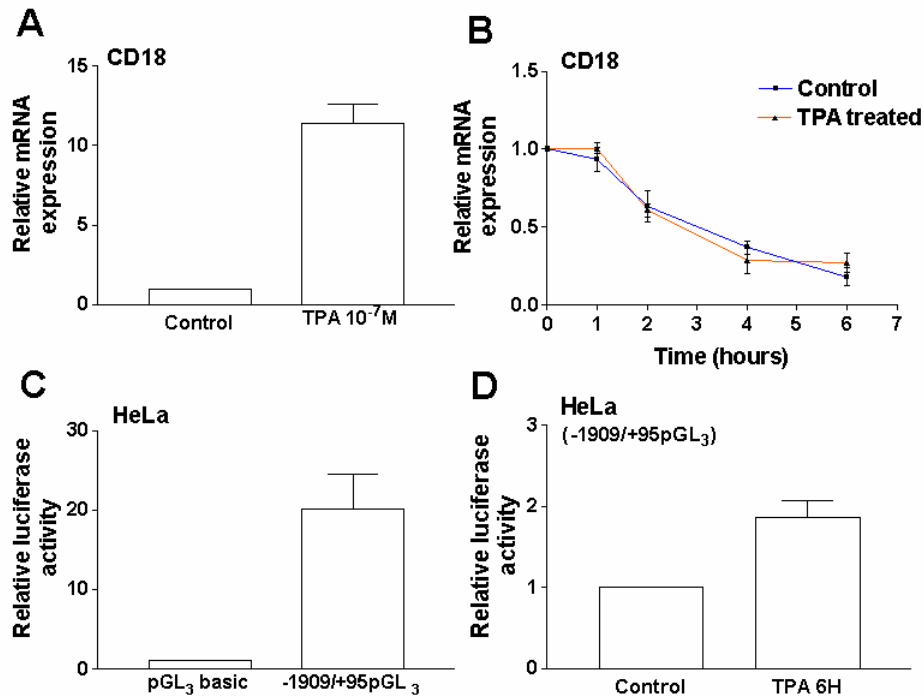


Figure 10. Induction of TTMP mRNA expression in CD18 cells after 4 hours of stimulation with $0.1\mu\text{M}$ TPA (A) and subsequent expression levels at various time points after addition of actinomycin D ($5\mu\text{g/ml}$) to both TPA stimulated and control cells (B). Results shown are means of two separate experiments \pm SEM on real-time quantitative PCR after normalization with GAPDH expression levels. Luciferase activity of 2kb TTMP promoter construct (-1909/+95pGL₃) compared to promoterless pGL₃-basic control vector (C). Induction of promoter activity of TTMP promoter construct after 6 hours of stimulation with $0.1\mu\text{M}$ TPA (D). Results for (C) and (D) are means \pm SEM from three separate experiments after normalization with protein concentration. Each experiment was carried out in triplicates.

4.1.6 Conclusions

Using oligonucleotide microarray analysis, we have identified a novel gene that is upregulated following treatment with TPA in the pancreatic cancer cell line CD18. Real-time PCR validated the microarray results in CD18 and HeLa cells, and showed that upregulation of the gene is time- and concentration dependent. In-silico analysis shows that this is a single-pass transmembrane protein of 217 residues that is localized to the

endoplasmic reticulum, thus the name TPA induced Trans-Membrane Protein (TTMP). A luciferase reporter assay demonstrates that upregulation of TTMP by TPA is triggered at the promoter level.

4.2 EXPRESSION, STRUCTURAL AND FUNCTIONAL CHARACTERIZATION OF TTMP

The aim of the second part of this study was to study the expression pattern, as well as perform structural and functional characterization of the novel gene product, TTMP, which we had previously described (204).

4.2.1 Expression of TTMP in Normal Pancreas and Cancer Cell Lines

We first examined the level of expression of the TTMP mRNA in both normal human tissues and cancer cell lines. Commercially available, multiple human tissue cDNA panels were used to quantify expression of TTMP by real-time PCR. Expression of TTMP was higher in human pancreas than in all of the other 18 tissues, and about twice as high as in the second highest expressing tissue (the rectum). Indeed, expression of TTMP in normal pancreas was more than 5-fold higher than the mean expression level of the 19 different tissues analyzed (Fig 11A). Expression levels in the human cervical carcinoma cell line HeLa and the human pancreatic cancer cell line CD18 were more than 10-fold lower than that of normal pancreatic tissue (Fig 11B). The low expression levels of TTMP observed in these cancer cell lines made them suitable for the subsequent forced expression studies.

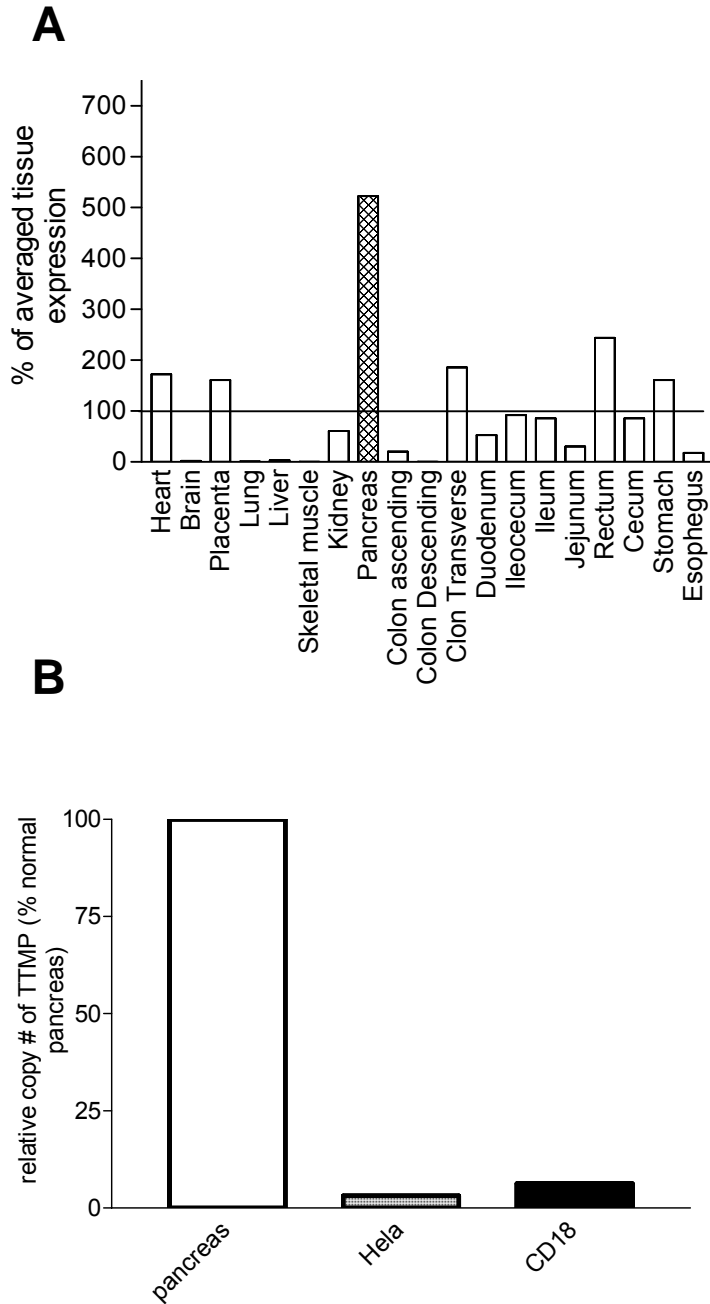


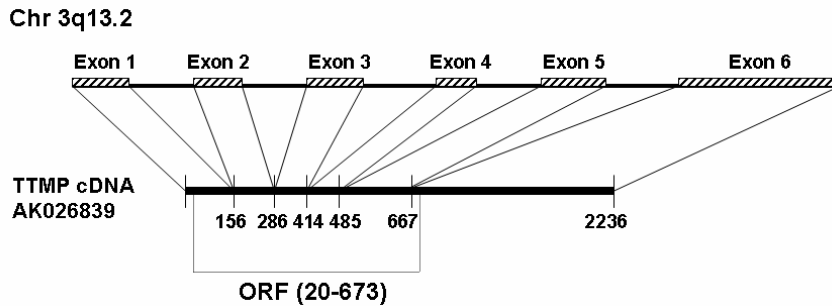
Figure 11. (A) Expression profile of TTMP in different normal tissues. TTMP mRNA expression was measured by real-time RT-PCR, using cDNA from different normal human tissues, and normalized by GAPDH and beta-actin. TTMP expression in the pancreas was the highest among the 18 different tissues examined. (B) Comparison of TTMP expression in normal human pancreas, CD18 pancreatic cancer cells and HeLa cells. TTMP mRNA expression was measured by real-time RT-PCR and normalized to GAPDH. Expression of TTMP mRNA in cancer cell lines is substantially lower than that in the normal pancreas.

4.2.2 Identification of Translation Start Site and Molecular Size of TTMP

The predicted open reading frame from the TTMP cDNA sequence (AY830714) is from nucleotides 20 to 673, and codes for a 217 amino acid product with an estimated molecular mass of 24.30 kDa and a pI of 4.02 (Fig 12A & B). Interestingly, a possible alternative translation start site was identified at nucleotide position 221, giving a putative in-frame N-terminal truncated protein that is 67 residues shorter than the full length protein with a predicted molecular mass of 16.88 kDa (Fig 12B). Thus, either of these putative translation start sites could be recognized by the ribosomal translational machinery to initiate the start of translation. To confirm which site was responsible, we cloned the 5' UTR of the TTMP cDNA together with the coding sequence into pcDNA3.1 (5'UTR/TTMPpcDNA3.1). This cloning strategy abolished the exogenous Kozak sequence on the pcDNA3.1 expression vector, making the start of translation dependent on the native Kozak sequence found in the 5' UTR (Fig 12C). The full-length TTMP (TTMPpcDNA3.1) as well as the truncated sequence (CT-TTMPpcDNA3.1) were both directionally cloned into pcDNA3.1 as described under experimental procedures. Cell lysates were obtained following transient transfection in TSA201 cells and expression of the protein products from the various constructs were identified by immunoblotting against the surrogate marker V5. A ~17 kDa product was obtained from the CT-TTMPpcDNA3.1 construct, and this was in agreement with its predicted molecular weight (Fig 13). However, the protein product obtained from the full-length TTMP construct (TTMPpcDNA3.1) was ~48 kDa in size, approximately twice the size of its predicted molecular mass of 24.3 kDa (Fig 13). The protein product obtained from 5'UTR/TTMPpcDNA3.1 is of identical size to that expressed by the full-length coding sequence that is cloned in tandem with the exogenous Kozak sequence (CACC) on pcDNA3.1 (Fig 13). This shows that the 5' UTR directs translation from the same start site as that predicted from bioinformatics (204). The lower band seen in the lane

representing TTMP could represent an incomplete post-translational modification (glycosylation) of the native TTMP protein (see next section).

A



B

```

1                               AGACTTTCCTGCCGGCACA   19

20  ATGGACCTGGCCAAACCTCACAGCCAGTAGACGAGCTGGAGCTCTCGGTGCTCGAGCGG   70
1   M D L A Q P S Q P V D E L E L S V L E R   20
80  CAGCCAGAAGAGAAACACGCCTCTCAATGGTGCCGACAAGGTCTTCCCTTCTTTGGACGAG   140
21  Q P E E N T P L N G A D K V F P S L D E   40
140 GAGGTCCCCCGGCCGAGGCTAACCAAGGAAAGCCCTGGAGCTCCTGTAATAAGAATGTG   200
41  E V P P A E A N K E S P W S S C N K N V   60
200 GTTGGAAAGATGCAAACTGTGGATGATCATCACCTCCATTTTCTAGGTGCATTACAGTG   260
60  V G R C K L W M I I T S I F L G V I T V   80
260 ATCATCATAGGCTTATGTCTTGCTGCAGTAACTTATGTTGATGAAGATGAAAATGAAATA   320
80  I I I G L C L A A V T Y V D E D E N E I   100
320 CTTGAATTATCATCAAAACAAAACATTCTTCATCATGCTGAAGATTCAGAGGAGTGTGTT   380
100 L E L S S N K T F F I M L K I P E E C V   120
380 GCTGAAGAGGAATTGCCTCACCTGCTCACCGAAAGGCTCACAGATGTGTACAGTACATCG   440
120 A E E E L P H L L T E R L T D V Y S T S   140
440 CCCTCTCTGAGTCGTTATTTTACTTCAGTTGAAATAGTGGACTTCAGTGGTGAAAATGCC   500
140 P S L S R Y F T S V E I V D F S G E N A   160
500 ACAGTAACTATGACCTGCAATTTGGGGTCCATCAGATGATGAAAATTTTATGAAGTAT   560
160 T V T Y D L Q F G V P S D D E N F M K Y   180
560 ATGATGAGTGAGGAGTTGGTGTGGGCATTTTGCTACAGGATTTCCGTGATCAGAATATA   620
180 M M S E E L V L G I L L Q D F R D Q N I   200
620 CCTGGTTGTGAGAGTCTGGGGCTTGATCCAACATCCCTTTGCTCTATGAATGA   674
200 P G C E S L G L D P T S L L L Y E *   217
  
```

C

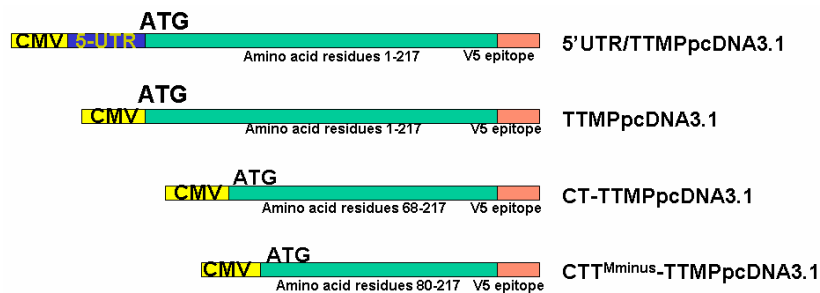


Figure 12. (A) Genomic organization of TTMP. The full length cDNA of TTMP was identified by 5'-rapid amplification of cDNA end (5'-RACE) and 3'-RACE. The full length cDNA sequence (gene bank accession #: AY830714) was used to blast human genomic sequence in Genbank. *TTMP* gene spans 21.8 kb and comprises 6 exons and 5 introns. The assembled human TTMP comprises a putative TATA-less promoter region, start/stop codons and a poly-A signal. (B) The open reading frame (ORF) of TTMP

encodes a protein with 217 amino acid residues. It contains one single transmembrane domain. No putative conserved functional domains were detected. (C) Diagram showing different pcDNA3.1 expression constructs. The cDNA sequence of TTMP ORF was fused with the V5 epitope tag at the 3'-end to enable the immunodetection of TTMP.

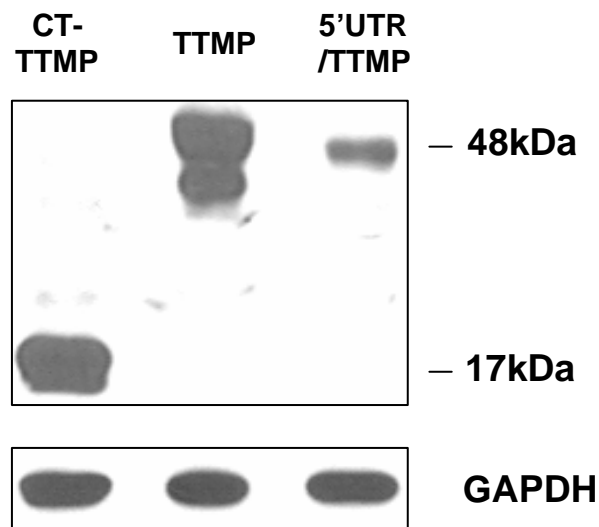


Figure 13. The molecular size of TTMP. The TTMP pcDNA3.1 expression plasmid was transfected into TSA201 cells for 24 hours. Proteins in the cell lysate were separated on SDS-PAGE gels transferred onto nitrocellulose membranes and the TTMP-V5 epitope was detected using a V5 antibody. The predicted molecular size of full-length TTMP without post-translational modification is 24 kDa. The observed molecular weight was ~48 kDa as shown here. In contrast, the molecular size of the C-terminal fragment of TTMP (CT-TTMP) was 17 kDa, which was identical to the predicted molecular size of this 150 amino acid peptide.

4.2.3 TTMP is N-Glycosylated and also Contains Sialic Acid

Western blot analysis of exogenously expressed TTMP showed a band corresponding to ~48 kDa. This is approximately twice the predicted molecular weight of 24 kDa. The TTMP open-reading frame contains several putative N-glycosylation sites (<http://www.cbs.dtu.dk/services/NetNGlyc/>), suggesting that the native protein undergoes post-translational glycosylation (Fig 14). Of the various forms of protein glycosylation found in eukaryotic systems, the most important types are in the form of asparagine-linked (N-linked) or serine/threonine-linked (O-linked) oligosaccharides (206-208). N-linked glycosylation is a co-translational process involving the transfer of the precursor

oligosaccharide $\text{GlcNAc}_2\text{Man}_9\text{Glc}_3$ to asparagine residues in the protein chain (206-208). The asparagine usually occurs in a Asn-Xaa-Ser/Thr, where Xaa is not a Proline. O-linked glycosylation involves the post-translational transfer of an oligosaccharide to a serine or threonine residue (206-208). To deglycosylate oligosaccharides from a core protein, N-glycanase (PNGase F) is the most effective method of removing virtually all N-linked oligosaccharides from glycoproteins. Complete removal of the O-linked structure or its derivatives would require, in addition to sialidase A, a beta(1-4)-specific galactosidase and a beta-N-acetylglucosaminidase (206-208). To confirm the computational suggestion that TTMP is N-glycosylated, protein extracts from transiently transfected TSA201 cells were subjected to deglycosylation treatment with N-glycanase (PNGase F), sialidase, O-glycanase, and a mixture of these enzymes prior to western blot analysis. N-glycanase treatment caused a pronounced shift in molecular size to approximately 24 kDa, which is identical to the predicated molecular size of TTMP in the absence of any post-translational modification (Fig 15A). Sialidase treatment caused a smaller shift in molecular weight, while the combined treatment with all three enzymes had a similar effect on molecular weight to N-glycanase treatment alone (Fig 15A). These findings suggest that the N-linked carbohydrate chains contained some sialic acid residues in the region not immediately adjacent to the protein core and thus the glycosylation is comprised of complex oligosaccharide. No change in molecular size is seen following treatment with O-glycanase, either alone, or with combined sialidase treatment. Finally, there was no change in molecular size of the truncated protein CT-TTMP following treatment with any of the glycanases, indicating that asparagine residues in the N-terminal end but not the C-terminal end of TTMP are glycosylated (Fig 15A). Besides sialylation, O-linked modification containing $\beta(1-4)$ galactose and $\beta(1-6)$ -linked N-actylglucosamine can block the action of O-glycanase. However, treatment with $\beta(1-4)$ galactosidase and/or β -N-actylglucosaminidase, as well as together with O-

glycanase treatment did not lead to a shift in the molecular size of the native protein (Fig 15B). Thus we conclude that TTMP is not O-glycosylated, and the sialic acid residues are most likely N-linked.

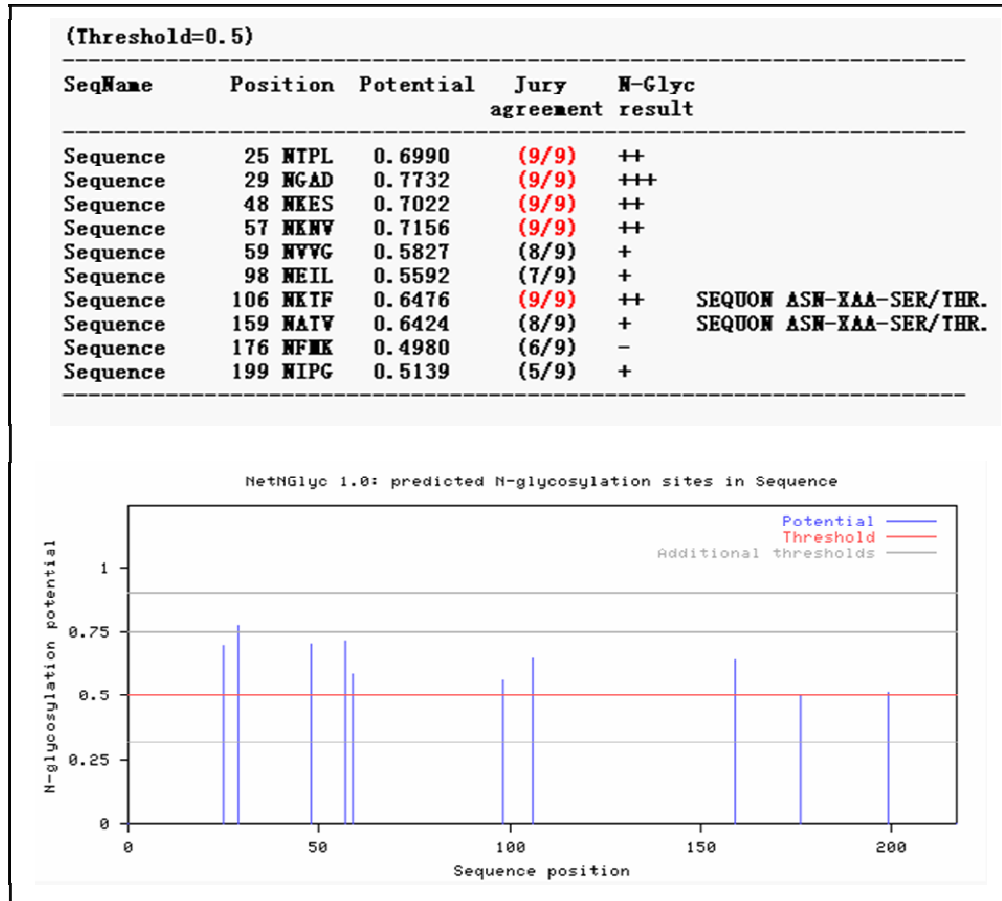


Figure 14. Oligosaccharides attached to Asn residues of secreted or membrane bound proteins are described as *N*-linked. The sequence motif Asn-Xaa-Ser/Thr (Xaa is any amino acid except Pro) has been defined as a prerequisite for *N*-glycosylation. Using online software for predicting the post-translational modifications of proteins, NetNGlyc 1.0 Server (<http://www.cbs.dtu.dk/services/NetNGlyc>), multiple asparagine residues in the TTMP protein that were likely to be *N*-glycosylated were identified. Of note, Asn25, Asn29, Asn48 and Asn57 have the highest score for *N*-glycosylation.

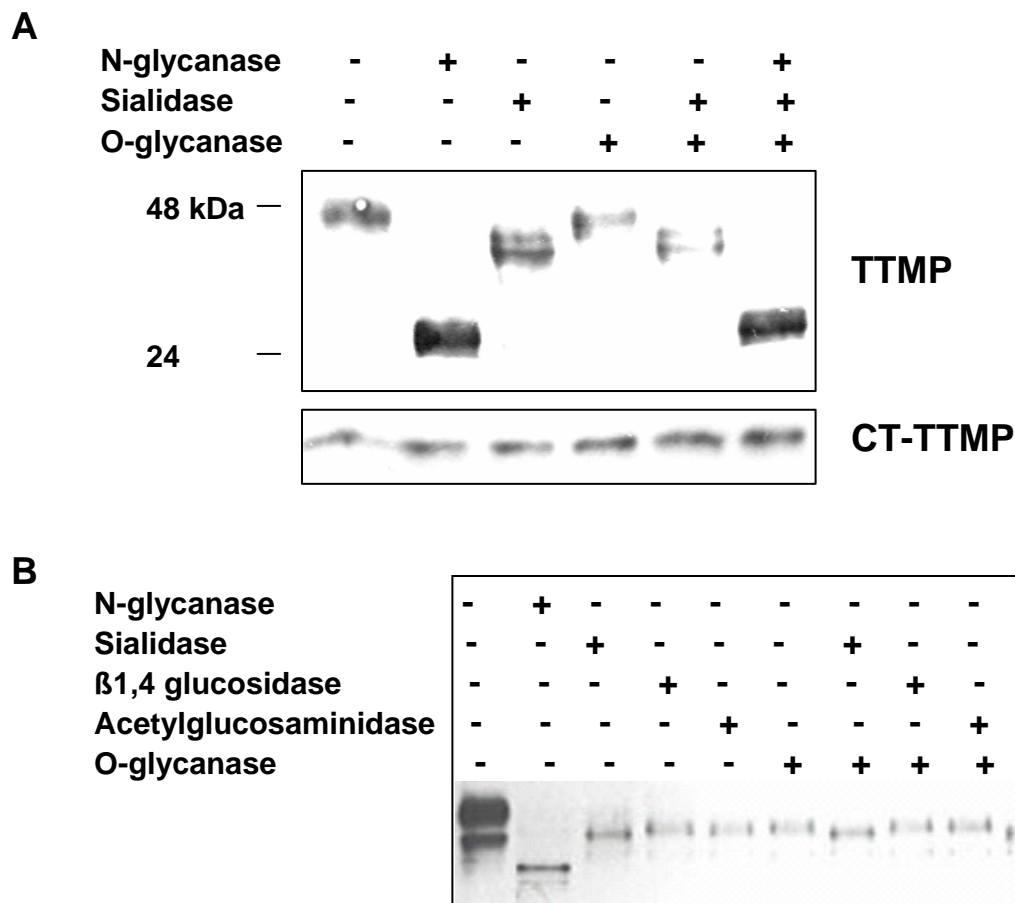


Figure 15. Glycosylation pattern of the TTMP protein: (A) Effect of digestion with N-glycanase, sialidase and O-glycanase or their combination on TTMP migration in a 12% SDS-PAGE gel. N-glycanase treatment reduced the size of full-length TTMP from 48 kDa to 24 kDa. O-glycanase treatment had no effect on migration of TTMP, while sialidase alone slightly reduced the size of TTMP. Combined digestion with the three enzymes reduced the size of TTMP to 24 kDa, the same as N-glycanase alone. None of the three enzymes individually or in combination altered the migration of the C-terminal fragment of the TTMP protein, indicating that TTMP is N-glycosylated in the N-terminal region. (B) The effect of digestion N-glycanase, sialidase, beta 1,4-glucosidase, acetylglucosaminidase or O-glycanase on TTMP migration in 12% SDS-PAGE gel. N-glycanase reduces molecular size of TTMP from 48 kDa to 24 kDa. None of the other enzymes changes the shift of TTMP except that sialidase slightly increases the migration of TTMP. Taken together, it was concluded that TTMP was N-glycosylated at its N-terminal end.

4.2.4 TTMP Localizes to the Endoplasmic Reticulum

Topological analysis of TTMP using two different web-based programs (<http://psort.nibb.ac.jp> and <http://www.cbs.dtu.dk/services/TMHMM>) predicted that TTMP is a single pass transmembrane protein localized to the endoplasmic reticulum (ER) (Fig

16A) (204). We initially used immunocytochemistry to study the subcellular localization of TTMP *in vitro*. HeLa cells transiently transfected with TTMPpcDNA3.1 were fixed and probed with V5-antibody. HRP staining of secondary antibody to V5-antibody showed that TTMP was localized to the perinuclear region in an eccentric fashion. This is in concordance with the in-silico prediction of TTMP's subcellular location. The ER localization of TTMP was further investigated by immunofluorescence co-localization experiments using laser scanning confocal microscopy. Co-immunofluorescence staining with the ER-specific protein disulphide isomerase (PDI) confirmed that TTMP is co-localized with the specific ER marker (Fig 16B). Furthermore, the ER localization was preserved when the truncated TTMP expression vector (CT-TTMPpcDNA3.1) which expresses the C-terminal end of TTMP with the transmembrane domain was transfected (Fig 16C). However, the C-terminal end of TTMP without transmembrane domain does not co-localize with the specific ER marker as seen by transfecting CT^{TMminus}-TTMPpcDNA3.1 expression plasmid into cells (Fig 16D).

4.2.5 TTMP Inhibits Proliferation of Pancreatic Cancer Cells

In screening studies for possible effects on cell growth dynamics of TTMP, we have observed that HeLa cells transiently transfected with TTMPpcDNA3.1 grow slower as compared to control cells transfected with the empty vector. Hence we investigated the growth effects of TTMP transfection in human pancreatic cancer cells. As pancreatic cancer cells have low transient transfection efficiencies, we first established CD18 pancreatic cancer cell clones with stable expression of TTMP, and identified multiple positive clones (Fig 17A). When observed under light microscopy, TTMP stable expression CD18 cell clones proliferate slower and do not have the same propensity to form large multi-cell adherent colonies as do native CD18 and empty vector-transfected cells (Fig 17B). This observation was also evident when the different colonies growing on

6-well plates were stained with crystal violet and photographed (Fig 17C). Furthermore, CD18 cells with stable expression of TTMP grow much slower as compared to control native and empty vector transfected cells when grown in collagen I gels (Fig 18).

The growth effects were then investigated in a quantitative manner. Cell proliferation assays was conducted using an automated cell counter. Equal numbers of TTMP stable expression cells, empty vector transfected cells, or native CD18 cells were seeded into 12-well plates and grown in serum-free conditions. Forced expression of TTMP resulted in a dramatic growth inhibition over the four days of study compared to empty vector-transfected cells or parental CD18 control cells (Fig 19A). By the end of 4 days, the decrease in the number of cells was more than 4-fold (clone 208 or clone 210 vs control CD18 cells or pcDNA3.1 empty vector-transfected cells at day 2: $P < 0.05$; clone 208 or clone 210 vs control CD18 cells or pcDNA3.1 empty vector-transfected cells at day 3 and 4: $P < 0.001$).

It could be argued that forced expression via the constitutionally active CMV promoter found on pcDNA3.1 could lead to supra-physiological levels of TTMP which may not truly reflect the nature of TTMP's action on the cell. Hence another strategy which we employed to quantify the growth inhibitory effect of TTMP was to use siRNA to silence the expression of endogenous TTMP and then investigate the effect on cell growth. A mixture of siRNA duplexes targeting to TTMP was able to effectively suppress the expression of TTMP by more than 50% compared with expression in control cells transfected with scrambled siRNA (Fig 19B, scrambled siRNA vs TTMP siRNA: $P < 0.05$). Cell counting assays, performed in parallel revealed a 2-fold increase in the number of TTMP siRNA transfected cells as compared to control (Fig 19B), scrambled siRNA vs

TTMP siRNA: $P < 0.05$). This finding further strengthens the evidence that TTMP inhibits growth of pancreatic cancer cells.

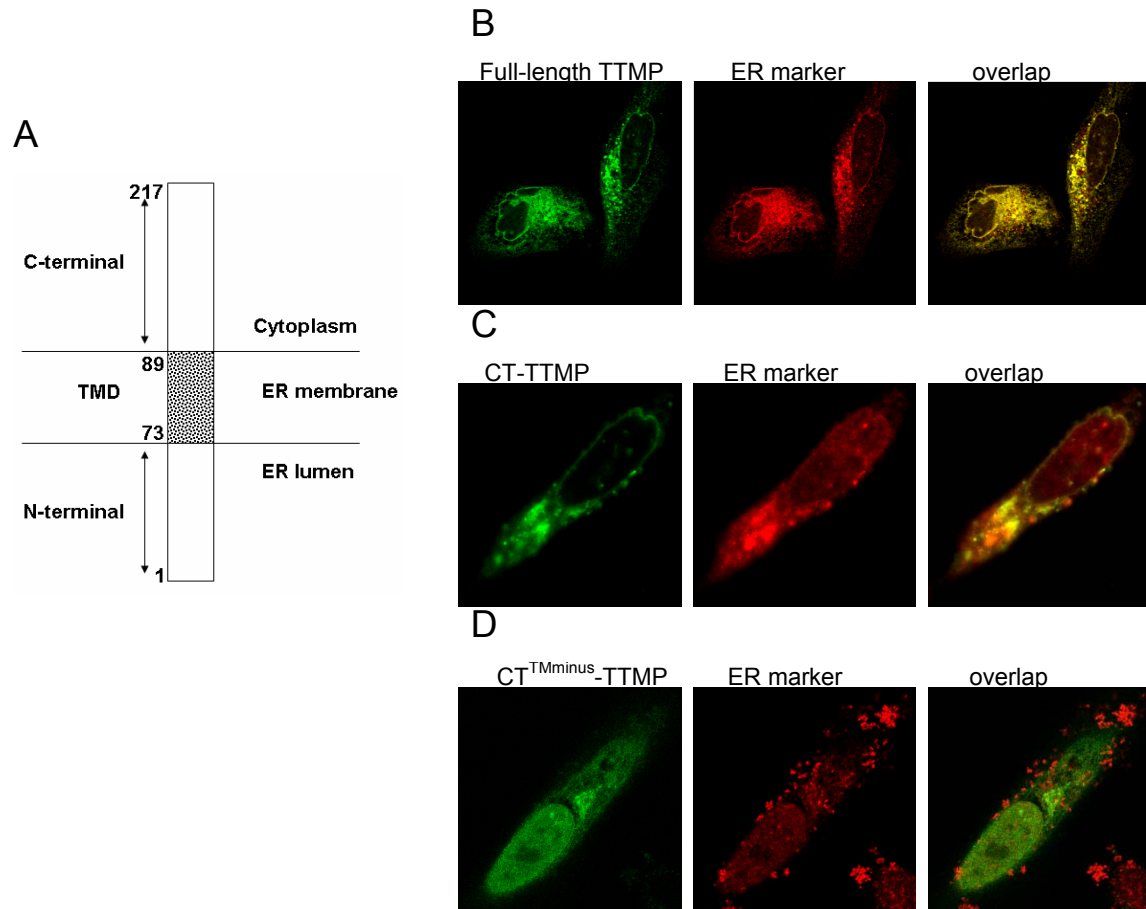


Figure 16. Immunofluorescence localization of TTMP in HeLa cells. HeLa cells were transfected with either pcDNA3.1 control plasmid, TTMPpcDNA3.1 plasmid, CT-TTMPpcDNA3.1 plasmid, or CT^{TMminus}-TTMPpcDNA3.1 plasmid for 24 hours. The cells were fixed with paraformaldehyde, with 0.2% Triton X-100, and co-stained with a V5 antibody (green) and an endoplasmic reticulum-specific marker, protein disulphide isomerase (PDI) antibody (red). Localization of TTMP and protein disulphide isomerase was visualized by confocal fluorescence microscopy. (A) Diagram depicting TTMP as an ER transmembrane protein with its N-terminal end in ER lumen and C-terminal end at cytoplasm. Immunofluorescence staining shows that TTMP is co-localized with protein disulphide isomerase (B), the C-terminal truncated TTMP with transmembrane domain (CT-TTMP) is co-localized with protein disulphide isomerase (C), and the C-terminal truncated TTMP without transmembrane domain (CT^{TMminus}-TTMP) is not co-localized with protein disulphide isomerase (D).

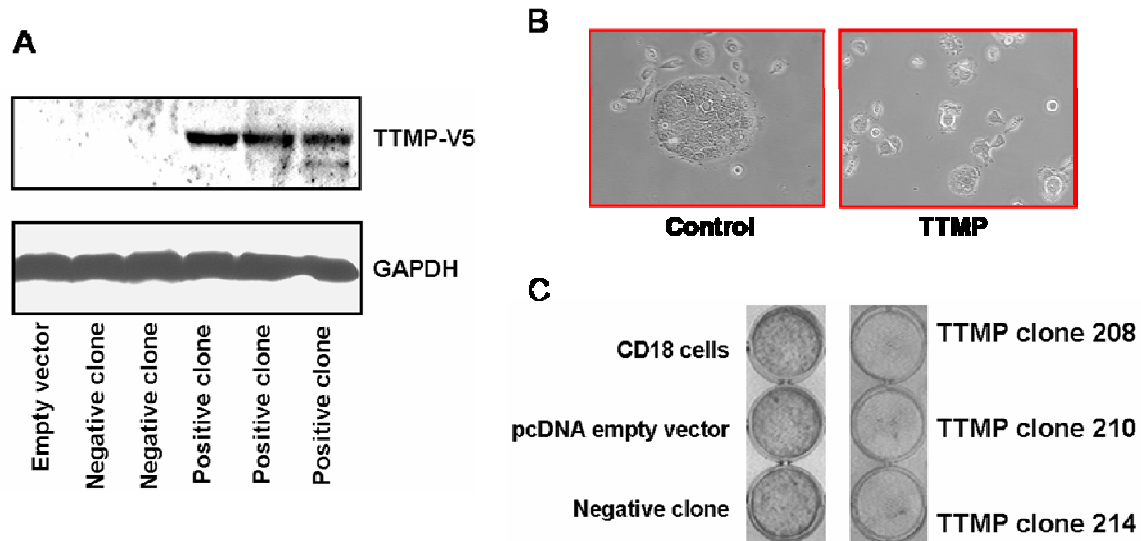


Figure 17. Effect of forced expression of TTMP on cell proliferation in CD18 pancreatic cancer cells. (A) CD18 cells were plated in T25 flask and transfected with either the TTMP pcDNA3.1 expression plasmid or the control pcDNA3.1 plasmid. The transfected cells were selected with G418 for 3 weeks. The G418 resistant clones were picked and TTMP-V5 expression in the stable clones was measured by western blotting. (B) Forced expression of TTMP inhibited CD18 cell proliferation and induced a marked morphological change with a lack of colony formation. (C) Equal amount of cells were plated into 12-well plates. The cells were grown for 4 days and stained with crystal violet. Compared with control CD18 cells or a TTMP negative clone, forced expression of TTMP significantly decreases cell proliferation as evidenced by cell density.

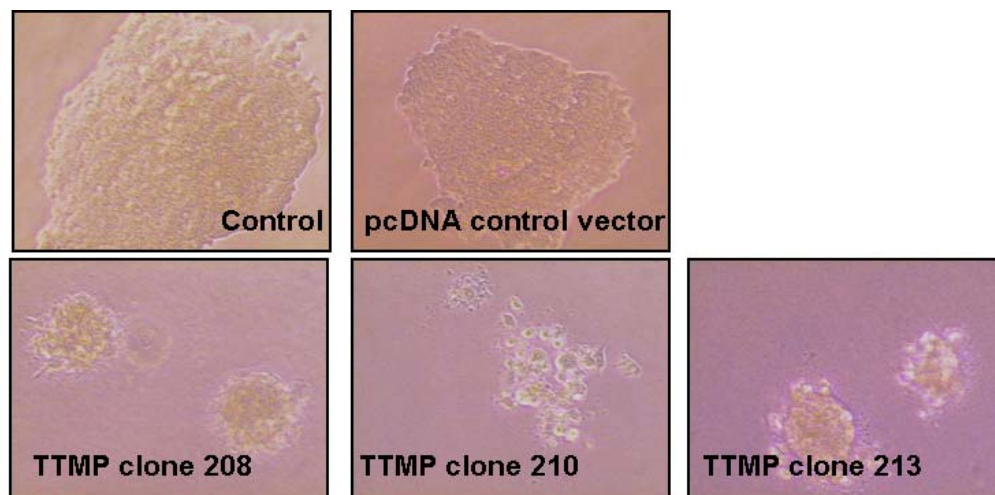


Figure 18. Effect of forced expression of TTMP on cell proliferation of CD18 cells in three-dimensional collagen 1 gels. 1500 CD18 cells were mixed into 0.5 ml of collagen 1 gel. The gels were plated into 24-well plates and fed with DMEM media containing 10% FBS. Pictures were taken 2 weeks following culture. Compared to control CD18 cells,

forced expression of TTMP significantly slowed down cell proliferation in the three-dimensional collagen 1 gels.

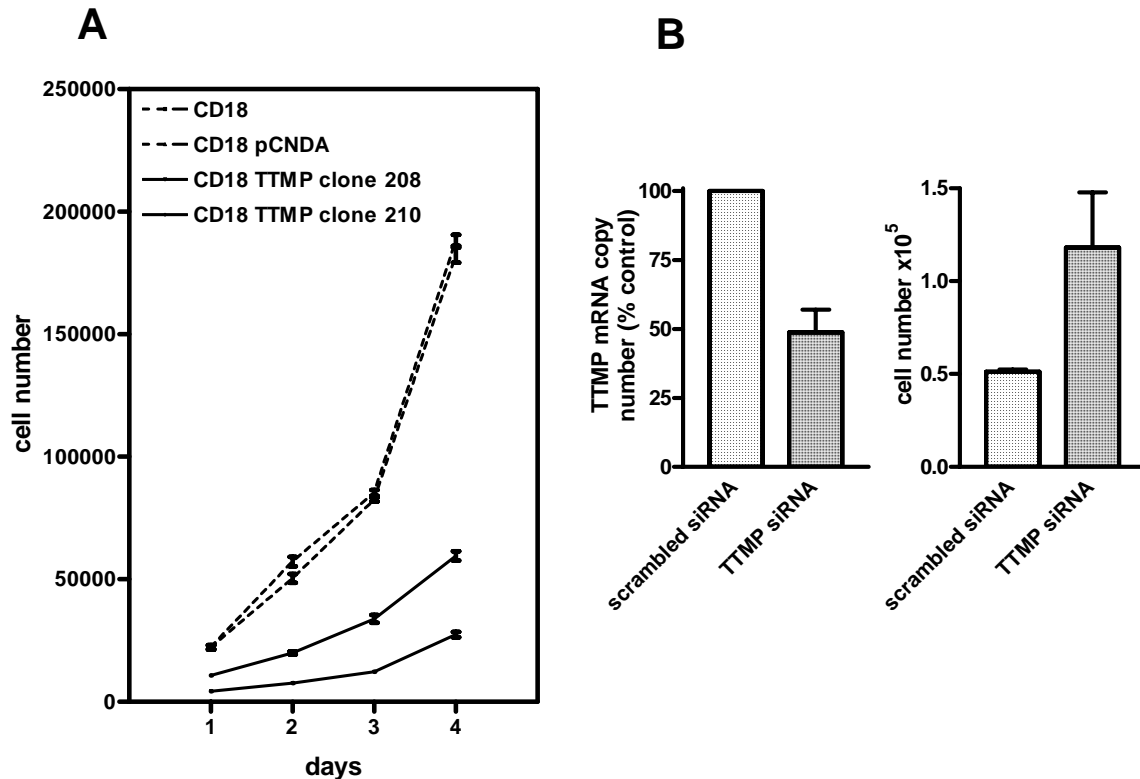


Figure 19. (A) Forced expression of TTMP inhibits cell proliferation in CD18 cells. CD18 control cells and TTMP stable expression clones were plated into 12-well plates. The cells were grown in DMEM and cell number was counted every 24 hours. Compared with control cells, forced expression of TTMP dramatically inhibited cell proliferation as measured by cell number. (B) siRNA duplexes targeted to TTMP enhance CD18 cell proliferation. TTMP siRNA duplexes decreased copy number of TTMP compared with scrambled siRNA transfection. Transfection with TTMP siRNA duplexes increased CD18 cell number compared with the control transfected cells.

4.2.6 CT-TTMP, an In-Frame N-Terminal Truncation of TTMP Enhances Pancreatic Cancer Cell Growth

As mentioned above, we examined the localization characteristics of CT-TTMP, an in-frame N-terminal truncation of TTMP, in parallel with the experiments conducted with the full-length protein. We next examined the functional differences between full-length TTMP and C-terminal end of TTMP with the transmembrane domain, following

establishment of CD18 cell clones that stably express the C-terminal end of TTMP by transfecting the CT-TTMPpcDNA3.1 plasmid and G418 selection (Fig 20A). The two clones (clone 6 and 18) with the highest expression of CT-TTMP were used for subsequent cell proliferation studies. Interestingly, forced expression of CT-TTMP stimulated pancreatic cancer cell growth, in contrast to the growth inhibition seen in response to transfection with full-length TTMP (Fig 20B). At the end of 4 days, the fastest growing clone exhibited a more than 2-fold increase in the number of cells compared with all 3 controls (Fig 20B, control CD18, pcDNA empty vector transfected CD18 cells, or the negative expression clone vs clone 6 or clone 18 at 4 days: $P < 0.01$; control CD18, pcDNA empty vector transfected CD18 cells, or the negative expression clone vs clone 6 or clone 18 at 3 days: $P < 0.05$).

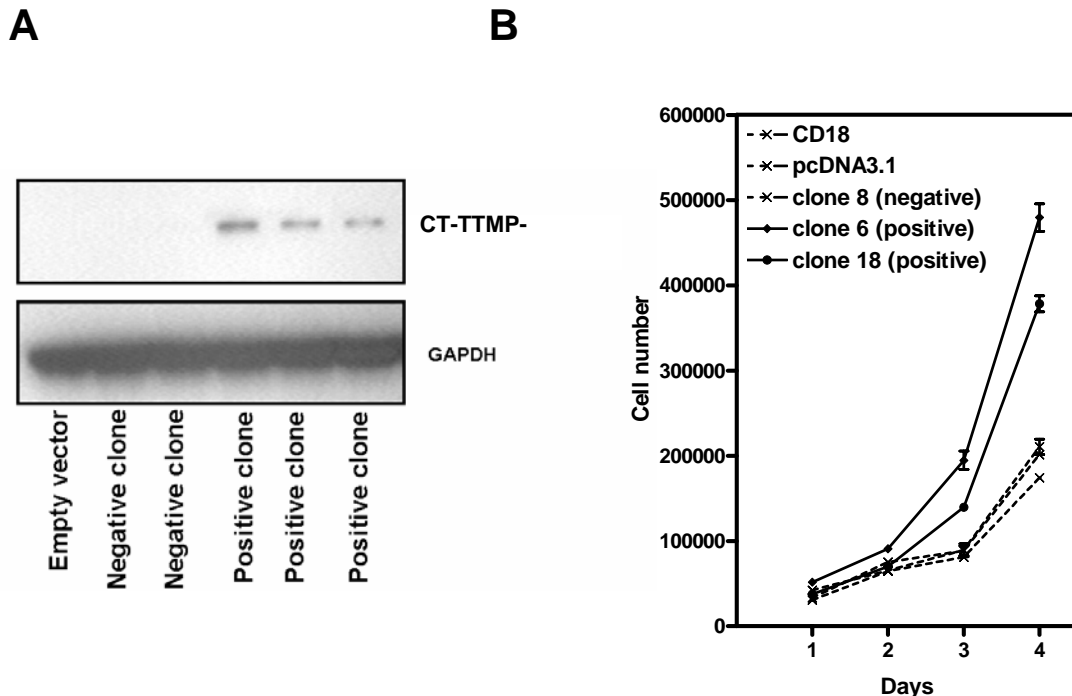


Figure 20. Effect of forced expression of the C-terminal fragment of TTMP on cell proliferation in CD18 pancreatic cancer cells. (A) CD18 cells were plated in T25 flasks and transfected with either the CT-TTMP pcDNA3.1 expression plasmid or the control pcDNA3.1 plasmid. The transfected cells were selected with G418 for 3 weeks. The G418 resistant clones were picked and CT-TTMP-V5 expression in the stable clones was measured by western blotting. Multiple clones stably expressing CT-TTMP were

identified by western blot. (B) CD18 control cells and CT-TTMP stable expression clones were plated into 12-well plates. The cells were grown in DMEM and cell number was counted every 24 hours. Compared to control cells, forced expression of CT-TTMP markedly increased cell proliferation, measured by cell number.

4.2.7 Forced Expression of TTMP Induces G₁ Phase Growth Arrest in CD18 Pancreatic Cancer Cells

We next investigated the effect of TTMP on cell cycle changes in pancreatic cancer cells. Cell cycle analysis was compared in CD18 cells with stable expression of TTMP and empty vector-transfected cells. Compared to control, a larger proportion of TTMP-expressing cells were in the G₀/G₁ phase, indicating G₁ phase growth arrest in these cells (Fig 21).

4.2.8 Forced Expression of TTMP Inhibits HeLa Cell Proliferation

To further confirm that forced expression of TTMP inhibits cell proliferation, we also looked at effect of TTMP on proliferation of another cancer cell line, HeLa cells. Stable TTMP expression clones of HeLa cells were established following transfection of TTMP expression vector and G418 selection. Forced expression of TTMP also significantly inhibited proliferation of HeLa cells, assessed by cell number in regular cell culture, further suggesting that TTMP negatively regulates cell proliferation (Fig 22).

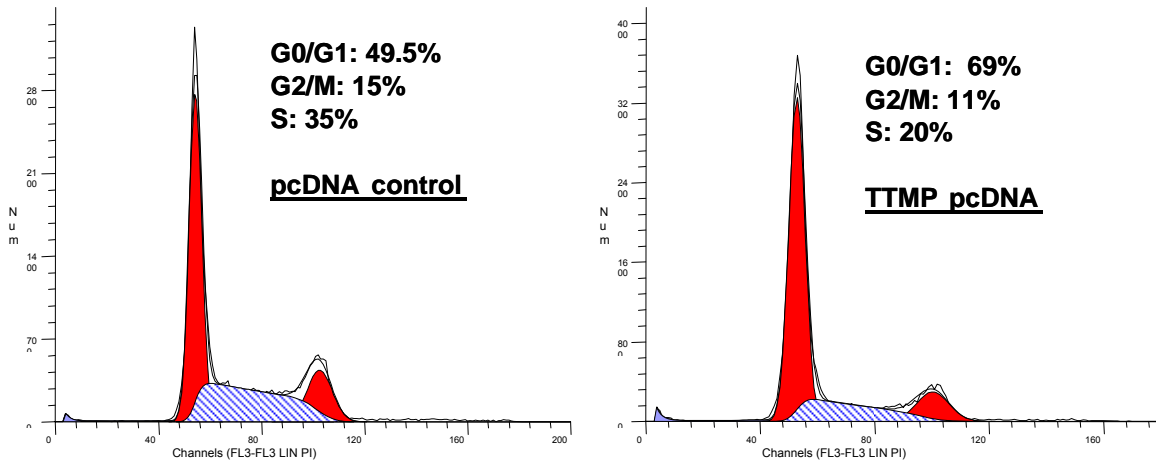


Figure 21. Forced expression of TTMP causes G0/G1 phase cell cycle arrest in CD18 pancreatic cancer cells. Both pcDNA3.1 empty vector-transfected cells and TTMP stable expression cells were trypsinized and fixed in 70% alcohol for 15 minutes. The cells were stained with propidium iodide in PBS for 2 hours. The DNA content of the cells was assessed by flow cytometry. Forced expression of TTMP significantly increased the G0/G1 cell population compared to that of control.

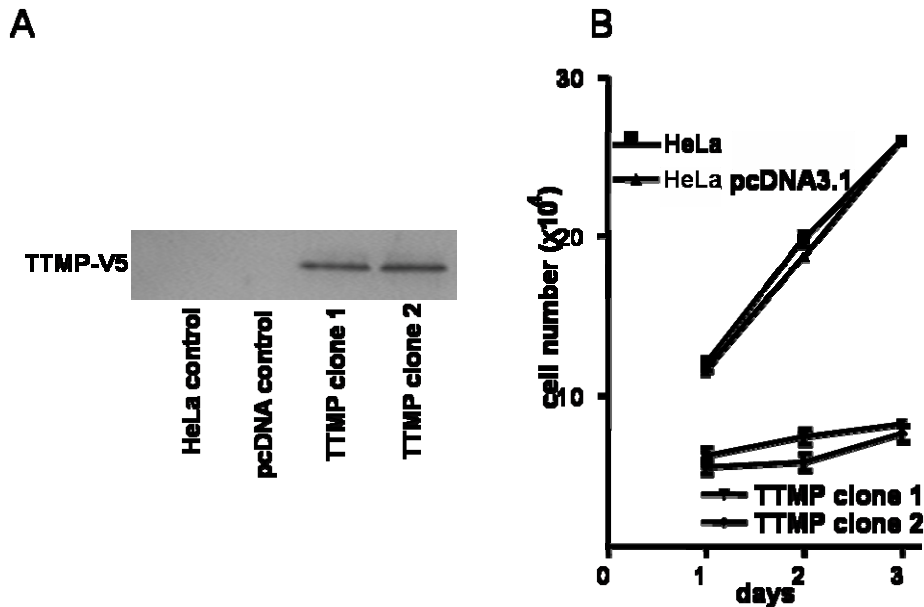


Figure 22 Effect of TTMP on HeLa cell proliferation. (A) HeLa cells were plated in T25 flask and transfected with either the TTMP pcDNA3.1 expression plasmid or the control pcDNA3.1 plasmid. The transfected cells were selected with G418 for 3 weeks. The G418 resistant clones were picked and TTMP-V5 expression in the stable clones was measured by western blotting. (B) HeLa control cells and TTMP stable expression clones were plated into 12-well plates. The cells were grown in DMEM and cell number was counted every 24 hours. Compared to control cells, forced expression of TTMP dramatically inhibited cell proliferation, measured by cell number.

4.2.9 Conclusion

In conclusion, we have shown that TTMP is highly expressed in human pancreatic tissue, but has a low expression in both pancreatic cancer cell line CD18 and the cervical carcinoma cell line, HeLa. TTMP undergoes post-translational modification via N-glycosylation and addition of sialic acid moieties, and is localized to the endoplasmic reticulum within the cell. We have also demonstrated that TTMP has growth inhibitory effects in pancreatic cancer cells by inducing G₁ phase growth arrest.

4.3 TRANSCRIPTIONAL REGULATION OF TTMP PROMOTER

The final aim of the study was to investigate the basal transcriptional activity of TTMP.

4.3.1 Sequence Analysis of the 5'-Flanking Region of TTMP

We have previously performed 5'-RACE to identify the transcription start site of TTMP. In order to assess the promoter activity of the 5'-flanking region of the TTMP gene, we analyzed up to 2-kb genomic NT_005612.14 sequence upstream of the determined transcription start site. Computer analysis was performed using online software TFSearch (<http://www.cbrc.jp/research/db/TFSEARCH.html>). Putative binding sites for transcription factors deltaE (core similarity 0.88), p300 at nucleotide position -150 to -138 (core similarity 0.91), Sp1 at nucleotide position -77 to -68 (core similarity 0.89), a second p300 site at nucleotide position -65 to -52 (core similarity 0.88), an overlapping Sp1/Mzf1 at nucleotide position -54 to -45 (core similarity 0.92/ 0.93), and an USF site (core similarity 0.87) were identified in the proximal sequence, as shown in Fig 23. There were neither TATA box nor CAAT box found in this region, and the proximal sequence had a high GC content. This analysis suggests that a TATA-less promoter is

used for TTMP expression, which is likely to be regulated by multiple transcription factors.

4.3.2 Functional Characterization of the TTMP Promoter

To localize the DNA elements that are important for promoter activity, a series of unidirectional deletion analysis of the 2-kb 5'-flanking region of the TTMP gene was carried out. Deletion fragments having 5' ends ranging from -1909 to -219 and 3' ends at +95 were generated by PCR and cloned into the promoterless pGL₃-basic, a luciferase reporter vector. Each resulting recombinant construct was then transiently transfected into human cervical carcinoma cell line HeLa. After 48 hours, cell extracts were prepared and luciferase activity was measured. The resulting luciferase reporter gene activities were then normalized to protein concentration of the individual lysates. As shown in Fig 24A, deletion from -1909 to -766 led to a reduction in luciferase activity from 20.2 fold to 13.7 fold. This would indicate the presence of enhancer elements in the region -766 to -1909. Similarly, the increase in luciferase activity seen when the constructs were deleted from -284 to -219 would indicate the presence of suppressor elements in that region. Of note, the construct -219/+95 exhibited a luciferase activity of 18.5 fold over that of the promoterless pGL₃-basic vector. This would be consistent with the definition of a proximal promoter being within 200 nucleotides flanking the transcription start site in both directions i.e. upstream and downstream.

To determine the functional importance of each potential transcription factor binding site that was located between -219 and +95 for promoter activity, further progressive deletion constructs were generated. Deletion from nucleotides -219 to +18 caused a reduction of luciferase activity from 23.5 fold of basal activity of the pGL₃ basic vector, to almost that of the basal activity of pGL₃ basic vector alone (Fig 24B).

Remarkably, deletion from nucleotides -135 to -69, which removed a putative Sp1 binding site, and deletion from nucleotides -69 to +18, which removed a putative p300 site and a putative Mzf1/Sp1 binding site, reduced the promoter activity by 46% and 85% respectively (Fig 24B). This would suggest that the putative transcription factor binding sites for Sp1, p300 and the overlapping Mzf1/Sp1 located between nucleotides -135 to +18 may contribute to the basal promoter activity.

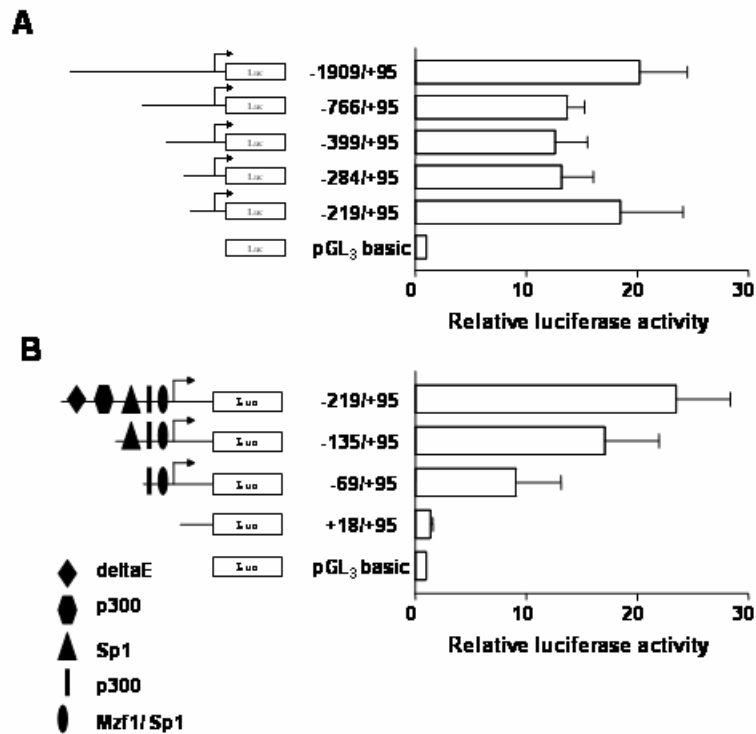


Figure 24. Deletion analysis of the 5' flanking region of the hTTMP gene. Progressive deletions of the hTTMP 5' region was amplified by PCR and then ligated to the luciferase reporter gene. The 3' terminus of each construct is nucleotide +95 relative to the transcription start site (TSS). The arrow marks the transcription start site (position +1). 0.25 μ g of each luciferase reporter construct with 1 μ l of Metafectane was transfected into HeLa cells. Cells were lysed after 48 hours post-transfection. The luciferase activity was measured and normalized to protein concentration. The normalized activity was expressed as fold change relative to the activity of the empty vector pGL₃ basic which is set at 1 and is presented as the mean \pm S.E. of three independent experiments. (A) Deletion analysis of the 5' flanking region spanning nucleotide -1909 to -219 relative to the TSS. (B) Deletion analysis of the proximal promoter region spanning nucleotide -219 to +18 relative to the TSS. The locations of putative transcription factor binding sites on the proximal promoter region are indicated.

4.3.3 Site-Directed Mutagenic Analysis of the Putative Transcription Factor

Binding Sites Responsible for Basal Promoter Activity of TTMP

To additionally confirm the involvement of the putative Sp1 and overlapping Mzf1/Sp1 transcription factor binding sites in the basal promoter activity of TTMP, we introduced mutations into these sites using site-directed mutagenesis as described

above. The mutated constructs were then transfected into HeLa cells and the luciferase activities determined. As shown in Fig 25, individual mutations in the Sp1 site as well as the Mzf1/Sp1 site led to a dramatic reduction in luciferase activities as compared to the wild-type -135/+95 construct. This reduction was even more pronounced when mutations were introduced into both sites on the same construct. The shorter -69/+95 construct did not contain the first Sp1 site. When the overlapping Mzf1/Sp1 site was mutated on this construct, again this led to a reduction in luciferase activity when compared to control. Putting these data together, we can conclude that both the Sp1 site as well as the Mzf1/Sp1 site plays a role in the high basal activity of the region spanning nucleotides -135 to +95.

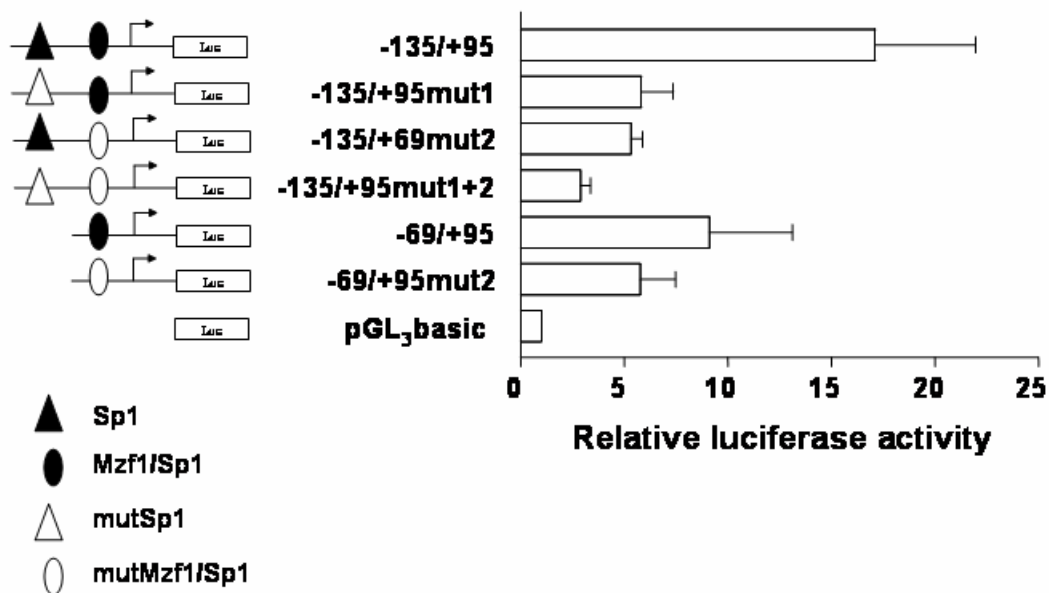


Figure 25. Mutational analysis of the proximal promoter region of the hTTMP gene. Site directed mutagenesis of the putative transcriptional factor binding sites was performed using PCR based technology. Transfection protocol and measurement of luciferase activity were as described in Fig 2. The wild type transcriptional factor binding sites are represented by the blackened symbols and the mutated transcriptional factor binding sites

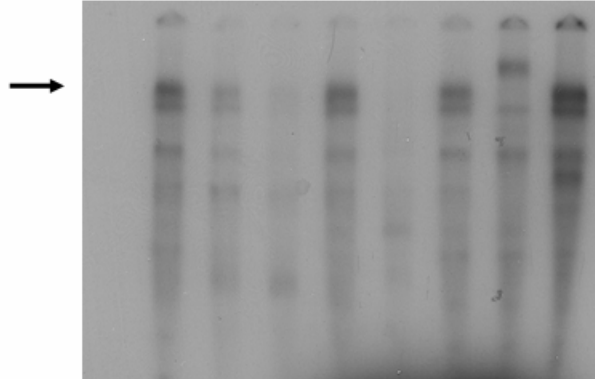
are represented by the white symbols. Mut1 represents mutation of the Sp1 site, mut2 represents mutation of the overlapping Mzf1/Sp1 site.

4.3.4 Electrophoretic Mobility Shift Analyses of Physical Binding of Transcription Factor Sp1 to Putative Cis-Elements on TTMP Promoter

We next sought to determine the physical interaction between trans-acting transcriptional factor Sp1 and the putative transcriptional factor binding sites identified above that conferred high basal promoter activity on functional studies. The wild-type sequences -77 to -68 and -54 to -45 which contained the putative Sp1 binding site (site 1) and overlapping Mzf1/Sp1 binding site (site 2) were used as probes, and together with nuclear extracts from HeLa cells, was analyzed for DNA-protein complexes formation. At least 4 complexes were formed with oligonucleotide sequence -77 to -68 (Fig 26A, lane ii). The specificity of these complexes was shown by competition assay, in which the highest band was dramatically reduced by the addition of unlabelled wild type oligonucleotides (lane iii and iv), as well as unlabelled Sp1 consensus sequence (lane vi). This competition is not seen with the addition of mutant native sequence (lane v) as well as a non-specific sequence (lane vii). To characterize these complexes further, supershift EMSA was conducted using Sp1 antibody. The result clearly identified a supershifted band upon addition of the Sp1 antibody (lane viii), but not with a non-specific antibody (lane ix). This experiment was next conducted with the nucleotide sequence -54 to -45 encompassing the overlapping Mzf1/Sp1 binding site. We see in fig 4B that the band indicated was similarly competed out by the addition of unlabelled wild type oligonucleotide, as well as both consensus Sp1 and Mzf1 sequences, but not the mutant native sequence and non-specific sequence (Fig 26B, lane ii to viii). Thus we conclude that the putative Sp1 and Mzf/Sp1 binding sites on TTMP promoter behave similarly to the consensus Sp1 and Mzf1 binding sequences.

A

		← Sp1(site 1) →								
Probe										
Supershift	-	-	-	-	-	-	-	Sp1	NSp	
Competitor	-	-	WT	WT	Mut	Con	NSp	-	-	
			25x	100x		Sp1				
		i	ii	iii	iv	v	vi	vii	viii	ix



WT : 5' ACGCCCCGCCAGTG 3'

Mutant : 5' ACGCCCATACCAGTG 3'

Sp1 Consensus : 5' TCGACGGGGCGGGGCTTA 3'

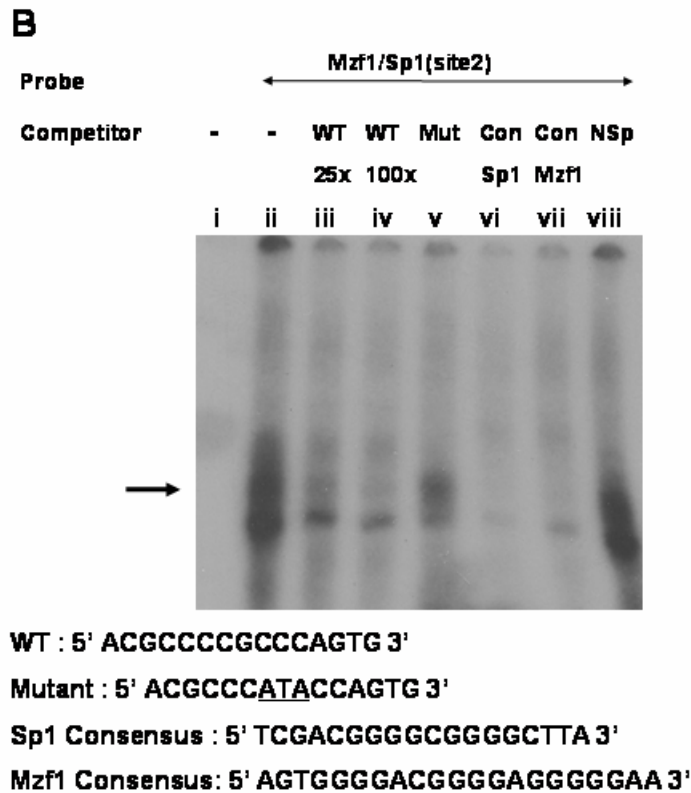


Figure 26. Electrophoretic mobility shift analysis of nuclear protein interactions with DNA fragments derived from the hTTMP proximal promoter. Oligonucleotides containing the putative transcription factor binding sites (A) Sp1 (nucleotide -77 to -68) and (B) Mzf1/Sp1 (nucleotide -54 to -45) were radiolabelled and used as probes to incubate with nuclear protein extracts obtained from HeLa cells in EMSA. Lane (i) of both experiments represents the control lane without addition of radioactive probe. Lane (ii) of both experiments represents the radiolabelled probes without competition. Unlabelled competitor wild type oligonucleotides were present in 25-fold and 100-fold excess (lane iii and iv), and unlabelled mutant, consensus and non-specific oligonucleotides (lane v, vi and vii for fig 26A; lane v, vi, vii and viii for fig 26B) were present in 100-fold excess in competition reactions. Anti-Sp1 antibody was used for supershifting (fig 26A lane viii), and non-specific antibody was used in lane ix. The arrows point to specific complexes that could be competitively blocked in the presence of unlabelled specific probes or unlabelled oligonucleotides containing consensus sequences for binding of transcription factors. Figures shown are representative of 3 separate experiments respectively.

4.3.5 Conclusion

In conclusion, we have cloned the 5' flanking region of TTMP, as well as identified the core promoter responsible for basal activity of the gene. Using mutational

analysis and EMSA, we have demonstrated that Sp1 binds to its counterpart cis-elements on the core promoter and activates transcription.

5. DISCUSSION, CONCLUSIONS AND FUTURE WORK

The effects of TPA and the role of PKC in pancreatic cancer are mixed and contrasting. Our own observations concurred with inhibition of pancreatic cancer cell growth following stimulation with TPA and subsequent activation of PKC (142,143). From the central and often diametric roles that PKC plays in cancer cell growth dynamics, we hypothesized that novel genes that play a causal role in this process would be identified by differential expression analysis of cancer cells following phorbol ester treatment.

TTMP is a novel TPA inducible gene. Using oligonucleotide microarray technology, we have identified a novel TPA-inducible gene which codes for a putative transmembrane protein which we called TPA-induced Trans-Membrane Protein (TTMP) (204). The full length cDNA of TTMP is comprised of 2252 base pairs of nucleotides. The gene was mapped to human chromosome 3q13.2 with 6 exons and 5 introns. A GeneBank search revealed several homologous expressed sequence tags (ESTs) from many different human tissues and organisms for this novel gene, suggesting that the gene is ubiquitously expressed in higher eukaryotes.

TTMP is highly expressed in normal pancreas tissue but is under-expressed in pancreatic and cervical cancer cells. TTMP was originally identified by oligonucleotide microarray analysis of pancreatic cancer cells treated with TPA. We examined the distribution of TTMP in normal human tissue and interestingly, normal pancreatic tissue had the highest expression of TTMP compared to the other tissues studied. This suggests that TTMP may play a role in maintaining normal pancreatic cellular homeostasis and function. More importantly, we see that the expression of TTMP is substantially reduced in pancreatic cancer cells and cervical carcinoma cells, compared to normal pancreatic tissue. This suggests that expression of TTMP may be

lost during the development of pancreatic cancer, underlying the potential anti-mitogenic nature of TTMP.

TTMP is a transmembrane protein that undergoes post-translational glycosylation. Computational analysis predicted that TTMP contains a single transmembrane domain and is localized on the ER membrane. This is confirmed by co-immunofluorescence staining with an antibody targeting to a specific ER membrane protein, protein disulfide isomerase. Interestingly, we found that ER localization of TTMP is dependent on the single transmembrane domain of TTMP, since the TTMP protein with deletion of this domain is not co-localized with the ER marker. Immunoblotting of TTMP shows that the molecular size of TTMP from cell extracts is 48 kDa, which is twice the size of the predicted protein product, suggesting that the protein undergoes post-translational modification. From our deglycosylation assays, we have demonstrated that the protein is N- but not O-glycosylated with complex oligosaccharides, including some sialic acid residues. Post-translational modifications are critical for the regulating functions of both intracellular and extracellular proteins. It has been known that glycosylation could alter proteolytic resistance, protein solubility, stability, local structure and lifespan of proteins (206-208). In the current study, we found that TTMP is heavily N-glycosylated at its N-terminal end but not the C-terminal end. Even though the functional effect of N-glycosylation on TTMP needs to be evaluated in future studies, we speculate that N-glycosylation might play an important role in TTMP-regulated cell proliferation because the truncated TTMP with deletion of its heavily N-glycosylated N-terminal end generates an opposite biological effect to that of the full-length TTMP in pancreatic cancer cells.

TTMP inhibits pancreatic cancer cell proliferation and induces a G₁ phase cell-cycle arrest. Possible role of TTMP in the UPR pathway. Even though it is not common that proteins localized on ER membrane are involved in cell division, multiple studies have shown that ER proteins can be involved in cell proliferation or apoptosis. For example, it has been shown that Ca⁺⁺ homeostasis endoplasmic reticulum protein (CHERP) regulates cellular DNA synthesis through Ca⁺⁺ homeostasis (209,210). Mediation of cellular apoptosis by ER proteins has also been well documented (211,212). The link between the unfolded protein response (UPR) and cancer has been a subject of much interest recently (213). UPR is a reaction to stress in the endoplasmic reticulum. An accumulation of unfolded or misfolded proteins within the ER, as well as outside stresses like nutrient and oxygen deprivation, trigger the UPR, leading to transcription of proteins in the nucleus that help cells cope with the stress. The UPR has both cytotoxic functions as well as cytoprotective ones. UPR activation can result in one of two outcomes: either regulated cell death triggered by apoptotic effectors or survival of the stress facilitated by beneficial UPR target genes. Prolonged activation of UPR results in decreased cellular proliferation from a cell cycle arrest in G₁ phase secondary to a decrease in translation of cyclin D1, and preventing cells from progressing through the cell cycle before ER homeostasis is re-established (214,215). This delay may allow a cell to pause in the cell cycle to determine whether adaptation to stressful conditions will be possible, and if not, to continue on toward apoptosis (216). Hypoxia is a common feature of solid tumours, notably pancreatic cancer, that display increased malignancy, resistance to therapy, and poor prognosis. Tumour cells need to adapt to the increasing hypoxic environment that surrounds them as they grow, and induction of the UPR is key to this response (217). The focus of this thesis is on structural and functional characterization of the novel gene TTMP, and little work has been done to dissect the molecular pathways acting upstream and downstream of TTMP. Motif scanning analysis

of TTMP did not find sequence homology with any conserved functional domain. Hence at this juncture, the role of TTMP in growth regulation and its mechanism of action can only be speculative at best. Effector genes of the UPR pathway has been found to be highly expressed in tissues that specialize in secretion such as the pancreas, salivary gland, and chondrocytes (218,219). A viable hypothesis is that the high expression of TTMP seen in normal pancreatic tissue represents a role for TTMP as a novel player in the UPR pathway in maintaining normal homeostasis of the pancreas as a secretory organ. Similar to known mediators of the UPR, namely IRE1, PERK and ATF6, TTMP is localized to the endoplasmic reticulum, has a single transmembrane domain and is N-glycosylated (220,221). In addition, full length TTMP inhibited pancreatic cancer cell growth and induces a G₁ phase growth arrest in pancreatic cancer cells, a phenomenon similar to the cellular effects of the known mediators of the UPR in other cell-types. Interestingly, the N-terminal truncated protein (CT-TTMP) induced cell proliferation, in contrast to the inhibition of cell proliferation seen with the full-length protein. This could be due to the absence of functional domains residing on the N-terminal of the protein, or due to lost of glycosylation of TTMP, or that the truncated protein behaves as a dominant negative mutant to oppose the effect of the full-length protein.

The TTMP promoter is a TATA-less promoter and is dependent on Sp1 for basal activity. In the last part of the study, we characterized the 5' flanking region of the TTMP gene, which is responsible for its transcriptional regulation in cell culture. We have focused mainly on the identification of the promoter elements involved in constitutive gene expression. Using luciferase reporter gene assays from transiently transfected cells, we have mapped a highly active proximal promoter region. The 5' region of the TTMP gene lacks a TATA box or a CAAT box, and has a high GC content, as well as the presence of potential binding sites for several well-characterized transcription factors.

The sequence around the transcription start site identified on TTMP is consistent with the consensus sequence of the initiator element (PyPy A N T/A PyPy), where A is the start site (161). Furthermore, the promoter of TTMP contains a GC rich region around the transcription start site, with putative binding motifs for transcription factor Sp1. This is consistent with previous report that the transcription of TATA-less promoters frequently involves the action of a proximal Sp1 site (223). We have determined that basal activity of the proximal promoter region is largely influenced by the putative Sp1 binding sites found on the TTMP promoter, as well as demonstrated physical association of Sp1 with these putative binding motifs. Studies have identified Sp1 sites in the promoters of multiple growth-regulated genes. Direct evidence for the ability of Sp1 sites to modulate transcription during changes in cell growth came with the demonstration that they are involved in the effects of serum stimulation of quiescent cells at the rep3a promoter (188) as well as at the hamster dihydrofolate reductase (DHFR) (224,225) and the ornithine decarboxylase promoters (205). Interestingly, studies have indicated that depending on the promoter, upregulation of Sp1 site dependent transcription can be related to positive and negative changes in cell growth. For example, whereas Sp1 sites in the rep3a and DHFR promoters support the upregulation of transcription following growth stimulation of quiescent cells, Sp1 sites in the p21^{WAF1/CIP1} promoter are involved in transcriptional upregulation related to growth inhibition (226).

Conclusions and future work. In summary, we have identified a novel gene, TTMP, which is up-regulated in pancreatic cancer cells following exposure to the phorbol ester, TPA. Functional studies have shown that TTMP inhibits pancreatic cancer cell proliferation, and that it is a transmembrane protein that localizes to the endoplasmic reticulum. Promoter studies have also identified a TATA-less 5'-flanking region that is dependent on Sp1 for basal activity. Correlation of our data with tissue expression of

TTMP in human cancer specimens is important. However, as this is a novel gene, antibodies to the protein product is not available. The first task henceforth is to raise antibodies, to both the full-length as well as the N-term truncated protein, to study its expression in pancreatic and other cancers. Furthermore, animal experiments should be conducted to investigate the effects of down-regulation or over-expression of this gene in-vivo. Certainly, further studies are necessary to elucidate the molecular mechanisms and signal molecules that mediate TTMP-induced inhibition of cell proliferation. Preliminary work on the TTMP promoter will provide the basis for future studies wherein the objectives are to elucidate the mechanisms underlying the upregulation of TTMP by TPA, the tissue specific expression of the gene, as well as the possible downregulation of promoter activity during carcinogenesis of pancreatic as well as other cancers.

REFERENCES

1. Folsch, U. R. (1984) *Clin Gastroenterol* **13**(3), 679-699
2. Hellerstrom, C., and Swenne, I. (1991) *Diabetes* **40 Suppl 2**, 89-93
3. Herrington, M. K., and Adrian, T. E. (1998) *Biol Neonate* **73**(2), 129-136
4. Otonkoski, T., Beattie, G. M., Rubin, J. S., Lopez, A. D., Baird, A., and Hayek, A. (1994) *Diabetes* **43**(7), 947-953
5. Swenne, I., Heldin, C. H., Hill, D. J., and Hellerstrom, C. (1988) *Endocrinology* **122**(1), 214-218
6. Blackstock, A. W., Cox, A. D., and Tepper, J. E. (1996) *Oncology (Williston Park)* **10**(3), 301-307; discussion: 308-323
7. Hunstad, D. A., and Norton, J. A. (1995) *Surg Oncol* **4**(2), 61-74
8. Jemal, A., Murray, T., Ward, E., Samuels, A., Tiwari, R. C., Ghafoor, A., Feuer, E. J., and Thun, M. J. (2005) *CA Cancer J Clin* **55**(1), 10-30
9. Mangray, S., and King, T. C. (1998) *Front Biosci* **3**, D1148-1160
10. Rosemurgy, A. S., and Serafini, F. M. (2000) *Cancer Control* **7**(5), 437-451
11. A Seow, W. P. K., K S Chia, L M Shi, H P Lee, K Shanmugaratnam. (2004) Trends in Cancer Incidence 1998-2002. In. *Singapore Cancer Registry Report No. 6*
12. Anderson, K., Potter, JD, Mack, TM. (1996), Oxford University Press, New York
13. Lynch, H. T., Smyrk, T., Kern, S. E., Hruban, R. H., Lightdale, C. J., Lemon, S. J., Lynch, J. F., Fusaro, L. R., Fusaro, R. M., and Ghadirian, P. (1996) *Semin Oncol* **23**(2), 251-275
14. Jaffee, E. M., Hruban, R. H., Canto, M., and Kern, S. E. (2002) *Cancer Cell* **2**(1), 25-28
15. Eberle, M. A., Pfutzer, R., Pogue-Geile, K. L., Bronner, M. P., Crispin, D., Kimmey, M. B., Duerr, R. H., Kruglyak, L., Whitcomb, D. C., and Brentnall, T. A. (2002) *Am J Hum Genet* **70**(4), 1044-1048
16. Lowenfels, A. B., Maisonneuve, P., DiMagno, E. P., Elitsur, Y., Gates, L. K., Jr., Perrault, J., and Whitcomb, D. C. (1997) *J Natl Cancer Inst* **89**(6), 442-446
17. Whitcomb, D. C., Gorry, M. C., Preston, R. A., Furey, W., Sossenheimer, M. J., Ulrich, C. D., Martin, S. P., Gates, L. K., Jr., Amann, S. T., Toskes, P. P., Liddle, R., McGrath, K., Uomo, G., Post, J. C., and Ehrlich, G. D. (1996) *Nat Genet* **14**(2), 141-145
18. Sakorafas, G. H., and Tsiotou, A. G. (1999) *Eur J Surg Oncol* **25**(6), 562-565
19. Bale, A. E. (1994) *Endocrinol Metab Clin North Am* **23**(1), 109-115
20. Longnecker, D. S. (1999) *Ann N Y Acad Sci* **880**, 74-82
21. Bos, J. L. (1989) *Cancer Res* **49**(17), 4682-4689
22. Hahn, S. A., and Schmiegel, W. H. (1998) *Digestion* **59**(5), 493-501
23. Kinzler, K. W., and Vogelstein, B. (1996) *Cell* **87**(2), 159-170
24. Cubilla, A. L., and Fitzgerald, P. J. (1976) *Cancer Res* **36**(7 PT 2), 2690-2698
25. Klein, W. M., Hruban, R. H., Klein-Szanto, A. J., and Wilentz, R. E. (2002) *Mod Pathol* **15**(4), 441-447
26. Hruban, R. H., Adsay, N. V., Albores-Saavedra, J., Compton, C., Garrett, E. S., Goodman, S. N., Kern, S. E., Klimstra, D. S., Kloppel, G., Longnecker, D. S., Luttges, J., and Offerhaus, G. J. (2001) *Am J Surg Pathol* **25**(5), 579-586
27. Moskaluk, C. A., Hruban, R. H., and Kern, S. E. (1997) *Cancer Res* **57**(11), 2140-2143

28. Yamano, M., Fujii, H., Takagaki, T., Kadowaki, N., Watanabe, H., and Shirai, T. (2000) *Am J Pathol* **156**(6), 2123-2133
29. Luttges, J., Zamboni, G., Longnecker, D., and Kloppel, G. (2001) *Am J Surg Pathol* **25**(7), 942-948
30. Wilentz, R. E., Iacobuzio-Donahue, C. A., Argani, P., McCarthy, D. M., Parsons, J. L., Yeo, C. J., Kern, S. E., and Hruban, R. H. (2000) *Cancer Res* **60**(7), 2002-2006
31. Heinmoller, E., Dietmaier, W., Zirngibl, H., Heinmoller, P., Scaringe, W., Jauch, K. W., Hofstadter, F., and Ruschoff, J. (2000) *Am J Pathol* **157**(1), 83-92
32. Klimstra, D. S., Rosai, J., and Heffess, C. S. (1994) *Am J Surg Pathol* **18**(8), 765-778
33. Rozenblum, E., Schutte, M., Goggins, M., Hahn, S. A., Panzer, S., Zahurak, M., Goodman, S. N., Sohn, T. A., Hruban, R. H., Yeo, C. J., and Kern, S. E. (1997) *Cancer Res* **57**(9), 1731-1734
34. Biankin, A. V., Kench, J. G., Morey, A. L., Lee, C. S., Biankin, S. A., Head, D. R., Hugh, T. B., Henshall, S. M., and Sutherland, R. L. (2001) *Cancer Res* **61**(24), 8830-8837
35. Shields, J. M., Pruitt, K., McFall, A., Shaub, A., and Der, C. J. (2000) *Trends Cell Biol* **10**(4), 147-154
36. Korc, M., Chandrasekar, B., Yamanaka, Y., Friess, H., Buchler, M., and Beger, H. G. (1992) *J Clin Invest* **90**(4), 1352-1360
37. Barton, C. M., Hall, P. A., Hughes, C. M., Gullick, W. J., and Lemoine, N. R. (1991) *J Pathol* **163**(2), 111-116
38. Friess, H., Berberat, P., Schilling, M., Kunz, J., Korc, M., and Buchler, M. W. (1996) *J Mol Med* **74**(1), 35-42
39. Watanabe, M., Nobuta, A., Tanaka, J., and Asaka, M. (1996) *Int J Cancer* **67**(2), 264-268
40. Overholser, J. P., Prewett, M. C., Hooper, A. T., Waksal, H. W., and Hicklin, D. J. (2000) *Cancer* **89**(1), 74-82
41. Sibilia, M., Fleischmann, A., Behrens, A., Stingl, L., Carroll, J., Watt, F. M., Schlessinger, J., and Wagner, E. F. (2000) *Cell* **102**(2), 211-220
42. Day, J. D., Diguseppe, J. A., Yeo, C., Lai-Goldman, M., Anderson, S. M., Goodman, S. N., Kern, S. E., and Hruban, R. H. (1996) *Hum Pathol* **27**(2), 119-124
43. Wagner, M., Cao, T., Lopez, M. E., Hope, C., van Nostrand, K., Kobrin, M. S., Fan, H. U., Buchler, M. W., and Korc, M. (1996) *Int J Cancer* **68**(6), 782-787
44. Whelan, A. J., Bartsch, D., and Goodfellow, P. J. (1995) *N Engl J Med* **333**(15), 975-977
45. Goldstein, A. M., Struewing, J. P., Chidambaram, A., Fraser, M. C., and Tucker, M. A. (2000) *J Natl Cancer Inst* **92**(12), 1006-1010
46. Lynch, H. T., Brand, R. E., Hogg, D., Deters, C. A., Fusaro, R. M., Lynch, J. F., Liu, L., Knezetic, J., Lassam, N. J., Goggins, M., and Kern, S. (2002) *Cancer* **94**(1), 84-96
47. Goldstein, A. M., Fraser, M. C., Struewing, J. P., Hussussian, C. J., Ranade, K., Zametkin, D. P., Fontaine, L. S., Organic, S. M., Dracopoli, N. C., Clark, W. H., Jr., and et al. (1995) *N Engl J Med* **333**(15), 970-974
48. Borg, A., Sandberg, T., Nilsson, K., Johannsson, O., Klinker, M., Masback, A., Westerdahl, J., Olsson, H., and Ingvar, C. (2000) *J Natl Cancer Inst* **92**(15), 1260-1266
49. Sherr, C. J. (2001) *Nat Rev Mol Cell Biol* **2**(10), 731-737

50. Liu, L., Dilworth, D., Gao, L., Monzon, J., Summers, A., Lassam, N., and Hogg, D. (1999) *Nat Genet* **21**(1), 128-132
51. Lal, G., Liu, L., Hogg, D., Lassam, N. J., Redston, M. S., and Gallinger, S. (2000) *Genes Chromosomes Cancer* **27**(4), 358-361
52. Sharpless, N. E., and DePinho, R. A. (1999) *Curr Opin Genet Dev* **9**(1), 22-30
53. Gorunova, L., Hoglund, M., Andren-Sandberg, A., Dawiskiba, S., Jin, Y., Mitelman, F., and Johansson, B. (1998) *Genes Chromosomes Cancer* **23**(2), 81-99
54. Harada, T., Okita, K., Shiraishi, K., Kusano, N., Kondoh, S., and Sasaki, K. (2002) *Cancer Res* **62**(3), 835-839
55. Maser, R. S., and DePinho, R. A. (2002) *Science* **297**(5581), 565-569
56. Artandi, S. E., Chang, S., Lee, S. L., Alson, S., Gottlieb, G. J., Chin, L., and DePinho, R. A. (2000) *Nature* **406**(6796), 641-645
57. Chin, L., Artandi, S. E., Shen, Q., Tam, A., Lee, S. L., Gottlieb, G. J., Greider, C. W., and DePinho, R. A. (1999) *Cell* **97**(4), 527-538
58. Gisselsson, D., Jonson, T., Petersen, A., Strombeck, B., Dal Cin, P., Hoglund, M., Mitelman, F., Mertens, F., and Mandahl, N. (2001) *Proc Natl Acad Sci U S A* **98**(22), 12683-12688
59. Suehara, N., Mizumoto, K., Muta, T., Tominaga, Y., Shimura, H., Kitajima, S., Hamasaki, N., Tsuneyoshi, M., and Tanaka, M. (1997) *Clin Cancer Res* **3**(6), 993-998
60. Kobitsu, K., Tsutsumi, M., Tsujiuchi, T., Suzuki, F., Kido, A., Okajima, E., Fukuda, T., Sakaki, T., and Konishi, Y. (1997) *Mol Carcinog* **18**(3), 153-159
61. Goggins, M., Hruban, R. H., and Kern, S. E. (2000) *Am J Pathol* **156**(5), 1767-1771
62. Venkitaraman, A. R. (2002) *Cell* **108**(2), 171-182
63. Sato, N., Mizumoto, K., Nakamura, M., Maehara, N., Minamishima, Y. A., Nishio, S., Nagai, E., and Tanaka, M. (2001) *Cancer Genet Cytogenet* **126**(1), 13-19
64. Mahlamaki, E. H., Hoglund, M., Gorunova, L., Karhu, R., Dawiskiba, S., Andren-Sandberg, A., Kallioniemi, O. P., and Johansson, B. (1997) *Genes Chromosomes Cancer* **20**(4), 383-391
65. Peltomaki, P., and de la Chapelle, A. (1997) *Adv Cancer Res* **71**, 93-119
66. Aarnio, M., Mecklin, J. P., Aaltonen, L. A., Nystrom-Lahti, M., and Jarvinen, H. J. (1995) *Int J Cancer* **64**(6), 430-433
67. Lynch, H. T., Voorhees, G. J., Lanspa, S. J., McGreevy, P. S., and Lynch, J. F. (1985) *Br J Cancer* **52**(2), 271-273
68. Goggins, M., Offerhaus, G. J., Hilgers, W., Griffin, C. A., Shekher, M., Tang, D., Sohn, T. A., Yeo, C. J., Kern, S. E., and Hruban, R. H. (1998) *Am J Pathol* **152**(6), 1501-1507
69. Yamamoto, H., Itoh, F., Nakamura, H., Fukushima, H., Sasaki, S., Perucho, M., and Imai, K. (2001) *Cancer Res* **61**(7), 3139-3144
70. Wilentz, R. E., Goggins, M., Redston, M., Marcus, V. A., Adsay, N. V., Sohn, T. A., Kadkol, S. S., Yeo, C. J., Choti, M., Zahurak, M., Johnson, K., Tascilar, M., Offerhaus, G. J., Hruban, R. H., and Kern, S. E. (2000) *Am J Pathol* **156**(5), 1641-1651
71. Hahn, S. A., Schutte, M., Hoque, A. T., Moskaluk, C. A., da Costa, L. T., Rozenblum, E., Weinstein, C. L., Fischer, A., Yeo, C. J., Hruban, R. H., and Kern, S. E. (1996) *Science* **271**(5247), 350-353
72. Massague, J., Blain, S. W., and Lo, R. S. (2000) *Cell* **103**(2), 295-309
73. Sirard, C., Kim, S., Mirtsos, C., Tadich, P., Hoodless, P. A., Itie, A., Maxson, R., Wrana, J. L., and Mak, T. W. (2000) *J Biol Chem* **275**(3), 2063-2070

74. Jonson, T., Albrechtsson, E., Axelson, J., Heidenblad, M., Gorunova, L., Johansson, B., and Hoglund, M. (2001) *Int J Oncol* **19**(1), 71-81
75. Dai, J. L., Schutte, M., Bansal, R. K., Wilentz, R. E., Sugar, A. Y., and Kern, S. E. (1999) *Mol Carcinog* **26**(1), 37-43
76. Giehl, K., Seidel, B., Gierschik, P., Adler, G., and Menke, A. (2000) *Oncogene* **19**(39), 4531-4541
77. Rowland-Goldsmith, M. A., Maruyama, H., Kusama, T., Ralli, S., and Korc, M. (2001) *Clin Cancer Res* **7**(9), 2931-2940
78. Hemminki, A., Markie, D., Tomlinson, I., Avizienyte, E., Roth, S., Loukola, A., Bignell, G., Warren, W., Aminoff, M., Hoglund, P., Jarvinen, H., Kristo, P., Pelin, K., Ridanpaa, M., Salovaara, R., Toro, T., Bodmer, W., Olschwang, S., Olsen, A. S., Stratton, M. R., de la Chapelle, A., and Aaltonen, L. A. (1998) *Nature* **391**(6663), 184-187
79. Giardiello, F. M., Brensinger, J. D., Tersmette, A. C., Goodman, S. N., Petersen, G. M., Booker, S. V., Cruz-Correa, M., and Offerhaus, J. A. (2000) *Gastroenterology* **119**(6), 1447-1453
80. Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A., Gocayne, J. D., Amanatides, P., Ballew, R. M., Huson, D. H., Wortman, J. R., Zhang, Q., Kodira, C. D., Zheng, X. H., Chen, L., Skupski, M., Subramanian, G., Thomas, P. D., Zhang, J., Gabor Miklos, G. L., Nelson, C., Broder, S., Clark, A. G., Nadeau, J., McKusick, V. A., Zinder, N., Levine, A. J., Roberts, R. J., Simon, M., Slayman, C., Hunkapiller, M., Bolanos, R., Delcher, A., Dew, I., Fasulo, D., Flanigan, M., Florea, L., Halpern, A., Hannenhalli, S., Kravitz, S., Levy, S., Mobarry, C., Reinert, K., Remington, K., Abu-Threideh, J., Beasley, E., Biddick, K., Bonazzi, V., Brandon, R., Cargill, M., Chandramouliswaran, I., Charlab, R., Chaturvedi, K., Deng, Z., Di Francesco, V., Dunn, P., Eilbeck, K., Evangelista, C., Gabrielian, A. E., Gan, W., Ge, W., Gong, F., Gu, Z., Guan, P., Heiman, T. J., Higgins, M. E., Ji, R. R., Ke, Z., Ketchum, K. A., Lai, Z., Lei, Y., Li, Z., Li, J., Liang, Y., Lin, X., Lu, F., Merkulov, G. V., Milshina, N., Moore, H. M., Naik, A. K., Narayan, V. A., Neelam, B., Nusskern, D., Rusch, D. B., Salzberg, S., Shao, W., Shue, B., Sun, J., Wang, Z., Wang, A., Wang, X., Wang, J., Wei, M., Wides, R., Xiao, C., Yan, C., Yao, A., Ye, J., Zhan, M., Zhang, W., Zhang, H., Zhao, Q., Zheng, L., Zhong, F., Zhong, W., Zhu, S., Zhao, S., Gilbert, D., Baumhueter, S., Spier, G., Carter, C., Cravchik, A., Woodage, T., Ali, F., An, H., Awe, A., Baldwin, D., Baden, H., Barnstead, M., Barrow, I., Beeson, K., Busam, D., Carver, A., Center, A., Cheng, M. L., Curry, L., Danaher, S., Davenport, L., Desilets, R., Dietz, S., Dodson, K., Doup, L., Ferriera, S., Garg, N., Gluecksmann, A., Hart, B., Haynes, J., Haynes, C., Heiner, C., Hladun, S., Hostin, D., Houck, J., Howland, T., Ibegwam, C., Johnson, J., Kalush, F., Kline, L., Koduru, S., Love, A., Mann, F., May, D., McCawley, S., McIntosh, T., McMullen, I., Moy, M., Moy, L., Murphy, B., Nelson, K., Pfannkoch, C., Pratts, E., Puri, V., Qureshi, H., Reardon, M., Rodriguez, R., Rogers, Y. H., Romblad, D., Ruhfel, B., Scott, R., Sitter, C., Smallwood, M., Stewart, E., Strong, R., Suh, E., Thomas, R., Tint, N. N., Tse, S., Vech, C., Wang, G., Wetter, J., Williams, S., Williams, M., Windsor, S., Winn-Deen, E., Wolfe, K., Zaveri, J., Zaveri, K., Abril, J. F., Guigo, R., Campbell, M. J., Sjolander, K. V., Karlak, B., Kejariwal, A., Mi, H., Lazareva, B., Hatton, T., Narechania, A., Diemer, K., Muruganujan, A., Guo, N., Sato, S., Bafna, V., Istrail, S., Lippert, R., Schwartz, R., Walenz, B., Yooseph, S., Allen, D., Basu, A., Baxendale, J., Blick, L., Caminha, M., Carnes-Stine, J., Caulk, P., Chiang, Y. H., Coyne, M., Dahlke, C., Mays, A., Dombroski, M., Donnelly, M., Ely, D., Esparham, S., Fosler, C., Gire,

- H., Glanowski, S., Glasser, K., Glodek, A., Gorokhov, M., Graham, K., Gropman, B., Harris, M., Heil, J., Henderson, S., Hoover, J., Jennings, D., Jordan, C., Jordan, J., Kasha, J., Kagan, L., Kraft, C., Levitsky, A., Lewis, M., Liu, X., Lopez, J., Ma, D., Majoros, W., McDaniel, J., Murphy, S., Newman, M., Nguyen, T., Nguyen, N., Nodell, M., Pan, S., Peck, J., Peterson, M., Rowe, W., Sanders, R., Scott, J., Simpson, M., Smith, T., Sprague, A., Stockwell, T., Turner, R., Venter, E., Wang, M., Wen, M., Wu, D., Wu, M., Xia, A., Zandieh, A., and Zhu, X. (2001) *Science* **291**(5507), 1304-1351
81. Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., McKernan, K., Meldrim, J., Mesirov, J. P., Miranda, C., Morris, W., Naylor, J., Raymond, C., Rosetti, M., Santos, R., Sheridan, A., Sougnez, C., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Wyman, D., Rogers, J., Sulston, J., Ainscough, R., Beck, S., Bentley, D., Burton, J., Clee, C., Carter, N., Coulson, A., Deadman, R., Deloukas, P., Dunham, A., Dunham, I., Durbin, R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Hunt, A., Jones, M., Lloyd, C., McMurray, A., Matthews, L., Mercer, S., Milne, S., Mullikin, J. C., Mungall, A., Plumb, R., Ross, M., Shownkeen, R., Sims, S., Waterston, R. H., Wilson, R. K., Hillier, L. W., McPherson, J. D., Marra, M. A., Mardis, E. R., Fulton, L. A., Chinwalla, A. T., Pepin, K. H., Gish, W. R., Chissoe, S. L., Wendl, M. C., Delehaunty, K. D., Miner, T. L., Delehaunty, A., Kramer, J. B., Cook, L. L., Fulton, R. S., Johnson, D. L., Minx, P. J., Clifton, S. W., Hawkins, T., Branscomb, E., Predki, P., Richardson, P., Wenning, S., Slezak, T., Doggett, N., Cheng, J. F., Olsen, A., Lucas, S., Elkin, C., Uberbacher, E., Frazier, M., Gibbs, R. A., Muzny, D. M., Scherer, S. E., Bouck, J. B., Sodergren, E. J., Worley, K. C., Rives, C. M., Gorrell, J. H., Metzker, M. L., Naylor, S. L., Kucherlapati, R. S., Nelson, D. L., Weinstock, G. M., Sakaki, Y., Fujiyama, A., Hattori, M., Yada, T., Toyoda, A., Itoh, T., Kawagoe, C., Watanabe, H., Totoki, Y., Taylor, T., Weissenbach, J., Heilig, R., Saurin, W., Artiguenave, F., Brottier, P., Bruls, T., Pelletier, E., Robert, C., Wincker, P., Smith, D. R., Doucette-Stamm, L., Rubenfield, M., Weinstock, K., Lee, H. M., Dubois, J., Rosenthal, A., Platzer, M., Nyakatura, G., Taudien, S., Rump, A., Yang, H., Yu, J., Wang, J., Huang, G., Gu, J., Hood, L., Rowen, L., Madan, A., Qin, S., Davis, R. W., Federspiel, N. A., Abola, A. P., Proctor, M. J., Myers, R. M., Schmutz, J., Dickson, M., Grimwood, J., Cox, D. R., Olson, M. V., Kaul, R., Raymond, C., Shimizu, N., Kawasaki, K., Minoshima, S., Evans, G. A., Athanasiou, M., Schultz, R., Roe, B. A., Chen, F., Pan, H., Ramser, J., Lehrach, H., Reinhardt, R., McCombie, W. R., de la Bastide, M., Dedhia, N., Blocker, H., Hornischer, K., Nordsiek, G., Agarwala, R., Aravind, L., Bailey, J. A., Bateman, A., Batzoglou, S., Birney, E., Bork, P., Brown, D. G., Burge, C. B., Cerutti, L., Chen, H. C., Church, D., Clamp, M., Copley, R. R., Doerks, T., Eddy, S. R., Eichler, E. E., Furey, T. S., Galagan, J., Gilbert, J. G., Harmon, C., Hayashizaki, Y., Haussler, D., Hermjakob, H., Hokamp, K., Jang, W., Johnson, L. S., Jones, T. A., Kasif, S., Kasprzyk, A., Kennedy, S., Kent, W. J., Kitts, P., Koonin, E. V., Korf, I., Kulp, D., Lancet, D., Lowe, T. M., McLysaght, A., Mikkelsen, T., Moran, J. V., Mulder, N., Pollara, V. J., Ponting, C. P., Schuler, G., Schultz, J., Slater, G., Smit, A. F., Stupka, E., Szustakowski, J., Thierry-Mieg, D., Thierry-Mieg, J., Wagner, L., Wallis, J., Wheeler, R., Williams, A., Wolf, Y. I., Wolfe, K. H., Yang, S. P., Yeh, R. F., Collins, F., Guyer, M. S., Peterson, J., Felsenfeld, A., Wetterstrand, K. A., Patrinos, A., Morgan, M. J., de Jong, P., Catanese, J. J., Osoegawa, K., Shizuya, H., Choi, S., and Chen, Y. J. (2001) *Nature* **409**(6822), 860-921

82. (2004) *Nature* **431**(7011), 931-945
83. Linzer, D. I., and Levine, A. J. (1979) *Cell* **17**(1), 43-52
84. O'Farrell, P. H. (1975) *J Biol Chem* **250**(10), 4007-4021
85. Takahashi, N., Kaji, H., Yanagida, M., Hayano, T., and Isobe, T. (2003) *J Nutr* **133**(6 Suppl 1), 2090S-2096S
86. Fung, E. T., Thulasiraman, V., Weinberger, S. R., and Dalmasso, E. A. (2001) *Curr Opin Biotechnol* **12**(1), 65-69
87. Sargent, T. D. (1987) *Methods Enzymol* **152**, 423-432
88. Masiakowski, P., Breathnach, R., Bloch, J., Gannon, F., Krust, A., and Chambon, P. (1982) *Nucleic Acids Res* **10**(24), 7895-7903
89. Manzari, V., Gallo, R. C., Franchini, G., Westin, E., Ceccherini-Nelli, L., Popovic, M., and Wong-Staal, F. (1983) *Proc Natl Acad Sci U S A* **80**(1), 11-15
90. Zimmermann, C. R., Orr, W. C., Leclerc, R. F., Barnard, E. C., and Timberlake, W. E. (1980) *Cell* **21**(3), 709-715
91. Hedrick, S. M., Cohen, D. I., Nielsen, E. A., and Davis, M. M. (1984) *Nature* **308**(5955), 149-153
92. el-Deiry, W. S., Tokino, T., Velculescu, V. E., Levy, D. B., Parsons, R., Trent, J. M., Lin, D., Mercer, W. E., Kinzler, K. W., and Vogelstein, B. (1993) *Cell* **75**(4), 817-825
93. Liang, P., and Pardee, A. B. (1992) *Science* **257**(5072), 967-971
94. Liang, P., Bauer, D., Averboukh, L., Warthoe, P., Rohrwild, M., Muller, H., Strauss, M., and Pardee, A. B. (1995) *Methods Enzymol* **254**, 304-321
95. McCarthy, S. A., Samuels, M. L., Pritchard, C. A., Abraham, J. A., and McMahon, M. (1995) *Genes Dev* **9**(16), 1953-1964
96. Zhang, R., Tan, Z., and Liang, P. (2000) *J Biol Chem* **275**(32), 24436-24443
97. You, M., Ku, P. T., Hrdlickova, R., and Bose, H. R., Jr. (1997) *Mol Cell Biol* **17**(12), 7328-7341
98. Park, B. W., O'Rourke, D. M., Wang, Q., Davis, J. G., Post, A., Qian, X., and Greene, M. I. (1999) *Proc Natl Acad Sci U S A* **96**(11), 6434-6438
99. Schena, M., Shalon, D., Davis, R. W., and Brown, P. O. (1995) *Science* **270**(5235), 467-470
100. Chee, M., Yang, R., Hubbell, E., Berno, A., Huang, X. C., Stern, D., Winkler, J., Lockhart, D. J., Morris, M. S., and Fodor, S. P. (1996) *Science* **274**(5287), 610-614
101. Ramaswamy, S., Tamayo, P., Rifkin, R., Mukherjee, S., Yeang, C. H., Angelo, M., Ladd, C., Reich, M., Latulippe, E., Mesirov, J. P., Poggio, T., Gerald, W., Loda, M., Lander, E. S., and Golub, T. R. (2001) *Proc Natl Acad Sci U S A* **98**(26), 15149-15154
102. Yeoh, E. J., Ross, M. E., Shurtleff, S. A., Williams, W. K., Patel, D., Mahfouz, R., Behm, F. G., Raimondi, S. C., Relling, M. V., Patel, A., Cheng, C., Campana, D., Wilkins, D., Zhou, X., Li, J., Liu, H., Pui, C. H., Evans, W. E., Naeve, C., Wong, L., and Downing, J. R. (2002) *Cancer Cell* **1**(2), 133-143
103. Alizadeh, A. A., Eisen, M. B., Davis, R. E., Ma, C., Lossos, I. S., Rosenwald, A., Boldrick, J. C., Sabet, H., Tran, T., Yu, X., Powell, J. I., Yang, L., Marti, G. E., Moore, T., Hudson, J., Jr., Lu, L., Lewis, D. B., Tibshirani, R., Sherlock, G., Chan, W. C., Greiner, T. C., Weisenburger, D. D., Armitage, J. O., Warnke, R., Levy, R., Wilson, W., Grever, M. R., Byrd, J. C., Botstein, D., Brown, P. O., and Staudt, L. M. (2000) *Nature* **403**(6769), 503-511
104. Garber, M. E., Troyanskaya, O. G., Schluens, K., Petersen, S., Thaessler, Z., Pacyna-Gengelbach, M., van de Rijn, M., Rosen, G. D., Perou, C. M., Whyte, R.

- I., Altman, R. B., Brown, P. O., Botstein, D., and Petersen, I. (2001) *Proc Natl Acad Sci U S A* **98**(24), 13784-13789
105. Sorlie, T., Perou, C. M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., Hastie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Thorsen, T., Quist, H., Matese, J. C., Brown, P. O., Botstein, D., Eystein Lonning, P., and Borresen-Dale, A. L. (2001) *Proc Natl Acad Sci U S A* **98**(19), 10869-10874
106. Dhanasekaran, S. M., Barrette, T. R., Ghosh, D., Shah, R., Varambally, S., Kurachi, K., Pienta, K. J., Rubin, M. A., and Chinnaiyan, A. M. (2001) *Nature* **412**(6849), 822-826
107. Wooster, R. (2000) *Trends Genet* **16**(8), 327-329
108. Petricoin, E. F., 3rd, Hackett, J. L., Lesko, L. J., Puri, R. K., Gutman, S. I., Chumakov, K., Woodcock, J., Feigal, D. W., Jr., Zoon, K. C., and Sistare, F. D. (2002) *Nat Genet* **32 Suppl**, 474-479
109. Kothapalli, R., Yoder, S. J., Mane, S., and Loughran, T. P., Jr. (2002) *BMC Bioinformatics* **3**, 22
110. Kuo, W. P., Jenssen, T. K., Butte, A. J., Ohno-Machado, L., and Kohane, I. S. (2002) *Bioinformatics* **18**(3), 405-412
111. Goryachev, A. B., Macgregor, P. F., and Edwards, A. M. (2001) *J Comput Biol* **8**(4), 443-461
112. King, H. C., and Sinha, A. A. (2001) *Jama* **286**(18), 2280-2288
113. Adams, M. D., Dubnick, M., Kerlavage, A. R., Moreno, R., Kelley, J. M., Utterback, T. R., Nagle, J. W., Fields, C., and Venter, J. C. (1992) *Nature* **355**(6361), 632-634
114. Velculescu, V. E., Zhang, L., Vogelstein, B., and Kinzler, K. W. (1995) *Science* **270**(5235), 484-487
115. Brenner, S., Johnson, M., Bridgham, J., Golda, G., Lloyd, D. H., Johnson, D., Luo, S., McCurdy, S., Foy, M., Ewan, M., Roth, R., George, D., Eletr, S., Albrecht, G., Vermaas, E., Williams, S. R., Moon, K., Burcham, T., Pallas, M., DuBridge, R. B., Kirchner, J., Fearon, K., Mao, J., and Corcoran, K. (2000) *Nat Biotechnol* **18**(6), 630-634
116. Takai, Y., Kishimoto, A., Inoue, M., and Nishizuka, Y. (1977) *J Biol Chem* **252**(21), 7603-7609
117. Blumberg, P. M. (1991) *Mol Carcinog* **4**(5), 339-344
118. Mischak, H., Goodnight, J. A., Kolch, W., Martiny-Baron, G., Schaechtle, C., Kazanietz, M. G., Blumberg, P. M., Pierce, J. H., and Mushinski, J. F. (1993) *J Biol Chem* **268**(9), 6090-6096
119. Cacace, A. M., Guadagno, S. N., Krauss, R. S., Fabbro, D., and Weinstein, I. B. (1993) *Oncogene* **8**(8), 2095-2104
120. Weinstein, I. B. (1991) *Environ Health Perspect* **93**, 175-179
121. Ojeda, F., Guarda, M. I., Maldonado, C., Folch, H., and Diehl, H. (1992) *Int J Radiat Biol* **61**(5), 663-667
122. de Vente, J. E., Kukoly, C. A., Bryant, W. O., Posekany, K. J., Chen, J., Fletcher, D. J., Parker, P. J., Pettit, G. J., Lozano, G., Cook, P. P., and et al. (1995) *J Clin Invest* **96**(4), 1874-1886
123. Garzotto, M., White-Jones, M., Jiang, Y., Ehleiter, D., Liao, W. C., Haimovitz-Friedman, A., Fuks, Z., and Kolesnick, R. (1998) *Cancer Res* **58**(10), 2260-2264
124. Emoto, Y., Manome, Y., Meinhardt, G., Kisaki, H., Kharbanda, S., Robertson, M., Ghayur, T., Wong, W. W., Kamen, R., Weichselbaum, R., and et al. (1995) *Embo J* **14**(24), 6148-6156
125. Denning, M. F., Wang, Y., Nickoloff, B. J., and Wrono-Smith, T. (1998) *J Biol Chem* **273**(45), 29995-30002

126. Li, L., Lorenzo, P. S., Bogi, K., Blumberg, P. M., and Yuspa, S. H. (1999) *Mol Cell Biol* **19**(12), 8547-8558
127. Livneh, E., and Fishman, D. D. (1997) *Eur J Biochem* **248**(1), 1-9
128. Soh, J. W., Lee, E. H., Prywes, R., and Weinstein, I. B. (1999) *Mol Cell Biol* **19**(2), 1313-1324
129. Zhou, W., Takuwa, N., Kumada, M., and Takuwa, Y. (1993) *J Biol Chem* **268**(31), 23041-23048
130. Livneh, E., Shimon, T., Bechor, E., Doki, Y., Schieren, I., and Weinstein, I. B. (1996) *Oncogene* **12**(7), 1545-1555
131. Sasaguri, T., Ishida, A., Kosaka, C., Nojima, H., and Ogata, J. (1996) *J Biol Chem* **271**(14), 8345-8351
132. Frey, M. R., Saxon, M. L., Zhao, X., Rollins, A., Evans, S. S., and Black, J. D. (1997) *J Biol Chem* **272**(14), 9424-9435
133. Barth, H., and Kinzel, V. (1994) *Exp Cell Res* **212**(2), 383-388
134. Arita, Y., Buffolino, P., and Coppock, D. L. (1998) *Exp Cell Res* **242**(2), 381-390
135. Hass, R., Gunji, H., Hirano, M., Weichselbaum, R., and Kufe, D. (1993) *Cell Growth Differ* **4**(3), 159-166
136. Kelekar, A., and Thompson, C. B. (1998) *Trends Cell Biol* **8**(8), 324-330
137. Ishino, K., Fukazawa, H., Shikano, M., Ohba, M., Kuroki, T., and Uehara, Y. (2002) *Mol Carcinog* **34**(4), 180-186
138. Hirata, M., Tsuchida, A., Iwao, T., Sasaki, T., Matsubara, K., Yamamoto, S., Morinaka, K., Kawasaki, Y., Fujimoto, Y., Inoue, H., Kariya, K., and Kajiyama, G. (1999) *Int J Oncol* **14**(6), 1129-1135
139. Franz, M. G., Norman, J. G., Fabri, P. J., and Gower, W. R., Jr. (1996) *Ann Surg Oncol* **3**(6), 564-569
140. Denham, D. W., Franz, M. G., Denham, W., Zervos, E. E., Gower, W. R., Jr., Rosemurgy, A. S., and Norman, J. (1998) *Surgery* **124**(2), 218-223; discussion 223-214
141. Detjen, K. M., Brembeck, F. H., Welzel, M., Kaiser, A., Haller, H., Wiedenmann, B., and Rosewicz, S. (2000) *J Cell Sci* **113** (Pt 17), 3025-3035
142. Wintroub AB, T. W., Cluck MW, Wang QJ, Mushinski JF, Adrian TE. (1998) *Pancreas* **17**, 459
143. Ingle PI, W. A., Ding XZ, Adrian TE. (2001) *Pancreas* **23**, 443
144. Salabat MR, D. X., Talamonti MS, Bell RH, Adrian TE. (2004) *Pancreas* **29**, 334
145. Trauzold, A., Wermann, H., Arlt, A., Schutze, S., Schafer, H., Oestern, S., Roder, C., Ungefroren, H., Lampe, E., Heinrich, M., Walczak, H., and Kalthoff, H. (2001) *Oncogene* **20**(31), 4258-4269
146. Trauzold, A., Schmiedel, S., Sipos, B., Wermann, H., Westphal, S., Roder, C., Klapper, W., Arlt, A., Lehnert, L., Ungefroren, H., Johannes, F. J., and Kalthoff, H. (2003) *Oncogene* **22**(55), 8939-8947
147. Farrow, B., Thomas, R. P., Wang, X. F., and Evers, B. M. (2002) *Int J Gastrointest Cancer* **32**(2-3), 63-72
148. Laudanna, C., Sorio, C., Tecchio, C., Butcher, E. C., Bonora, A., Bassi, C., and Scarpa, A. (2003) *Lab Invest* **83**(8), 1155-1163
149. Neid, M., Datta, K., Stephan, S., Khanna, I., Pal, S., Shaw, L., White, M., and Mukhopadhyay, D. (2004) *J Biol Chem* **279**(6), 3941-3948
150. Smale, S. T. (2001) *Genes Dev* **15**(19), 2503-2508
151. Blackwood, E. M., and Kadonaga, J. T. (1998) *Science* **281**(5373), 60-63
152. Bulger, M., and Groudine, M. (1999) *Genes Dev* **13**(19), 2465-2477
153. West, A. G., Gaszner, M., and Felsenfeld, G. (2002) *Genes Dev* **16**(3), 271-288
154. Breathnach, R., and Chambon, P. (1981) *Annu Rev Biochem* **50**, 349-383

155. Singer, V. L., Wobbe, C. R., and Struhl, K. (1990) *Genes Dev* **4**(4), 636-645
156. Suzuki, Y., Tsunoda, T., Sese, J., Taira, H., Mizushima-Sugano, J., Hata, H., Ota, T., Isogai, T., Tanaka, T., Nakamura, Y., Suyama, A., Sakaki, Y., Morishita, S., Okubo, K., and Sugano, S. (2001) *Genome Res* **11**(5), 677-684
157. Smale, S. T., Schmidt, M. C., Berk, A. J., and Baltimore, D. (1990) *Proc Natl Acad Sci U S A* **87**(12), 4509-4513
158. Burley, S. K., and Roeder, R. G. (1996) *Annu Rev Biochem* **65**, 769-799
159. Hultmark, D., Klemenz, R., and Gehring, W. J. (1986) *Cell* **44**(3), 429-438
160. Struhl, K. (1987) *Cell* **49**(3), 295-297
161. Javahery, R., Khachi, A., Lo, K., Zenzie-Gregory, B., and Smale, S. T. (1994) *Mol Cell Biol* **14**(1), 116-127
162. Kaufmann, J., and Smale, S. T. (1994) *Genes Dev* **8**(7), 821-829
163. Burke, T. W., and Kadonaga, J. T. (1996) *Genes Dev* **10**(6), 711-724
164. Kaufmann, J., Ahrens, K., Koop, R., Smale, S. T., and Muller, R. (1998) *Mol Cell Biol* **18**(1), 233-239
165. Chalkley, G. E., and Verrijzer, C. P. (1999) *Embo J* **18**(17), 4835-4845
166. Carcamo, J., Buckbinder, L., and Reinberg, D. (1991) *Proc Natl Acad Sci U S A* **88**(18), 8052-8056
167. Weis, L., and Reinberg, D. (1997) *Mol Cell Biol* **17**(6), 2973-2984
168. Burke, T. W., and Kadonaga, J. T. (1997) *Genes Dev* **11**(22), 3020-3031
169. Kutach, A. K., and Kadonaga, J. T. (2000) *Mol Cell Biol* **20**(13), 4754-4764
170. Zhou, T., and Chiang, C. M. (2001) *J Biol Chem* **276**(27), 25503-25511
171. Lagrange, T., Kapanidis, A. N., Tang, H., Reinberg, D., and Ebright, R. H. (1998) *Genes Dev* **12**(1), 34-44
172. Evans, R., Fairley, J. A., and Roberts, S. G. (2001) *Genes Dev* **15**(22), 2945-2949
173. Brandeis, M., Frank, D., Keshet, I., Siegfried, Z., Mendelsohn, M., Nemes, A., Temper, V., Razin, A., and Cedar, H. (1994) *Nature* **371**(6496), 435-438
174. Macleod, D., Charlton, J., Mullins, J., and Bird, A. P. (1994) *Genes Dev* **8**(19), 2282-2292
175. Emami, K. H., Navarre, W. W., and Smale, S. T. (1995) *Mol Cell Biol* **15**(11), 5906-5916
176. Kadonaga, J. T., Carner, K. R., Masiarz, F. R., and Tjian, R. (1987) *Cell* **51**(6), 1079-1090
177. Pugh, B. F., and Tjian, R. (1991) *Genes Dev* **5**(11), 1935-1945
178. Zenzie-Gregory, B., Khachi, A., Garraway, I. P., and Smale, S. T. (1993) *Mol Cell Biol* **13**(7), 3841-3849
179. Blake, M. C., Jambou, R. C., Swick, A. G., Kahn, J. W., and Azizkhan, J. C. (1990) *Mol Cell Biol* **10**(12), 6632-6641
180. Jackson, S. P., and Tjian, R. (1988) *Cell* **55**(1), 125-133
181. Jackson, S. P., MacDonald, J. J., Lees-Miller, S., and Tjian, R. (1990) *Cell* **63**(1), 155-165
182. Hagen, G., Muller, S., Beato, M., and Suske, G. (1992) *Nucleic Acids Res* **20**(21), 5519-5525
183. Kingsley, C., and Winoto, A. (1992) *Mol Cell Biol* **12**(10), 4251-4261
184. Philipsen, S., and Suske, G. (1999) *Nucleic Acids Res* **27**(15), 2991-3000
185. Turner, J., and Crossley, M. (1999) *Trends Biochem Sci* **24**(6), 236-240
186. Sif, S., Capobianco, A. J., and Gilmore, T. D. (1993) *Oncogene* **8**(9), 2501-2509
187. Eastman, H. B., Swick, A. G., Schmitt, M. C., and Azizkhan, J. C. (1991) *Proc Natl Acad Sci U S A* **88**(19), 8572-8576
188. Schilling, L. J., and Farnham, P. J. (1995) *Cell Growth Differ* **6**(5), 541-548

189. Black, A. R., Jensen, D., Lin, S. Y., and Azizkhan, J. C. (1999) *J Biol Chem* **274**(3), 1207-1215
190. Fandos, C., Sanchez-Feutrie, M., Santalucia, T., Vinals, F., Cadefau, J., Guma, A., Cusso, R., Kaliman, P., Canicio, J., Palacin, M., and Zorzano, A. (1999) *J Mol Biol* **294**(1), 103-119
191. Kennett, S. B., Udvadia, A. J., and Horowitz, J. M. (1997) *Nucleic Acids Res* **25**(15), 3110-3117
192. Tanaka, T., Kanai, H., Sekiguchi, K., Aihara, Y., Yokoyama, T., Arai, M., Kanda, T., Nagai, R., and Kurabayashi, M. (2000) *J Mol Cell Cardiol* **32**(11), 1955-1967
193. Thiesen, H. J., and Bach, C. (1990) *Nucleic Acids Res* **18**(11), 3203-3209
194. Hagen, G., Dennig, J., Preiss, A., Beato, M., and Suske, G. (1995) *J Biol Chem* **270**(42), 24989-24994
195. Shields, J. M., and Yang, V. W. (1998) *Nucleic Acids Res* **26**(3), 796-802
196. Dang, D. T., Pevsner, J., and Yang, V. W. (2000) *Int J Biochem Cell Biol* **32**(11-12), 1103-1121
197. Imataka, H., Sogawa, K., Yasumoto, K., Kikuchi, Y., Sasano, K., Kobayashi, A., Hayami, M., and Fujii-Kuriyama, Y. (1992) *Embo J* **11**(10), 3663-3671
198. Dennig, J., Beato, M., and Suske, G. (1996) *Embo J* **15**(20), 5659-5667
199. Yet, S. F., McA'Nulty, M. M., Folta, S. C., Yen, H. W., Yoshizumi, M., Hsieh, C. M., Layne, M. D., Chin, M. T., Wang, H., Perrella, M. A., Jain, M. K., and Lee, M. E. (1998) *J Biol Chem* **273**(2), 1026-1031
200. Black, A. R., Black, J. D., and Azizkhan-Clifford, J. (2001) *J Cell Physiol* **188**(2), 143-160
201. Iacobuzio-Donahue, C. A., Maitra, A., Shen-Ong, G. L., van Heek, T., Ashfaq, R., Meyer, R., Walter, K., Berg, K., Hollingsworth, M. A., Cameron, J. L., Yeo, C. J., Kern, S. E., Goggins, M., and Hruban, R. H. (2002) *Am J Pathol* **160**(4), 1239-1249
202. Rosty, C., Christa, L., Kuzdzal, S., Baldwin, W. M., Zahurak, M. L., Carnot, F., Chan, D. W., Canto, M., Lillemoe, K. D., Cameron, J. L., Yeo, C. J., Hruban, R. H., and Goggins, M. (2002) *Cancer Res* **62**(6), 1868-1875
203. Han, H., Bearss, D. J., Browne, L. W., Calaluce, R., Nagle, R. B., and Von Hoff, D. D. (2002) *Cancer Res* **62**(10), 2890-2896
204. Chan, C. Y., Salabat, M. R., Ding, X. Z., Kelly, D. L., Talamonti, M. S., Bell, R. H., Jr., and Adrian, T. E. (2005) *Biochem Biophys Res Commun* **329**(2), 755-764
205. Kumar, A. P., and Butler, A. P. (1998) *Biochem Biophys Res Commun* **252**(2), 517-523
206. Edge, A. S. (2003) *Biochem J* **376**(Pt 2), 339-350
207. Nagayama, Y., Nishihara, E., Namba, H., Yamashita, S., and Niwa, M. (2000) *J Pharmacol Exp Ther* **295**(1), 404-409
208. Sairam, M. R. (1989) *Faseb J* **3**(8), 1915-1926
209. Laplante, J. M., O'Rourke, F., Lu, X., Fein, A., Olsen, A., and Feinstein, M. B. (2000) *Biochem J* **348 Pt 1**, 189-199
210. O'Rourke, F. A., LaPlante, J. M., and Feinstein, M. B. (2003) *Biochem J* **373**(Pt 1), 133-143
211. Breckenridge, D. G., Germain, M., Mathai, J. P., Nguyen, M., and Shore, G. C. (2003) *Oncogene* **22**(53), 8608-8618
212. Rao, R. V., Ellerby, H. M., and Bredesen, D. E. (2004) *Cell Death Differ* **11**(4), 372-380
213. Garber, K. (2006) *J Natl Cancer Inst* **98**(8), 512-514
214. Brewer, J. W., and Diehl, J. A. (2000) *Proc Natl Acad Sci U S A* **97**(23), 12625-12630

215. Brewer, J. W., Hendershot, L. M., Sherr, C. J., and Diehl, J. A. (1999) *Proc Natl Acad Sci U S A* **96**(15), 8505-8510
216. Niwa, M., and Walter, P. (2000) *Proc Natl Acad Sci U S A* **97**(23), 12396-12397
217. Lee, K., Tirasophon, W., Shen, X., Michalak, M., Prywes, R., Okada, T., Yoshida, H., Mori, K., and Kaufman, R. J. (2002) *Genes Dev* **16**(4), 452-466
218. Clauss, I. M., Gravallesse, E. M., Darling, J. M., Shapiro, F., Glimcher, M. J., and Glimcher, L. H. (1993) *Dev Dyn* **197**(2), 146-156
219. Tirasophon, W., Welihinda, A. A., and Kaufman, R. J. (1998) *Genes Dev* **12**(12), 1812-1824
220. Hong, M., Luo, S., Baumeister, P., Huang, J. M., Gogia, R. K., Li, M., and Lee, A. S. (2004) *J Biol Chem* **279**(12), 11354-11363
221. Liu, C. Y., Wong, H. N., Schauerte, J. A., and Kaufman, R. J. (2002) *J Biol Chem* **277**(21), 18346-18356
222. Werner, T. (1999) *Mamm Genome* **10**(2), 168-175
223. Azizkhan, J. C., Jensen, D. E., Pierce, A. J., and Wade, M. (1993) *Crit Rev Eukaryot Gene Expr* **3**(4), 229-254
224. Jensen, D. E., Black, A. R., Swick, A. G., and Azizkhan, J. C. (1997) *J Cell Biochem* **67**(1), 24-31
225. Noe, V., Chen, C., Alemany, C., Nicolas, M., Caragol, I., Chasin, L. A., and Ciudad, C. J. (1997) *Eur J Biochem* **249**(1), 13-20
226. Gartel, A. L., and Tyner, A. L. (1999) *Exp Cell Res* **246**(2), 280-289