Charakterisierung des 95 kDa Spermienadhäsions proteins

Dissertation
zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultäten
der Georg-August-Universität zu Göttingen

Vorgelegt von Leonard Bull
aus Yengema Freetown/Sierra Leone

Goettingen 2003

D7

Referent: Prof. Dr. H.-J. Fritz

Korreferent: Prof. Dr. med. W. Engel

Tag der mündlichen Prüfung: 2004-01-22

PUBLICATION:

1. Bull, L., S, Jansen, C, Knorr, F, Herbarmann, B, Brenig. Assignment of the sperm protein zona receptor tyrosine kinase gene (SPRMTK)to porcine chromosome SSC3q11-q12 by fluorescence *in situ* hybridization and by analysis of somatic cell and radiation hybrid panels. Cytogenet. Cell. Genet. 2003;101 (2):178

I LIST OF CONTENTS:

II Abbreviations:

AMV = avian myeloblastosis
approx. = approximately
APS = ammonium persulfate
bp = base pair(s)
BSA = bovine serum albumin
°C = degree Celsius
CAM = chloramphenicol
cDNA = complementary DNA
cM = centi Morgan
cm = centimetre
cpm = counts per min
CIA = chloroform isoamylalcohol
Da = Dalton
DEPC = diethyl pyrocarbonate
dATP = deoxyadenosine triphosphate
dCTP = deoxycytidine triphosphate
dGTP = deoxyguanosine triphosphate
dTTP = deoxythymidine triphosphate
dNTPs = deoxyribonucleic triphosphates
DNA = deoxyribonucleic acid
EDTA = ethylenediamine tetra acetic acid
EtBr = ethidium bromide
FISH = fluorescence *in situ* hybridization
Flqter = fractional length
G = gram
x g = centrifugation force
h = hour

Hcl = hydrochloric acid
HUMSPMRTK = *Homo sapiens* sperm protein receptor tyrosine kinase
H_2O = distilled water
INRA = Institut National de la Recherche Agronomique
Kb = kilobase pair(s)
kDa = kilo Dalton
L = litre
LB = Luria-Bertani medium
LD PCR = long distance polymerase chain reaction
M = molarity
MOPS = 3-(N-morpholino)propanesulfonic Acid)
Min = minutes
NaCl = sodium chloride
NaOH = sodium hydroxide
nm = nano metre
OD = optical density
o/n = over night
PAC = P1-derived artificial chromosome
PCR = polymerase chain reaction
PP = plate pool
PTK = protein tyrosine kinase
RACE = rapid amplification of cDNA ends
RNA = ribonucleic acid
rRNA = ribosomal ribonucleic acid
Rnase = ribonuclease
rpm = revolutions *per* min
RT = reverse transcription
SDS = sodium dodecyl sulfate
Sec = second
SP = super pool
SPMRTK = sperm protein receptor tyrosine kinase
SSPMRTK = swine sperm protein receptor tyrosine kinase
SSC = saline-sodium citrate buffer
SSP = super/super pool
TBE = Tris-borate-EDTA buffer

TE = Tris-EDTA buffer
TEMED = N,N,N,N-tetra methyl ethylene diamine
TM = Tris-$MgCl_2$ buffer
Tm = annealing temperature
Tris = Tris(hydroxymethl)amino methane
U = unit
µl = micro litre
µM = micro Molar
UV = Ultraviolet
V = Volt
v/v = volume *per* volume
w/v = weight *per* volume
XCFF = xylene cyanol FF
X-Gal = 5-Bromo-4-chloro—3-indolylß-D galactopyranoside
ZP = Zona pellucida

III Table and Figure captions:

1 Introduction

1.1 Early stages of fertilization

The bringing together of two gametes in mammals is a complex process that starts with cells moving through the reproductive tracts of both the male and female organisms until they arrive close to each other in the female reproductive tract. The subsequent interaction between the two gametes result in gamete fusion to produce a zygote. This union involves a series of carefully orchestrated cellular interactions; involving moieties associated with the nuclear and cytoplasmic components of both haploid germ cell. This interaction is mediated by a number of specialised molecules; Adhesion, signalling and effector molecules are located in the membrane and on the upper surface layers. The creation of an embryo involves not only adhesion (or binding) of the two morphologically different cell types (spermatozoa and oocyte) but also membrane fusion, accompanied by signal transduction events that trigger cell cycle progression and transform the egg from its quiescent state to an activated one-cell embryo. In mammals, the initial interaction between the oocyte and the spermatozoa occurs at an egg-specific extracellular matrix, the zona pellucida (ZP). The matrix is composed of three glycoproteins, which are encoded by three different genes (ZPA, ZPB and ZPC). The zona proteins involved have been characterized in some detail, with ZP3 and ZP2 generally acknowledged for the initial (primary) and secondary interactions, respectively. Contrary to the female gamete, the spermatozoa posses an overabundance of molecules some of which have been defined both chemically and biologically. The presence of receptors (spermatozoa) and ligands (oocyte), however, is not the only prerequisite for

successful fertilization; their stereochemistry plays an essential role in the recognition and fusion of both gametes as introduced by proacrosin, a sperm adhesion ligand (Jansen *et al.*, 1995, 1993).

1.2 Morphology of the oocyte

The zona pellucida is a porous extracellular network-like matrix enveloping the oocyte and the early embryo that exerts several important functions during fertilization and early embryonic development. It forms a viscous border between oolemma and the innermost layer of follicle cells. Each of the glycoproteins is heterogeneous with respect to molecular weight (Mr) due to extensive and heterogeneous glycosylation of a unique polypeptide with both asparagines (N)-linked (complex-type) and serine/threonine-(O)-linked oligosaccharides (Wassarman *et al.*, 1998). The oligossacharides are both sulphated and sialylated (Noguchi and Nakano, 1992; Lui *et al.*, 1997), which additionally contributes to Mr heterogeneity (see table 1) and makes all three glycoproteins relatively acidic (Nakano and Yonezawa, 2001). The most obvious interspecies difference is the thickness of the ZP varying from 1 to 2 μm in the opossum, 5 μm in mouse, from 13 to 16 μm in human and pig and 27 μm in cow (Dunbar *et al.*, 1989, 1994). The ZP of most mammalian species, build its typical fibrogranular structure (Wassarman and Mortillo, 1991) by noncovalent interactions. The zona pellucida provides receptors for the relatively species-specific attachment, likewise species-selectivity (Jansen *et al.*, 2001) and binding of capacitated sperm and is also involved in the subsequent induction of the sperm acrosome reaction (Florman and Storey, 1982; Cherr *et al.*, 1986; Berger *et al.*, 1989). In most species, certain exposed oligosaccharide chains of the ZP and complementary carbohydrate binding proteins on the sperm-oocyte interface

mediate, at least in part, the initial binding and recognition between the sperm and the ZP. Removal of the ZP eliminates the barrier to in vitro fertilization of egg by sperm from different species (Wassarman *et al.*, 1998). After penetration of the first spermatozoon, the ZP is modified and serves as a site of secondary block to pathological polyspermy, a lethal condition (Wassarman, 1982; Yanagimachi, 1988, 1994). In addition to its role in fertilization the matrix proteins of the ZP appear to be important for the organisation and differentiation of granulosa cells and folliculogenesis and provide protection from physical and environmental damage.

Table 1 Molecular weights (Mr) of ZP proteins in different mammalian species

Species	Gene family /ZP protein	Molecular weight, kDa	References
Human	ZPB/ZP1	80-92	Bercegeay *et al.*, (1995)
	ZPA/ZP2	58-66	Harris *et al.*, (1994)
	ZPC/ZP3	54-62	
Mouse	ZPB/ZP1	200	Bleil & Wassarman (1980)
	ZPA/ZP2	120	Harris *et al.*, (1994)
	ZPC/ZP3	83	
Rabbit	ZPB/ZP1	68-125	Dunbar *et al.*, (1989)
	ZPA/ZP2	81-100	Harris *et al.*, (1994)
	ZPC/ZP3	100-132	
Pig	ZPA/ZP1	80-90	Hedrick & Wardrip (1986)
	ZPA/ZP2	60-65	Harris *et al.*, (1994)
	ZPB/ZP3α	55	
	ZPC/ZP3	55	
	ZPA/ZP4	20-25	
Cow	?/ZP1	78	Noguchi *et al.*, (1994)
	ZPA/ZP2	64	Harris *et al.*, (1994)
	ZPB/ZP3α	21	
	ZPC/ZP3	21	

1.3 Sperm morphology

The sperm cells are specialised cells, generated in the testes and then stored in the epididymis until maturity. They are extremely streamlined and highly polarised cells and contains only elements for essential functions such as motility and a few critical enzymes to ensure efficient transmission of the paternal genome. Each sperm consists of two morphologically and functionally distinct regions. A head containing an usually highly condensed haploid nucleus and a tail. The DNA in the nucleus is inactive and extremely tightly packed as a result of its association with highly charged proteins, the protamine. The head also contains a large secretory vesicle called acrosome localized as a cap-like structure at its tip. The mid-piece between the head and tail forms the power-house, containing mitochondria and a single centriole. The mitochondria (approx. 50-100), which provide the energy for sperm metabolic needs, are derived from the maternal side; they have been proved indispensable in tracing human lineages.

1.3.1 Sperm adhesion molecules involved in fertilization prior to fusion

The adhesion molecules mediate a critical role in binding of the gametes. Therefore, an essential task in the study of sperm-egg interaction is an examination of the capabilities of a distinct set of surface proteins. On gametes, these proteins act in a sequential pattern to orchestrate the close approach and ultimate fusion of the two cells. A plethora of these molecules has been located on the male gamete of which some have already being structurally elucidated. In most mammals that have been studied, an acrosome-intact spermatozoon usually makes contact with the egg to undergo a specific gamete recognition process. Gamete

recognition is mediated by complementary molecules associated with the extracellular matrix of the oocyte (zona pellucida) and the plasma membrane on the sperm head. Subsequent to initial (primary) sperm-zona binding, the fertilizing spermatozoon undergoes the acrosome reaction. This exposes the acrosomal contents of the spermatozoon, the fertilizing process will then undergo the secondary sperm-zona binding before penetration of the zona pellucida.

1.3.1.1 Sperm proteins involved in primary interactions with ZP3

ZP3 is the main oocyte protein involved in the initial (primary) sperm-zona binding in most species. By definition, proteins involved in primary binding are located on the principal segment of the plasma membrane overlying the sperm's acrosome and will be lost from the cell as a consequence of acrosomal exocytosis. Despite this cautionary note, investigators have been successful in identifying sperm proteins involved in primary ZP3 binding, and the five following sperm proteins are discussed below:

Spermadhesins

Spermadhesins belong to a novel family of low molecular weight (12-16 kDa) secretory proteins that are expressed in the male epididymal tract. Spermadhesins are major products of seminal plasma and has been characterized most extensively in pigs/swine. They are believed to act as primary sperm-binding proteins owing to their capacity to interact with both O- and N-linked oligosaccharides and, in addition, they possess heparin-binding activity (Töpfer-Petersen *et al.*, 1998).

Sp56, a peripheral membrane protein

A mouse lectin-like 56 kDa zona oligosaccharide-interacting peripheral membrane protein (sp56), which is a member of a superfamily of protein receptors that includes the α subunits of complement 4B-binding protein, has also shown to interact directly with ZP3 in cross-linking studies. Sp56 is believed to interact with the terminal galactose residues of ZP3 O-linked oligosscharides and is species-restricted (present in mice and hamsters, absent in guinea-pigs and humans) (Bleil and Wassarman, 1990; Cheng *et al*., 1994; Bookbinder *et al*., 1995).

β-1,4-Galactosyltransferase (GalTase)

Galactosyltransferase comprise a group of enzymes involved with the intracellular synthesis of complex carbohydrates in the golgi apparatus and endoplasmic reticulum (McLeskey *et al*., 1998). Although there is no direct evidence for interaction with zona proteins, an isoform of GalTase present on the surface of mouse spermatozoa over the head region form association in a lectin-like manner by specifically recognizing and binding terminal N-acetyglucosamine (GlcNAc) residues of ZP3 O-linked oligo-saccharides (Miller *et al*., 1992).

Zonadhesin

Zonadhesin has been characterized in pig/swine and shown to be similar to the blood von Willebrand factor, with repeats that might facilitate multiple interaction with the zona, possibly by carbohydrate binding. This transmembrane protein has been shown to possess high-affinity species-specific binding to the zona and is the only sperm protein in which such species specificity has been observed at a molecular level (Hardy and Garbers, 1995).

Sperm protein receptor tyrosine kinase (*The 95 kDa tyrosine kinase receptor)*

The 95 kDa tyrosine kinase receptor, which is our candidate gene was first characterized in mice and later in humans. It has been shown to be phosphorylated in response to zona protein and to bind ZP3 directly (Leyton and Saling, 1989b). In humans, a monoclonal antibody (mAb97.25) has been shown to inhibit sperm-zona binding and immunoprecipitation experiments revealed that this antibody recognized a 95 kDa phosphotyrosine-containing protein expressed only in testicular germ cells (Moore *et al*., 1987; Burks *et al*., 1995). The 95 kDa protein exhibits phosphorylation on tyrosine residues; the level of tyrosine phosphorylation increases 4-fold when sperm are exposed to ZP, indicating that a protein tyrosine kinase (PTK) is activated in response to ligand binding (Leyton and Saling, 1989b), leading to the biochemical pathway that results in the acrosome reaction. There is also evidence that such a moiety is somehow involved in initial gamete recognition (Burks *et al*., 1995; Brewis *et al*., 1998). Thus, these findings raise the possibility that the transducing machinery used by the sperm to convey extracellular signals to the cell interior may be similar to that used by somatic cells possessing PTK receptors. For the later type of receptor, ligand binding to the extracellular domain stimulates receptor aggregation (Ikari *et al*., 1988; O'Brien *et al*., 1987), which activates an intrinsic tyrosine kinase and leads to phosphorylation of the receptor as well as other intracellular substrate (Ullrich *et al*., 1990; Pawson, 1992). This phosphorylation effect appears to be conserved in all mammalian spermatozoa.

1.3.1.2 Sperm proteins involved in secondary interactions with ZP2

Until this date, the component of the acrosome-reacted spermatozoon involved in secondary binding with ZP2 remains unidentified. However recent developments suggests that the acrosome reaction may sequentially release and/or expose specific components of the acrosomal matrix, which may stabilize sperm adhesion to the zona matrix. The most likely candidate for a ZP2 receptor in the spermatozoon is proacrosin, although the evidence is actually not conclusive (Brewis and Moore, 1997).

Proacrosin

Acrosin is localized in the acrosomal matrix as an enzymatically inactive zymogen, proacrosin (Notland *et al.*, 1989; Hardy *et al.*, 1991; Westbrook-Case *et al.*, 1994).
It exhibits two distinct activities:
I). Enzymatic activity as a trypsin-like serine protease
II). A lectin-like carbohydrate-binding activity.
Thus, acrosin has been postulated to function both in the limited proteolysis of the zona pellucida, and in the maintenance of binding of acrosome-reacted sperm to the zona (Töpfer-Petersen *et al.*, 1985; Töpfer-Petersen and Henschen 1987; Fries *et al.*, 1987). These properties, binding mechanism and the enzymic activities of the acrosin, are located at different sites of the molecule (Töpfer-Petersen *et al.*, 1990b; Urch and Patel, 1991; Jones and Jansen, 1993; Francavilla *et al.*, 1994).

PH-20, a glycosyl phosphatidylinositol

PH-20 is a glycosyl phosphatidylinositol-anchored membrane protein initially found on guinea pig sperm (Phelps *et al.*, 1988). In these cells, PH-20 is represented in two

populations: on the sperm surface in the postacrosomal region of the sperm head as well as within the acrosome. In accordance to these two areas of localization, the PH-20 polypeptide displays two distinct activities and appears to play two independent roles during sperm-egg interaction. PH-20 demonstrates hyaluronidase activity (Hunnicutt *et al.*, 1996) and the intra-acrosomal population participates in secondary binding to the ZP (Benoff, 1997)

1.4 Molecular Basis (/Mechanism) of sperm and egg interaction

The sperm-zona interaction might occur either as protein-protein, protein-carbohydrate, carbohydrate-carbohydrate interactions or a combination of these binding phenomena.
The question here is how do free-swimming sperm and ovulated eggs from the same species recognize one another and bind to each other? The answer to these questions is of great interest since binding of sperm to eggs initiates a pathway that leads to fertilization of eggs and development of a new individual of the species. In mammals, fertilization begins with the binding of sperm to the oligosaccharides coat of ovulated eggs, called the zona pellucida (Gwatkin, 1977; Wassarman, 1987, 1999a; Yanagimachi, 1994; Snell and White, 1996). The binding of sperm to eggs in mammals is a carbohydrate-mediated event (Wassarman, 1992; Litscher and Wassarman, 1993; Miller and Shur, 1994; Chapmann and Barratt, 1996; Benoff, 1997; Tulsiani *et al.*, 1997; Sinowatz *et al.*, 1998; Töpfer-Petersen, 1999a; Wassarman, 1999a; Prasad *et al.*, 2000). Apparently, sperm recognize and bind to specific oligosaccharides associated with a specific egg-zp glycoprotein. It is proposed that changes in the structure of these oligosaccharides (e.g., their composition, sequence, linkage, and/or modification) together with changes in the egg-binding protein(s) on sperm

that recognize ZP3 oligosaccharides, account for species-specific binding.

1.4.1 Physiological aspects

The events that occur during the early stages of fertilization may be viewed as an exquisite example of the specificity of intercellular and cell surface-extracellular matrix interaction. Cell-cell recognition mechanisms in many somatic cell systems involve carbohydrate side chains of membrane glycoproteins (Gabius, 1998; Gabius and Sinowatz, 1998; Solis *et al*., 2001) and several observations indicate that similar molecules may play a pivotal role in spermatozoa-oocyte binding in mammals. Compelling evidence has now demonstrated that carbohydrate-binding proteins on the sperm surface mediate gamete recognition by binding with high affinity and specificity to complex glycoconjugates of the zona pellucida (Yanagimachi, 1994; Chapmann and Barratt, 1996). The ZP3, involved in primary binding, is the principal physiological inducer of the human sperm acrosome reaction.

1.4.2 Biochemical aspects

A very strong bond is achieved through interaction of the ZP3 and complementary sperm-binding proteins present in the plasma membrane. The ZP3 triggers the acrosome reaction, which is then followed by a secondary binding process involving the ZP2 and the inner acrosomal sperm membrane leading to zona penetration (Bleil and Wassarman, 1980, 1983, 1988; Nixon *et al*., 2001; Wassarman and Litcher, 2001). The acrosome reaction-triggering activity of ZP3 seems to depend upon the integrity of the protein backbone (Wassarman, 1990; Saling, 1991; Nixon *et al*., 2001;

Wassarman and Litcher, 2001). Peptides synthesized based upon the published DNA sequence of ZP3 proteins are able to induce acrosomal exocytosis in some species (Hinsch *et al.*, 1998).

1.5 Objective of this study

The adhesive properties of gametes are key to their changing interactions with various surfaces contacted on the route to fertilization. Thus understanding the molecular basis of the gamete cross talk and gamete-oviduct interactions in mammals needs to be explained. Furthermore, the detailed, knowledge of the genetic and molecular aspects underlying these adhesive molecules is of importance in providing the basis for understanding the mechanism of gamete interaction.

Here, our aim is to characterize a 95 KDa sperm adhesion protein in *porcine*, which has already been identified in *Homo sapiens* and *murine* as a putative ligand for the ZP glycoprotein, ZP3 (Burks *et al.*, 1995). The 95 kDa sperm protein is a tyrosine kinase subtrate, with phosphorylation on tyrosine stimulated upon zona pellucida binding (Leyton and Saling, 1989b). The initial focus of this project is the physical mapping of the 95 kDa sperm adhesion gene. We intend to isolate the gene by screening a *porcine* PAC library and subsequently sequence the 95 kDa sperm adhesion gene. There is no doubt that genetic factors play an important role in understanding infertility in mammals. It therefore underscores the need to locate the chromosomal sites and structure of the genetic determinants responsible for fertilization in order to rationally intervene in better treatment of gene-related infertility. We intend to create the cDNA-sequence of the gene using the recombination DNA-technique and then transfer into an appropriate expression system: eukaryotic, prokaryotic or mammalian

expression system. We aim more so to express the catalytic domain where the phosphotyrosine residues are located, which is stimulated when binding occurs. We intend also to carry out expression studies to find out in which tissue/tissues the protein is expressed.

Future perspective:

We hypothesise that binding affinity to the complementary gamete – either permanently or temporally – would become explainable, if the assessments of the effect of phosphorylation and non-phosphorylation conditions upon binding are studied. Setting up binding assays, binding of SPRMTK-ZP3 could be examined using the aforementioned conditions.

Applying the mass spectroscopy, circular dicroism spectroscopy and fourier transform infrared spectroscopy, the secondary structure and thus folding of the protein could be explained which is important in explaining gamete binding.

Also the linkage analysis of the gene, in which the gene is assigned to SSC3 (Bull *et al.*, in press) and obtaining a detailed knowledge of the primary structure of the sperm receptor SPRMTK is not only interesting for scientific reasons; it is also significant in our thinking for the diagnosis of idiopathic sterility and serve as a rational starting point for a sperm-targeted contraceptive development. This applies to human and certainly in a modified form to domestic, wild and zoo animals.

2 General Materials and Methods

This chapter describes the general materials and methods used in more than one chapter in this thesis. Common laboratory techniques, materials, and general procedures for preparation of reagents presented here are based on Molecular Cloning: A laboratory Manual (Sambrook *et al.*, 1989), unless otherwise mentioned.
All glassware and plastic ware were sterilised by autoclaving at 120°C for 30 min. Chemical reagents used were Analytical Reagent (AR) or Molecular Biology grades unless otherwise specified.

2.1 Materials

2.1.1 Bacteria, Plasmid and Gene Bank

All bacteria, plasmids and Gene Bank used in this project belong to the Institute of Veterinary Medicine, University of Goettingen.

2.1.1.1 Bacteria

E.coli XL1-Blue
(*mcr*A)183 (*mcr*CB-*hsd* SMR-*mrr*)173, *end*A1, *sup* E44, *thi*-1 *gyr*A96, *tell*A1, lac[F'*pro*AB, lacI^{q}ZM15, Tn10(Tet^{r})] from Stratagene, Heidelberg

E.coli Top 10
F-*mcr*A, (*mrr-hsd*RMS-*mcr*BC), 80*lac*ZM15, *lac*X74, *rec*A1 *deo*R, *ara*D139, (*ara-leu*)7697, *gal*K, *rps*L, (Str^r), *end*A1, *nup*G from Invitrogen, Groningen

E.coli BL21(DE3)pLysS
F', ompT, hsdSB(r_B^- m_B^-)*gal*, dcm (DE3), pLysS, (Cam^R), from Invitrogen, Groningen

2.1.1.2 Plasmid

PGEM®4-Z from Promega GmbH, Mannheim
PCR®2.1-T0PO® from Invitrogen, Groningen
PRSET A, PRSET B, PRSET C from Invitrogen, Groningen

2.1.1.3 Gene Bank

Porcine PAC Gene Bank (IVM PAC 714) (Al-Bayati *et al*., 1999)

2.1.2 Enzymes, DNA-Marker, Nucleotides, Chemicals and Kits

2.1.2.1 Enzymes and DNA-Marker

Enzymes were obtained from the following companies:
Boehringer-Mannheim GmbH, Mannheim
New England Biolabs, Schwalbach
Amersham Pharmacia Biotech Europe GmbH, Freiburg
Life Technologies, Karlsruhe
Qiagen GmbH, Hilden
Roche Diagnostics GmbH, Mannheim

Roth GmbH, Karlsruhe
GibcoBRL Life Technologies GmbH, Karlsruhe

2.1.2.2 Nucleotides

All primers were synthesized with highest purity by MWG-Biotech GmbH (Ebersberg). The dNTPs were purchased from Roth GmbH (Karlsruhe).

2.1.2.3 Chemicals

Chemicals were obtained from the following companies:
Bio-Rad Laboratories, MÜnchen
Fluka Feinchemikalien, Neu-Ulm
Merck AG, Darmstadt
Amersham Pharmacia Biotech Europe GmbH, Freiburg
Roth GmbH, Karlsruhe
Qiagen GmbH, Hilden
FMC BioProducts, Dänemark
Sigma Chemie, München

2.1.2.4 Kits

Amersham Pharmacia Biotech Europe GmbH, Freiburg: ECL DNA Direct Labelling System, Thermo Sequenase Fluorescent Labelled Cycle Sequencing Kit®
Qiagen GmbH, Hilden: Plasmid purification kit, QiaexII gel extraction Kit
Stratagene GmbH, Heidelberg: cDNA Synthesis kit
BD Biosciences Clontech, Heidelberg: Super Smart Race cDNA synthesis Kit

Clontech Laboratories GmbH, Heidelberg: Advantage® 2 PCR Enzyme System
Invitrogen, Groningen: TOPO TA Cloning® Kit

2.1.3 Equipment

Agarose gel tanks: MWG Biotech GmbH, Ebersberg
Bacterial shaker: A. Kühnert AG, Bottingen
Biofuge 28S, Megafuge 1.0R: HaraeusChrist, Osterode
Blotting filter paper: Schleicher & Schuell, Dassel
Blotting membrane: Schleicher & Schuell, Dassel
Blotting membrane, Nylon: Stratagene, Heidelberg
Sorvall RC 58 Centrifuge: Kendor Laboratory products GmbH, Hamburg
Eagle Eye II: Stratagene, Heidelberg
Pulsed field gel electrophoresis: Bethesda Research Laboratories, USA
Hybond-XL, nylon membrane: Schleicher & Schuell, Dassel
Photospectrometer: Perkin Elmer (UV—VIS 550), Weitestadt
Photo Developer: Agfa-Gevaret, Leverkusen
Pipette, 1-10 µl, 20-200 µl, 200-1000 µl: Eppendorf, Hamburg
Power supply (agarose gels): Pharmacia, Freiburg
Robo Cycler Gradient 2000: Stratagene, Heidelberg
UV-Transilluminator: Bachofer Labor equipments, Reutlingen
Vortex Genie 2: Bender & Hobein, Zürich
Water bath DC3: Haake, USA
X-ray films, Kodak X-OMAT: Sigma (F-5513), USA

2.1.4 Preparation of buffers

2.1.4.1 Commonly used buffers

Tris-HCL

To make 1M Tris-HCL with a required pH (7.5 or 8.0), 121.1 g of Tris base was initially dissolved in 800 ml distilled water and the pH was adjusted to the desired value by adding concentrated HCL at room temperature. The volume of the solution was then adjusted to 1000 ml by adding distilled water and the mixture was then autoclaved at 120°C for 30 min.

EDTA

The stock solution of 0.5 M EDTA (pH 8.0) was made by adding 186.1 g of EDTA to an initial volume of 800 ml distilled water. About 20 g sodium hydroxide (NaOH) was added to hasten dissolving of EDTA. The solution was stirred on a magnetic stirrer for at least 3h. Finally, the pH was adjusted to 8.0 with further NaOH and volume was made up to 1000 ml. The solution was then autoclaved at 120°C for 30 min

TBE

A 10 x TBE (Tris-borate EDTA) stock was made by mixing 108 g of Tris base, 55 g of boric acid and 40 ml of 0.25 M EDTA (pH 8.0) and then adjusting to final volume of 1000 ml by adding distilled water.

SSC

A 20 x SSC stock was prepared by mixing 175.3 g of NaCl and 88.2 g of sodium citrate in 800 ml distilled water. The solution was then adjusted to pH 7.0 with few drops of NaOH.

The final volume of 1000 ml was obtained by adding distilled water.

2.1.4.2 Buffers used for DNA extraction

Cell lysis buffer

To make 100 ml of cell lysis buffer, 1 ml of Triton X-100 (1% final concentration), 100 µl of 1 M Tris-Hcl pH 8.0 (1 mM final concentration) and 500 µl of 1 M $MgCl_2$ (5 mM final concentration) were mixed, and distilled water was added to adjust the total volume to 100 ml. Prior to use, 32 ml 1 M sucrose was added to 68 ml of the solution (0.32 M final concentration). This buffer was freshly made and used for lysing cells when genomic DNA was extracted from blood samples.

Solution I

A stock of 100 ml solution I was prepared by mixing 2 ml of 0.5 M EDTA pH 8.0 (10 mm final concentration) and 2.5 ml of 1 M Tris-Hcl pH 8.0 (25 mM final concentration). The final volume was adjusted to 100 ml with distilled water. This buffer was used for extracting DNA from bacterial cells. Prior to use 100 g/ml DNase free RNase A was added and storage at 4°C.

Solution II

This solution was freshly made immediately before use for extraction of plasmid DNA from bacterial cells. 1 ml stock of 10% SDS was mixed with 2 ml of 1 M NaOH. Distilled water was then added to obtain a 10 ml solution II, containing 1% SDS and 0.2 M NaOH.

Solution III

This solution was also required for extraction of plasmid DNA from bacterial cells. To prepare a 100 ml of this solution, 11.5 ml of acetic acid (11.5% final concentration) was added to 60 ml stock of 5 M potassium acetate (3 M final concentration) and adjusted with distilled water.

Buffer QBT

To prepare 1 litre of this solution, 43.83 g NaCl, 10.46 g MOPS were dissolved in 800 ml distilled water. The pH was adjusted to 7.0. 150 ml pure ethanol was added and then 15 ml 10% Triton X-100 solution. The volume was filled up to 1 litre with distilled water.

Buffer QC

1 litre of this solution, was prepared by dissolving 58.44 g NaCl and 10.46 g MOPS in 800 ml distilled water. The pH was adjusted to 7.0. 150 ml pure ethanol was added. The volume was filled up to 1 litre using distilled water.

Buffer QF

To prepare a 1 litre of this solution, 73.05 g NaCl and 6.06 g Tris base were dissolved in 800 ml distilled water. The pH was adjusted to 8.5 with HCl. 150 ml pure ethanol was then added. The volume was adjusted to 1 litre with distilled water.

CIA

Chloroform was mixed with isoamylalcohol in the ratio 24:1 (w/v). This solution was required to extract DNA from blood.

Phenol/CIA

In this case phenol, chloroform and isoamylacohol were mixed thoroughly in the ratio 25:24:1. This solution was also used in the extraction of DNA from blood.

2.1.4.3 Gel loading buffers

Agarose gel-loading buffer

To make 10 ml of 10 X stock: 0.025 g Bromophenol blue (0.25% final concentration), 0.025 g Xylene Cyanole FF (0.25% final concentration) and 5 ml of 30% Ficoll (15% final concentration) were mixed and distilled water was added to adjust the final volume to 10 ml.

RNA loading dye

To prepare 10 ml of 10 X stock; 0.025 g Bromophenol blue (0.25% final concentration), 0.025 g Xylene Cyanole FF (0.25% final concentration), 1 mM EDTA, 50% glycerol were mixed. The volume of 10 ml adjusted with distilled H_20.

2.1.5 Preparation of media

LB medium

For preparing 1 litre LB medium, 10 g Bacto-Tryptone (1% final concentration) 5 g Bacto-Yeast extract (0.5% final concentration) and 10 g NaCl (1% final concentration) were dissolved in 800 ml distilled water. The pH of the medium was adjusted to 7.0 by adding NaOH and the final volume was increased to 1 litre by adding distilled water. The medium was autoclaved at 120°C for 30 min and stored at 4°C.

LB plate with the appropriate antibiotic

In order to prepare LB plates 35 g LB-agar was dissolved in 1 litre distilled water. The solution was autoclaved for 30 min at 120°C and then cooled down to 50°C. 1 ml of the appropriate antibiotic was added (50 μg/ml final concentration) and mixed. This was poured into Petri dishes

and allowed to harden. Storage was done in an upside-down position at 4°C until used.

2.2 METHODS

2.2.1 Screening of *porcine* PAC library

2.2.1.1 Organization of the PAC library

The P1-derived artificial chromosome (PAC) library was constructed with *Mbo*I partially digested male Landrace pig genomic DNA (insert size of 120 Kb) that was cloned into the PAC vector pCYPAC 2. As illustrated by Al-Bayati *et al*. (1999), the PAC library is organized using a standard pooling scheme: The 90,240 clones are arrayed in 560 96-well microtiter master plates and 95 384-well microtiter master plates. From each microtiter master plate, DNA is pooled into ten plate pools (PP). Which are further combined into one super pool (SP). Five SPs are pooled into one super/super pool (SSP) bringing the number to 13 SSP. Identification of a positive PAC clone in a microtiter plate is performed by two-dimensional PCR screening of the 8-row and 12-column pools

2.2.1.2 PCR screening of the PAC library

Screening of the PAC library was performed by polymerase chain reaction of hierarchical DNA pools. Initially the SSPs were screened with specific oligonucleotides. Then the five SPs comprising the positive SSP were screened to determine the positive PPs. Once the appropriate PP was identified, a final round of amplification of the 8-row and 12-column pools was performed. The intersection of the appropriate row

and column on the microtiter plate identified the correct clone. Genomic DNA of a Large White pig was used as a positive control. A PCR reaction mixture without any DNA was used as a negative control.

2.2.2 Isolation of DNA

2.2.2.1 Extracting genomic DNA from blood

5 ml *porcine* whole blood sample was diluted using 10 ml 1 x SSC. After centrifugation for 10 min at a setting of 10.000 X g at room temperature the supernatant was discarded. The pellet consisting of leucocytes was suspended in 10 ml 1 x SSC and then centrifuged at maximum speed for 1 min. This step was repeated until the pellet of just leucocytes had a white/beige colour. The pellet was resuspended in 5 ml 0.2 M NaO-acetate, pH 7.0 and 0.5 ml 10% (v/v) SDS was added. An equal volume of phenol/CIA was added to the mixture. The different phases were separated by centrifugation at 10.000 x g for 10 min at room temperature. The aqueous phase was poured into a new tube and extracted once more by adding 5 ml CIA. The supernatant was transferred into a fresh tube and the organic phase discarded. Subsequently, an equivalent volume of isopropanol was carefully layered on the top, and the solution was gently mixed to precipitate the DNA. The DNA was spooled out, rinsed with 70% ethanol, partially dried in a gentle stream of air and finally dissolved using an appropriate volume of TE buffer.

2.2.2.2 Maxi-preparation of PAC DNA

The PAC DNA was extracted using the Qiagen-tip 500 Plasmid Purification kit (Qiagen). The protocol is based on a modified alkaline lysis procedure; the binding of PAC DNA to an anion-exchange resin under appropriate low salt range and pH conditions. RNA, protein, dyes and low-molecular-weight impurities are selectively removed by a medium-salt wash. PAC DNA is eluted in a high-salt buffer and then concentrated and desalted by isopropanol precipitation. The isolation was performed according to the manual provided with the kit. Nevertheless, before centrifugation the cell culture was divided into four polypropylene screw-cap centrifuge tubes of 150 ml. After precipitation the PAC DNA was then rinsed with 70% ethanol, dried briefly, and dissolved in 100.0-150.0 µl 10 mM Tris-HCL, pH 8.0 at room temperature o/n.

2.2.2.3 Mini-preparation of plasmid DNA

The alkaline minilysis method of Birnboim and Doly (Birnboim and Doly, 1979) was adapted for the preparation of plasmid DNA.
5 ml LB medium, supplemented with appropriate concentration of antibiotic, was inoculated using a single colony off a freshly streaked plate. The culture was incubated at 37°C with moderate agitation of 250 rpm for 10-18 hours until the cells were midway through the logarithmic phase of growth. The culture was then centrifuged at 4°C for 10 min at 6000 x g. The supernatant was discarded and the tube was inverted briefly on a paper towel to remove remaining liquid. The residual cell pellet was thoroughly resuspended in 150 µl of ice-cold miniprep solution I (see section

2.1.4.2) and then transferred into a 1.5 ml micro centrifuge tube. To lyse the cells, an equal volume of miniprep solution II (see section 2.1.4.2) was added to the cell suspension. The tube was mixed gently by inversion to ensure that the entire surface of the tubes came in contact with solution II, and then incubated for 5 min at room temperature. To precipitate proteins and cell debris, 250 µl of ice-cold miniprep solution III (see section 2.1.4.2) was added and then mixed. The tube was chilled in ice for 15 min to allow most of the protein, high molecular weight RNA and chromosomal DNA to precipitate. The tube was centrifuged at 10.000 x g for 20 min to yield a clear supernatant. The supernatant, which contains the plasmid DNA, was transferred to a fresh micro centrifuge tube. To precipitate the plasmid DNA, 1 ml absolute ethanol was added to the tube. This was mixed thoroughly by vortexing and then centrifuged at 10.000 x g for 15 min. The DNA pellet was rinsed with 1-20 ml 70% ethanol and then partially dried in vacuum for 10 min. The pellet was dissolved in 40—50 µl TE at room temperature.

2.2.2.4 Quantification of DNA

2.2.2.4.1 Spectrophotometric measurement

DNA concentration was determined with a spectrophotometer measuring the optical density (OD) at a wavelength of 260 nm and 280 nm. An OD_{260} of 1 corresponds to 50 µg/ml of double stranded DNA (Sambrook *et al.,* 1989). The OD_{260} and OD_{280} of the diluted DNA sample (2 µl DNA, diluted 1:200) were measured. Purified DNA has a ratio of OD_{260}/OD_{280} of approximately 1.8. When the ratio is less than 1.8, it indicates that the sample is contaminated with protein

(Sambrook *et al.*, 1989). The DNA concentrations were calculated as follows:

$$\text{DNA concentration } (\mu g/ml) = \text{measured } OD_{260} \times \frac{50\ \mu g/ml}{1 OD_{260}} \times \text{dilution factor}$$

2.2.2.4.2 Agarose gel quantification

To check the concentration of DNA or determine whether the DNA was sheared or otherwise degraded, 2 µl of extracted DNA was mixed with agarose loading dye and run on a 1% agarose gel at 100 V for 25 min. A single clear band corresponding to that of the DNA fragment size, verifies that the DNA is intact.

2.2.3 Manipulation and analysis of DNA

2.2.3.1 Restriction enzyme reaction

Digestion was performed according to the supplier's recommendations. For the digestion of 1 µg of DNA 10U of the respective enzyme were used. The volume of the enzyme should not exceed more than 10% of the total reaction volume. The following cocktail components were used to digest PCR products or plasmid DNA:

PCR products or plasmid DNA	= 0.5-2 µg
10 x restriction enzyme buffer	= 3 µl
Restriction enzyme	= 10 units
Distilled water up to 30 µl	

Incubation was done at 37°C/65°C for 1h. PAC-DNA was incubated for 16h.

2.2.3.2 Gel electrophoresis of DNA

Agarose gel electrophoresis was used to separate DNA fragments and was performed in a horizontal gel electrophoresis tank. The concentration of the gel used was dependent on the size of DNA fragments to be resolved. Agarose gels were poured containing between 0.8-1% (w/v) in 1 x TBE for genomic DNA and PAC DNA, 1—2% (w/v) for PCR fragments. EtBr was used for staining purposes, which was added to the gel prior to electrophoresis at final concentration of 0.4 g/ml EtBr. The DNA bands were detected by EtBr fluorescence using an ultraviolet light transilluminator. Gels were photographed using the Eagle eye II camera.

2.2.3.3 Pulsed field gel electrophoresis

Pulsed field gel electrophoresis was run using the BioRad CHEF, to analyse the size of the PAC Clone. Three litre 0.5 x TBE was prepared and 1% agarose gel with EtBr 1g/ml was poured. The rest of the buffer 0.5 x TBE was poured into the container making sure that the buffer covered the gel to a depth of 2-3cm. The temperature was equilibrated at 14°C. The PAC Clone was digested using *Not*1; 2 μg DNA was digested using 10U enzyme for 1h incubation at 37°C. The sample was loaded on the gel. The electrophoretic conditions were as follows:

Voltage: 6 volt
Pulse angel: 120°
Pulse/Ramp: 1-30 sec
Running time: 20h

2.2.3.4 Isolation of DNA from gel slices

The isolation of pure PCR products is essential for the cloning reaction and for direct sequencing. For this purpose, a Qiagen II Gel Extraction Kit was used.
Fragments were excised from the gel under UV light and extracted using the Qiagen II Gel Extraction Kit from Qiagen®
(http://www.qiagen.com/litreature/handbooks/qexII/qiaexII-agarose.pdf). Purification of DNA fragments with the QIAEX II system is based on solubilization of agarose and selective adsorption of nucleic acids onto QIAEX II silica-gel particles in the presence of chaotropic salt. DNA was quantified optically by loading 2 µl on an agarose gel or by the use of a photospectrometer.

2.2.3.5 Polymerase chain reaction (PCR)

Polymerase chain reaction was performed in a final reaction volume Of 25.0 µl as described by Saiki *et al.* (1988). The reaction medium consisted of 10–200 ng DNA, 1 x PCR buffer, 200 µM dNTP's, 0.4 µM primers and a variable amount of DNA polymerase, as indicated in the respective product description. After a pre-denaturation of 95°C for 2 min, the PCR was subjected to 25–35 cycles with denaturation at 95°C for 30–45 sec, annealing at the primer pair specific temperature for 30–45 sec elongation at 72°C for 30–90 sec followed by a final elongation step at 72°C for 10 min. In most cases a temperature gradient was performed to obtain the optimum annealing temperature of the primer pair.

2.2.3.6 Sequencing

Automated sequencing was carried out according to the manual supplied with the ABI PRISM sequencing kit. A total of 1.5—2.5 μg of PAC-DNA was added to 4.0 μl of Big Dye Mix, 10 pmoles of primer in a 10.0 μl reaction volume. After 3 min denaturation at 95°C, templates were subjected to 35 cycles of 30 sec at 95°C, 20 sec at 55°C (depending on the Tm of the primer), and 4 min at 60°C. The amount of target DNA in ng derived from PCR products was determined by calculating for every 100 bp of fragment length 5 ng of template.

2.2.3.7 Cytogenetic (Physical) mapping

Physical mapping identifies the location of genes or genetic markers on the individual chromosomes or regions of chromosomes without analysis of segregation between loci. There are three major physical mapping techniques which allow the identification of the chromosomal location of specific DNA fragments, namely use of I) somatic cell hybrids II) radiation hybrids, and III) fluorescence *in situ* hybridization.

2.2.3.7.1 Somatic cell hybrid and radiation hybrid mapping

For mapping of the SPRMTK gene obtained from the PAC library the INRA somatic hybrid panel, containing 27 pig x rodent cell hybrids were used (Yerle *et al.*, 1996) and a *porcine* whole genome radiation hybrid panel (Yerle *et al.*, 1998) of 118 cell lines were screened for *porcine* SPRMTK by PCR. Primers were designed from the published *Homo sapiens*

SPRMTK-mRNA sequence. PCR amplifications (see Section 2.2.3.5) were performed in a total volume of 25 µl with 25 ng of panel DNA as template. PCR results were evaluated using the interpreting web-pages http://imprh.toulouse.inra.fr (radiation hybrid panel) and http://toulouse.inra.fr/lgc/pig/hybrid.htm (somatic cell hybrid) at INRA.

2.2.3.7.2 Fluorescence *in situ* hybridization (FISH)

FISH was performed as described previously by Toldo *et al.*, (1993) and Solinas-Toldo *et al.*, (1995) using swine metaphase spreads (prepared from peripheral lymphocytes) obtained from a normal, healthy boar. Prior to FISH, the QFQ-banded spreads were photographed using a cooled CCD camera. Hybridization signals were detected and amplified by incubation with streptavidin-Cy3 (Rockland, Gilbertsville). The chromosomes were then DAPI counterstained (Sigma, Deisenhofen). The positions of the signals on chromosomes were measured considering relative fractional length from the short arm telomere to the hybridization signal enabling the calculation of the fractional length (Flqter).

2.2.3.8 Ligation

Ligation of PCR products was performed with the pGEM(T)-vector System (Promega, Heidelberg) according to the supplier's manual. Vector DNA (25.0 ng) was incubated with 2 x Rapid Ligation Buffer, 3 units of T4 DNA Ligase and the prepared insert (molar ratio insert:vector = 3:1) at 4°C o/n. Additionally, control DNA provided with the kit was used to determine whether the ligation proceeded

efficiently, while vectors without inserts were used to determine the number of background blue colonies. Using the pGEM™-4Z vector System, incubation was done o/n at 16°C.

2.2.3.9 Transformation

The XL1-Blue *E.coli* was used as host cells in transformation. 50.0 µl was thawed on ice. The ligation reaction of 5 µl was added, swirled to mix DNA evenly and then incubated on ice for 30 min. The cells were heat-shocked by placing the microcentrifuge tubes in a 42°C water bath for 45 sec, and then chilled immediately on ice to quench the heat-shock. 250 µl of pre-warmed LB medium was added and the microcentrifuge tubes were incubated at 37°C for 1h with moderate agitation at 225 rpm. From each tube, 20.0 µl and 200.0 µl were spread on agar plates containing the appropriate antibiotic and incubated at 37°C o/n. For blue-white selection, the agar plates were prepared with 4 mg/ml X-Gal and IPTG before plating the cells. The recombinant plasmids appear white, as the lac-Z gene expression was impossible.

2.2.3.10 Southern Blotting

2.2.3.10.1 Alkali transfer of DNA to a nitro-cellulose membrane

The PAC clone was digested using the appropriate restriction enzymes. According to Reed and Mann (1985), the digested DNA was separated on a 0.8% agarose gel. Denaturising of the DNA was done by treating the gel with 0.4 M NaOH for 10 min. Once denatured, the DNA was then transferred from the

electrophoresis gel to a sheet of nitrocellulose membrane. The transfer was performed o/n. Finally the nitrocellulose paper was removed and the wells marked using a colour pencil. Subsequently, the membrane was neutralized by washing with 2 x SSC and dried briefly by placing it between two Whatmann papers. The single-stranded target DNA was permanently immobilized to the membrane by baking at 80°C for 2h. The membrane was directly used for hybridization.

2.2.3.10.2 Hybridization using the ECL Direct Labelling System

The nitrocellulose membrane was pre-hybridized using the hybridization buffer for 1h at 42°C. The DNA sample used was labelled as follows: the DNA sample with a concentration of 100 ng was denatured at 100°C for 5 min and immediately cooled on ice for another 5 min to prevent renaturation. Then, an equivalent volume of labelling reagent was added and mixed gently. Likewise glutaraldehyde reagent was added and thoroughly mixed to obtain a homogenous mixture. The vial was centrifuged briefly to collect the contents at the bottom. Thereafter, the DNA probe was incubated for 20 min at 37°C for efficient labelling. The labelled probe was added to the pre-hybridized solution and incubated o/n for 12h at 42°C. After hybridization, the membrane was washed; 2 x 5 min with wash buffer I at high stringency (54°C) and subsequently, 2 x 5 min at room temperature with wash buffer II. Finally, the membrane was rinsed briefly with 2 x SSC solution, allowed to dry and then incubated for 1 min in a detection solution. The incubated nylon membrane was placed next to a sheet of X-ray film in a light container. The positions of hybridization signals were quantified via autoradiography. Exposure time varied between 30 min and 4h.

2.2.4 RNA methods

2.2.4.1 RNA extraction

Total RNA was extracted from 0.7 g of tissue using 10 ml Trizol. The tissue sample was first lysed and homogenized in the presence of a highly denaturing guanidine isothiocyanate-containing buffer, which immediately inactivates RNases to ensure isolation of intact RNA. The sample was then incubated at room temperature for 5 min. 2 ml chloroform was added and the mixture was mixed vigorously by hand for 15 sec. This was further incubated for 3 min and then centrifuged at 6.000 x g for 15 min at 4°C. The aqueous upper phase was carefully transferred into a new tube. For precipitation of the RNA, 5 ml isopropanol was added. The tube was inverted several times to mix thoroughly. The mixture was then incubated o/n at -20°C to obtain optimal precipitation. After which, the tube was centrifuged at 12.000 x g for 15 min at 4°C. The supernatant was carefully discarded and 10 ml of 75% ethanol (1 ml 75% ethanol/ml Trizol used) was added to wash off the contaminants. Subsequently, centrifugation was carried out at 7.500 x g for 5 min at 4°C. The supernatant was completely aspirated and the pellet was air dry for 15 min. Finally, the pellet was resuspended in 1 ml RNAse-free water. Aliquots of 200 µl were distributed in RNase-free tubes and stored at -20°C/-80°C.

2.2.4.2 Evaluation of RNA

2.2.4.2.1 Quantification of RNA

RNA concentration was determined with a spectrophotometer measuring the optical density (OD) at a wavelength of 260 nm and 280 nm. An OD_{260} of 1 corresponds to 40 µg/ml of single-stranded RNA (Birren *et al.*, 1997). The RNA concentrations were calculated as follows:

$$\text{RNA concentration (µg/ml)} = \text{measured } OD_{260} \times \frac{40 \text{ µg/ml}}{1 OD_{260}} \times \text{dilution factor}$$

Purity was regarded as sufficient when the ratio OD_{260}/OD_{280} ranged between 1.9 and 2.1.

2.2.4.2.2 Gel electrophoresis of RNA

Before loading on a 1.20% formaldehyde-agarose gel, 2.0 µg of RNA were denatured at 60°C for 15 min in a mixture containing MOPS, formamide and formaldehyde and then cooled on ice. The RNA was separated by electrophoresis in 1 x MOPS. In order to visualize the 18S and 28S rRNAs, the gel was stained with methylene blue and distained in water until the rRNA bands were visible. The 28S ribosomal RNA bands should be present with an intensity approximately twice that of the 18S RNA band.

2.2.4.3 Reverse transcription

Reverse transcription was carried out in a volume of 25.0 µl. A mixture containing 2.5 µg RNA, 3.0 µM poly-A specific

primer and DEPC treated ddH_2O up to 12.7 µl was incubated at 70°C for 5 min and cooled to room temperature for 10 min. The following substances were added: 5 x RT reaction buffer, 250 µM dNTP's, 25U of RNasin, 4 mM Na-pyrophosphate and 15U of AMV-reverse transcriptase. The mixture was incubated for 1h at 42°C, 10 min at 55°C and stopped for 10 min at 72°C. Subsequent polymerase chain reactions were performed using 1.0 µl as template.

2.2.4.4 5'- and 3'- RACE (Rapid amplification of cDNA Ends) using the Super-Smart cDNA synthesis kit and the Advantage® 2 PCR Enzyme System.

Rapid amplification of the cDNA ends (RACE) of the SPRMTK gene was performed. The Super-SMART PCR cDNA synthesis kit is a PCR-based method for producing high-quality cDNA from nanogram quantities of total RNA.

First-strand cDNA synthesis:

First strand cDNA is synthesized from total RNA using 3' SMART CDS primer II A (a modified oligo dT-primer), Power-Script reverse transcriptase and the deoxynucleotide. 1 µg of total RNA was added to 1 µl Smart primer and then incubated for 2 min at 72°C. The mixture was cooled on ice for another 2 min. After which the tube was briefly centrifuged to collect contents at the bottom. The following reagents were then added: 2 µl 5 x first-strand buffer, 1 µl DTT, 50 x dNTP and finally the RT. The tube was then incubated for 1h in a thermal cycler. After the first strand reaction the tube was placed on ice to terminate reaction. Using the NucleoSpin Extraction kit the Purification from unincorporated nucleotides and primers was carried out. Terminal transferase was used to add homopolymeric A-tail to the 3' end of the cDNA.

cDNA amplification by LD PCR:
The following reagents were added to the reaction tube; 10 µl 10 x advantage 2 PCR Buffer, 50 x dNTP mix, 5′ PCR Primer II A, 2 µl 50 x Advantage 2 polymerase Mix. This was well mixed and then briefly centrifuged. The tube was placed in a preheated thermal cycle and PCR carried out as explained in Section 2.2.3.5.

2.2.4.5 Northern Blot analysis

In the Northern blot analysis, total cellular RNA (25 µg) from various pig tissues were subjected to electrophoresis through a 2.2 M formaldehyde gel. The electrophorese tank was filled with 1 x MOPS running buffer and the gel ran for several hours between 100—200 V, until the xylene cyanol dye front migrated 3 to 4 cm into the gel and the bromophenol blue about 2/3^{rds} down the gel. After an appropriate distance, the gel was photographed including a ruler aligned with the wells. Preparation and transfer to a nylon membrane was carried out as in section 2.2.3.10 (Southern blotting) except that the gel was pre-soaked in 20 x SSC before transfer. The blot was hybridized using ECL (see section 2.2.3.10.2) labelled RT-PCR product.

2.2.5 Protein analysis

2.2.5.1 Preparation of the cDNA fragment for subcloning into pRSET vector

By using two mutagenic oligonucleotides with *Hin*dIII and *Bam*HI sites respectively, the fragment of interest was

generated. Since this is not the full-length cDNA-fragment methionine (start codon) is absent at the upstream end. Therefore, the primer designed for this region had a start codon ATG and a *Hin*dIII restriction site incorporated, whilst the downstream primer a *Bam*HI site after a stop codon TGA. The fragment was amplified using a standard PCR protocol, and the product analysed on a 1% agarose gel. The band was excised and purified using QiaexII kit. The fragment was then digested for 1h at 37°C using both enzymes. This was again analysed on a gel, excised and then purified. DNA sequencing was used to verify the presence of the two restriction sites and to assure that no other base substitutions were incorporated. The fragment was then subcloned into a pRSET A vector, which was initially cleaved using the same enzymes to ensure a uni-directional cloning of the fragment. The start codon of the fragment located after the T7 promoter region. Once subcloned the fragment was expressed in a prokaryotic system.

2.2.5.2 Expression in a prokaryotic system

The amplified fragment was subcloned in a pRSET A vector for expression. The targeting vector pRSET A includes a polyhistidine (6 X his) region, enterokinase recognition site, expressed epitope driven by the T7 promoter. Homologous recombination of this vector and a mutagenic amplified were analysed by restrictions digest (*Bam*HI, *Hin*dIII). A transformed E.coli bacteria BL 21(DE3)*pLys*S colony was used to inoculate 100 ml LB-medium (+ 50 μg/ml Ampicillin, 25 μg/ml CAM) and allowed to grow to an OD= 0.7. Synthesis of the protein was induced using 0.3 mM IPTG. The bacteria suspension was grown for 3h and then centrifuged using the GSA-rotor for 10 min, 10.000 x g at 4°C. 20 μl aliquots of the bacteria cell lysate were analysed on a SDS-gel.

2.2.5.3 Separation of the protein on a SDS-polyacrylamide gel electrophoresis.

Homogenous polyacrylamide gel (12.5%) containing 0.15% SDS were prepared according to Laemmli (1970) between 14 X 14 cm glass plates with 2 mm thick spacers. This forms a discontinuous system consisting of two contiguous, but distinct gels: a resolving or separating (lower) gel and a stacking (upper) gel. Casting of gel was as follows: The glass plates, spacers, combs and the upper buffer reservoir were thoroughly cleaned. The resolving gel was poured between the glass plates and layered with water. After 1h polymerization the water was removed. The stacking gel was then poured into the glass plates. A comb was put between the plates and the gel was allowed to polymerize. After which the wells were loaded with defined concentration of protein standards and the gels electrophoresed under standard conditions: 1x SDS-running buffer, 4h run at 25 mA or o/n at 70 V. 1% coomassie blue R-250 in 40% methanol 0.7% acetic 53% water staining was used to visualise proteins on the gels. Permanent records of stained gels were obtained by drying them o/n between cellophane foils.

The following cocktail components were used to prepare the gels:

12% resolving gel

Water	3.35 ml
1.5 M Tris-Cl, pH 8.8	2.5 ml
10% SDS	0.1 ml
Acrylamide/bis (30% T, 2.7% C)	4.0 ml
10% ammonium persulfate	50 µl
TEMED	5 µl
Total	10 ml

4% Stacking gel

0.5 M Tris-Cl, pH 6.8	2.5 ml
Acrylamide stock solution	1.3 ml
10% SDS	0.1 ml
Water	6.1 ml
Total	10 ml

2.2.5.4 Isolation and purification of the protein.

In order to isolate the expressed protein, 50 µl start buffer was added to the pellet and resuspended on ice. The pellet was then transferred into a fresh tube and homogenised using the Ultraturrax for 5 min, followed by 10 min centrifugation at 12.000 rpm. The supernatant was filtered using a 0.45 µm sieve and stored at –20°C for further analysis. 250 µl of the supernatant was thawed and then mixed with 125 µl 4 M Tris, 125 µl 4% SDS and 40 µl protein loading buffer. The mixture was denatured for 3 min at 95°C before separation on a SDS-PAGE. The purification of the protein was carried out using a nickel column based on the ion exchange chromatography. The column was washed with 5 ml distilled water and stabilized using 0.5 ml 0.1 M nickel-Hcl solution. The column was equilibrated using 10 ml Start buffer. The supernatant was then loaded in the column and the flow-through collected; 1 ml aliquots in microcentrifuge tubes collected. The column was equilibrated again and the elution of the protein with 4 ml elution buffer with increasing concentration of Imidazol was carried out. 1 ml aliquots were collected in microcentrifuge tubes and stored at –20°C. Finally the samples were analysed on a 12% SDS-gel.

25 ml Start buffer cocktail mix contained:

Phosphate buffer pH 7.4	3 ml

2 M Imidazol	0.12 ml
H_20 bidest.	21.88 ml

3 Results

3.1 Cytogenetic (physical) mapping

Pigs (*Sus scrofa)* have a diploid chromosome number of 38 (18 pair of autosomes and XY sex chromosome). The standard G-banding and R-banding karyotypes have been reported in dividing the chromosomes into cytogenetically defined regions by systematically described landmarks (Gustavsson, 1988). The standard R-banded karyotype has been used for mapping SPRMTK gene (e.g. FISH mapping).
Two standard methods were employed in mapping SPRMTK, namely: a) PCR screening of somatic cell hybrid panel and radiation hybrid panel, and b) Fluorescence *in situ* hybridization.
The FISH mapping experiment presented here was carried out by Dr. med Felix Habermann (Department of Animal Breeding and Molecular Genetics, Technical University of Munich, Munich, Germany). The cytogenetic mapping of the SPRMTK gene was carried out by screening the somatic hybrid panel and the radiation hybrid panel using gene specific oligonucleotides.

3.1.1 PCR screening of hybrid panels

A *porcine* rodent somatic cell hybrid panel (Yerle *et al.*, 1996) and a *porcine* whole genome radiation hybrid panel (Yerle *et al.*, 1998) were screened for *porcine* SPRMTK by PCR. Primers $HUMSPRMTK_1$ forward 5´- GCG AGA TGA CAT GAC TGT CT -3´ and $HUMSPRMTK_2$ reverse 5´- CTG TAC CCA TGG CAC GCA CA -3´ were designed from *Homo sapiens* SPRMTK cDNA sequence (Accession number: L08961). PCR amplification (see section 2.2.3.5) was carried out on a sample of pig

genomic DNA. The PCR product was about 254 bp in length and proved to be thoroughly consistent with the expected sequence. Subsequently, PCR amplification was performed on the somatic cell hybrid panel (27 cell lines) and radiation hybrid panel (118 cell lines) in a total volume of 25 µl with 25 ng of panel DNA as template. The products were analysed by gel electrophoreses and data were collected, according to the principle, as follows: '1' represented having PCR product, '0' represents having no product (see tables 2 and 3).

3.1.1.1 Somatic cell hybrid panel

Somatic cell hybrid panel analysis of the 27 pig x rodent hybrid cell clones gave the following code: 10100 11111 10000 11100 00100 00 (Table 2 and Figure 1). Statistical evaluations of the results revealed a significant correlation of 1.00 between SPRMTK and SSC3 (error risk lower than 0.5% and maximum correlation of 78%). Within SSC3, chromosome region q11 indicated the highest probability of 0.84 with a correlation of 77%. This result is expected since the regions within the SSC3 (see figure 2, regions A-I) are differently distributed within the hybrid lines.

Table 2 The distribution of PCR results in somatic cell hybrid panel; cell line versus signal. '1' and '0' indicate the positive and negative PCR products.

Cell line	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Signal	1	0	1	0	0	1	1	1	1	1	1	0	0	0
Cell line	15	16	17	18	19	20	21	22	23	24	25	26	27	
Signal	0	1	1	1	0	0	0	0	1	0	0	0	0	

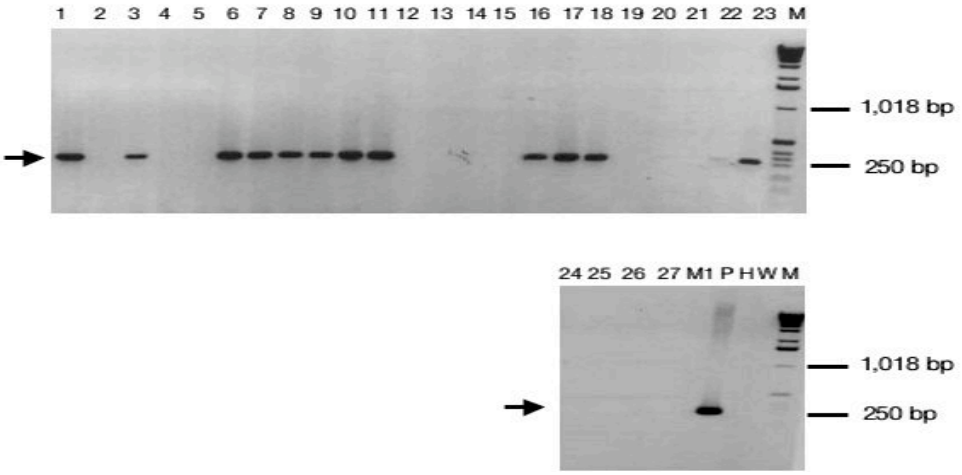

Figure 1 PCR amplification of SPRMTK product using gene specific primers on the 27-clone somatic cell hybrid panel. In this case a rodent background did not amplify. Pig (P) as positive control; hamster (H), mouse (M1) DNA and water (W) were used as negative controls. M: 1 Kb molecular weight marker. Arrows indicate the ampified SPRMTK product.

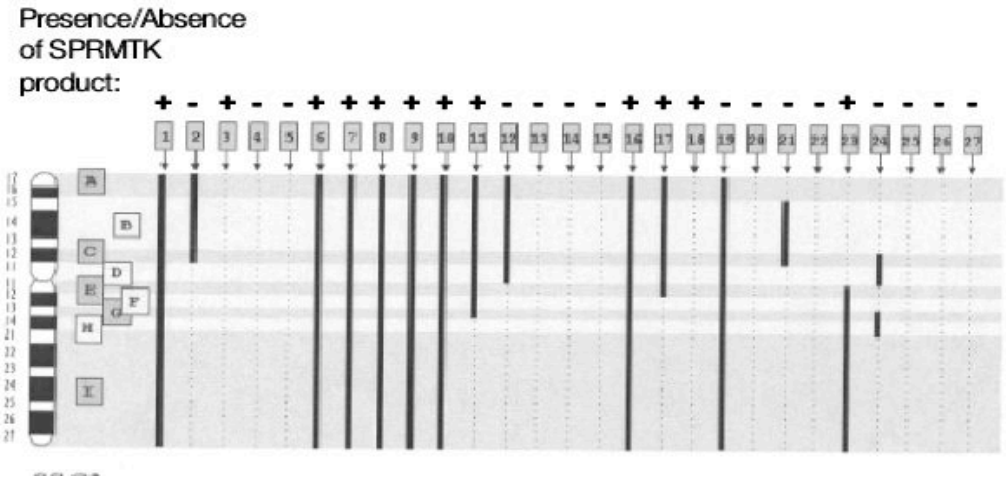

Figure 2 Diagram representing the proportion of pig chromosome 3 retained in each of the 27-clone. A, B, C, D, E, F, G, H, I indicate regions of SSC3 distinguishable by the panel. Clones positive for the *porcine* product are indicated by a plus (+). The presence of products for clones 3, 18 may be an artefact or may indicate that the descriptions of these clones are incorrect. However note that amplification in clones 1, 6, 7, 8, 9, 10, 11, 16, 17, 23 is consistent with the site being located in region E i.e. SSC3q11(distal)12.

3.1.1.2 Radiation hybrid panel

To determine the gene order and to obtain a finer resolution within the pig genome, a high-resolution mapping on a *porcine* whole-genome radiation hybrid panel (WG-RH) was used.

Radiation hybrid panel analysis resulted in the following vector (see table 3): 00000 10001 00100 00110 00101 10000 00101 10100 01000 00001 01010 00001 01000 00001 00010 00000 00001 00001 01110 01100 01101 10110 01011 010. The most significantly linked marker (two-point-analysis) is SWR978 (Rohrer *et al.*, 1994) on SSC3 (39cR; LOD = 11.08). Multi-point-analysis leads to linkage group SWR2069-SW2597-SWR978-**SPRMTK**-IL1B-SW1045-SW1436. Two-point LOD scores indicate that the gene is linked to the markers SWR978 and IL1B. Moreover, the cytogenetic localization of SW2597 (Alexander *et al.*, 1996) and IL1B (Mellink *et al.*, 1994) are 3q11-q12 and 3q11-q14 respectively, thus confirming the localization of SPRMTK to 3q11-q12. These data are the first to establish a chromosomal location for the SPRMTK locus and thus provide a framework map to further localize the gene.

The reliability of these results by statistical evaluation showed a low error risk. However the fluorescence in situ hybridization technique was performed to confirm these results.

Table 3 PCR Results of the Radiation hybrid panels. The Radiation hybrid panel: the letter L represents the hybrid cell line mice x pig, the letter F represents the hybrid cell line hamster x pig.

Cell line	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11	L12	L13	L14
Signal	0	0	0	0	0	1	0	0	0	1	0	0	1	0
Cell line	L15	L16	L17	L18	L19	L20	L21	L22	L23	L24	L25	L26	L27	L28
Signal	0	0	0	0	1	1	0	0	0	1	0	1	1	0
Cell line	L29	L30	L31	L32	L33	L34	L35	L36	L37	L38	L39	L40	L41	L42
Signal	0	0	0	0	0	1	0	1	1	0	1	0	0	0
Cell line	L43	L44	L45	L46	L47	L48	L49	L50	L51	L52	L53	L54	L55	L56
Signal	1	0	0	0	0	0	0	0	1	0	1	0	1	0
Cell line	L57	L58	L59	L60	L61	L62	L63	L64	L65	L66	L67	L68	L69	L70
Signal	0	0	0	0	1	0	1	0	0	0	0	0	0	0
Cell line	L71	L72	L73	L74	L75	L76	L77	L78	L79	L80	L81	L82	L83	L84
Signal	1	0	0	0	1	0	0	0	0	0	0	0	0	0
Cell line	L85	L86	L87	L88	L89	L90	L91	L92	L93	L94	L95	L96	L97	F98
Signal	0	1	0	0	0	0	1	0	1	1	1	0	0	1
Cell line	F99	F100	F101	F102	F103	F104	F105	F106	F107	F108	F109	F110	F111	F112
Signal	1	0	0	0	1	1	0	1	1	0	1	1	0	0
Cell line	F113	F114	F115	F116	F117	F118								
Signal	0	1	0	1	1	0								

3.1.1.3 Fluorescence *in situ* hybridization (FISH)

The precise chromosomal localization of the gene was assigned by the fluorescence *in situ* hybridization method, which complements the genetic mapping technique. The key advantage of FISH is its ability to target only those genetic sequences of interest.

The reference probe used for hybridization was the PAC-clone E11-410 (see section 3.5.1). The figure 3 below shows a clear fluorescent signal in the metaphase preparation of *porcine* chromosome and the *porcine* gene is a single copy gene located between band 11 and 12 of the long arm of chromosome 3 (3q11-12).

Evaluation as follows:

The most precise location: SSC3q11(distal)q12

Flqter: 0.56 ± 0.03

Chromosomes measured: 21

Mean chromosome length: 4.78; Range: 3.55–7.64 □m

Standard deviation: 1.02 □m

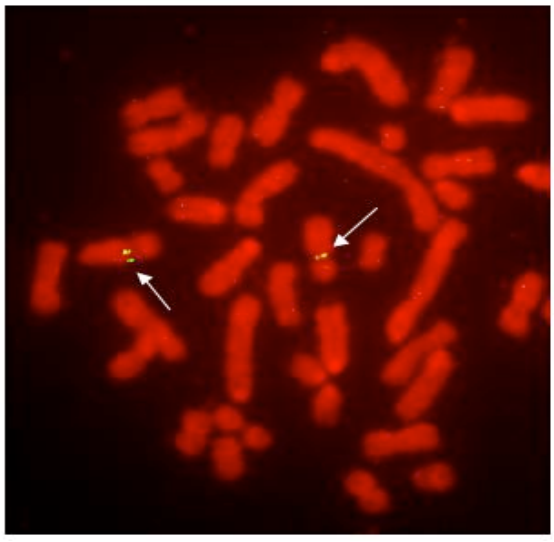

Figure 3 Fluorescence *in situ* hybridization of SPRMTK gene in *porcine*. The yellow/green fluorochrome spots indicate the position of the gene at the pericentromeric region of Chromosome SSC3q11-12.

3.2 Gene expression studies

This study was intended to verify in which *porcine* tissue/tissues the SPRMTK gene is expressed. The basic principle here is, genes, which code for proteins are transcribed into messenger RNA's (mRNA's) in the cell nucleus. Such information is crucial to our understanding of the gene networks that control developmental, physiological and pathological processes. Two classical applications were used to detect in which tissue SPRMTK-mRNA is expressed:

1) Reverse transcription-polymerase chain reaction
2) Northern blotting

3.2.1 Isolation of RNA from the *porcine* tissues

The five target tissues were isolated by dissection and then freezed in liquid nitrogen. In order to isolate the RNA the tissues were transferred to a lyses puffer containing RNase inhibitors, and then subjected to mechanical grinding (see section 2.2.4.1). Speed in isolation of RNA is critical for the study of gene expression. With proper pressure, most tissues cells were shattered into invisibility and released most of their RNA. Also, the liquid nitrogen followed by thawing releases most of the RNA from the cells. To eliminate contaminating DNA, the RNA extracts were treated with 1U DNase 1 for 15 minutes at room temperature. The DNase 1 was inactivated by adding 1 µg of 20 mmol/l EDTA solution to the reaction mixture and heating at 65°C for 10 min. 5 µl of the isolated RNA were analysed on a 1.2% denaturising gel to verify the integrity of the RNA (see figure 4).

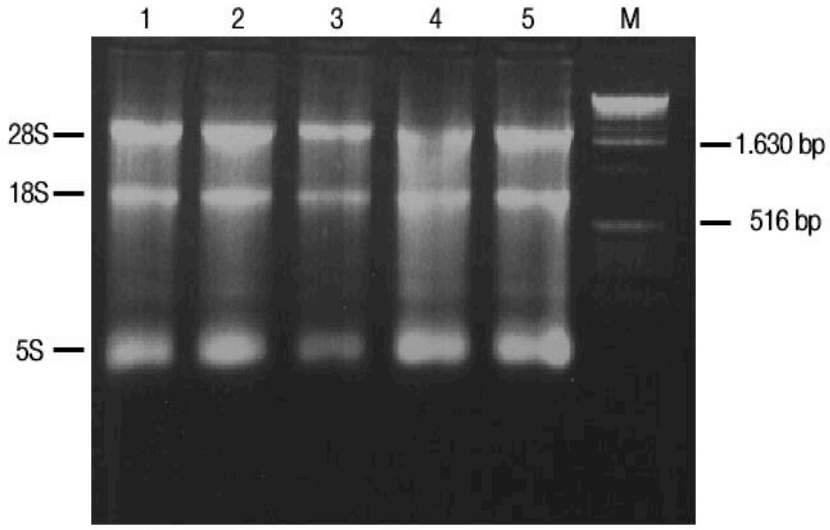

Figure 4 Shows intact and undegraded total RNA with virtually no genomic DNA. Sizes of RNA ribosomal bands indicated on the left. Lane 1-5: brain, testis, kidney, spleen, liver, respectively, M: 1 Kb molecular weight marker

3.2.2 Reverse transcription-polymerase chain reaction

The RT-PCR assay is the most sensitive technique, allowing the detection of even a few copies of an mRNA. This amplification technique, RT-PCR for short, is a combination of two interrelated steps: synthesis of cDNA from RNA by reverse transcription (RT) and amplification of a specific cDNA by polymerase chain reaction (PCR). The RT-PCR assay was performed on five *porcine* tissues: brain, testis, kidney, spleen and liver.

RT-PCR was carried out using the isolated RNA as template, with OligodT as anchor primer in the reverse transcription step (Section 2.2.4.3). The subsequent PCR using two gene-specific oligonucleotides SP1f 5′-GCG AGA TGA CAT GAC TGT CT-3′; SP1r 5′-CTG TAC CCA TGG CAC GCA CA-3′ (see Table 4), designed based on the published *Homo sapiens* HUMSPRMTK-mRNA sequence (Accession number: L08961), gave the expected

fragment length of about 150 bp (see figure 5) with the testis RNA. The sequence of the purified fragment is in conformity with the published sequence.

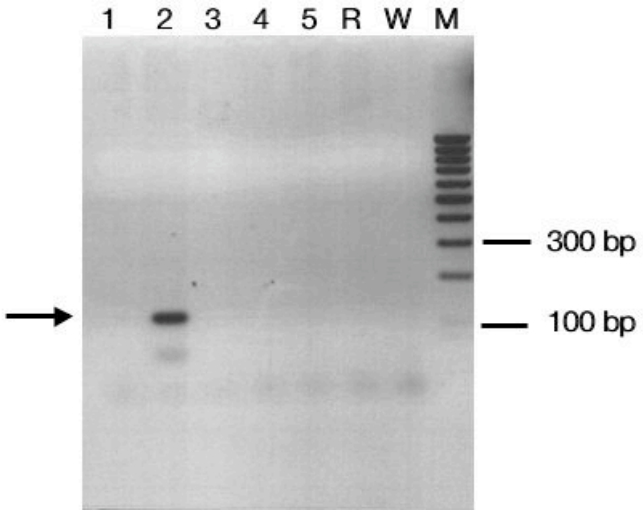

Figure 5 RT-PCR: the arrow indicates amplified fragment of testis RNA. Lane 1-5: brain, testis, kidney, spleen, liver respectively. R: no (RT) W: PCR negative control (no cDNA). M: 1 Kb molecular weight maker

Additional PCR using gene specific primers:
SP2f 5′-TGC TCA TCA TCX TXG GCT GC-3′; SP3f 5′-GAA ATC CTT CTG TCG GCG; SP4r 5′-ATG GTT ATG AAC TCC GGG-3′ (see Table 4) to quantify further the testis mRNA was carried out with the standard amount (25 ng) which differed by factor 10 for every primer pair see figure 6. The amplified fragments were consistent with the expected sequence. The negative control, where no RT was used, confirms no genomic-DNA contamination (see figure 6).

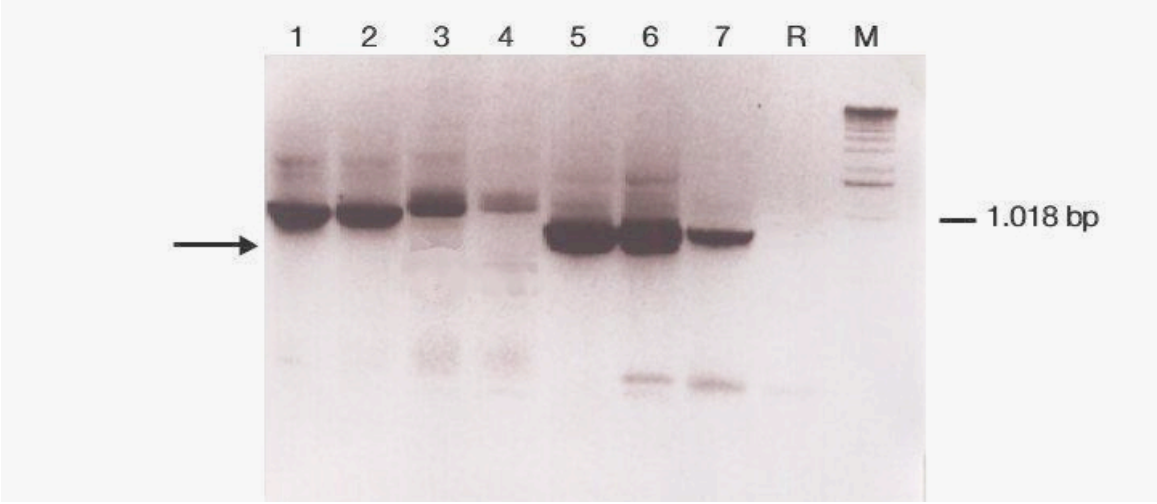

Figure 6 RT-PCR analysis of the RNA sample isolated from *porcine* testis. Concentration gradient of template (cDNA) used 25 ng, 15 ng, 5 ng and 2 ng for each primer-set. Lanes: 1-4 using primer-set SP2f and SP4r, and Lanes 5-7; semi-nested PCR using primer-set SP3f and SP4r; the arrow indicates amplified SPRMTK product. R: no (RT) negative control. M: 1 Kb molecular weight marker.

PCR product of lanes 5-7 (see figure 6) where excised and purified using QIAEX gel purification kit. 2 μl was subcloned into pGEM-T vector. Sequencing of the clone revealed a high degree of nucleotide sequence homology with mouse and human SPRMTK sequences 88.5% and 74%. That between mouse and human SPRMTK share a homology of 73.5%. Comparison of the PCR product amino acid sequence (which span the entire catalytic domain) and the published *Homo sapiens* sequence is shown in figure 7.

Table 4 PCR primer-set used in the RT-PCR; size of product and their respective annealing temperatures Tm

Oligonucleotide 5′-3′	Size of product (bp)	Tm°C
1. SP1f 5′-GCG AGA TGA CAT GAC TGT CT-3′		
2. SP1r 5′-CTG TAC CCA TGG CAC GCA CA-3′	150 bp	57
3. SP2f 5′-TGC TCA TCA TCX TXG GCT GC -3′		
4. SP4r 5′-ATG GTT ATG AAC TCC GGG -3′	990 bp	56
5. SP3f 5′-GAA ATC CTT CTG TCG GCG -3′		
6. SP4r 5′-ATG GTT ATG AAC TCC GGG –3′	800 bp	56

```
Human   209   KSFCRRAIELTHSLGVSEELQNKLEDVVIDRNLLILGKILGEGEKGTVYEGLWNIPEGK
pig     12    KSFCRRAIELTLQSLGVSEELQNKLEDVVIDRNLLILGKILGEGEFGSVMEGNLNQQDGT

Human   268   EVKIPVAIKTLKLDTMANKEI---LDEASVMKGFGNPHVVRLLGICMTSTIYVITE-YCL
pig     72    SQK--VAVKTMKLDNFSHREIEEFLSEAACMKDFNHPNVIRLLGVCIEMSPQGIPKPMVI

Human   324   LVYRRNKDKAEQHRSNCAELNP---PLQTLLKFMVDIALGMEYLSNRNFLHRDLAARNCM
pig     130   LPFMYGDLHTYLLYSRLDTGPKHIPLQTLLKFMVDIAQGMEYLSNRNFLHRDLAARNCM

Human   381   LRDDMTVCVADFGLSKKIYSGDYYRQGRIAKMPVKWIAIESLADRVYT-KSDVWAFGVTM
pig     190   LRDDTVCVADFGLSKKIYSGDYYRQGRIAKMPVKWIAIESLADRVYTSKSDVWAFGVTM

Human   440   WEIATTLRGMTPYP
pig     250   WEIA--RGMTPIP
```

Figure 7 Alignment of the deduced amino acid (aa) sequences between the human- and pig testis cDNA share a high homology of 74%. The mouse cDNA not included because it has only been partially sequenced. Y indicates potential autophosphorylation sites. Dashes means no similarity.

3.2.3 Northern Blotting

A Northern Blot containing RNA prepared from the *porcine* tissues - brain, testis, kidney, spleen, liver - was probed with ECL-labelled cDNA (see section 2.2.3.10.2). The cDNA region used represents the catalytic domain with a homology of 74% to human SPRMTK-mRNA. 50 µg of total RNA was resolved on a 1.2% agarose gel containing 2% formaldehyde. A Northern blotting prepared with β actin cDNA probe was done as a control. The hybridization gave no signals.

3.3 Expression of cDNA fragment encoding the intracellular domain.

The SPRMTK protein shows characteristics of a protein tyrosine kinase (PTK) and inhibition of the activity of PTK prevents acrosomal exocytosis and blocks fertilization (Burks *et al.*, 1995). Thus the intracellular domain ranging from amino acids Leu 240 to Ala 515, with potential phosphorylation sites plays a key role in fertilization.

3.3.1 Choice of expression vector

In order to obtain a high-level expression the pRSET vector was employed, which is controlled by the presence of a strong bacteriophage T7 promoter. The pRSET vector is offered in three forms A, B, C. Each having the N-terminal tag coding sequence in a different reading frame relative to the multiple cloning site. In this expression, the pRSET A vector was used which allows a simplified in-frame cloning of the cDNA fragment.

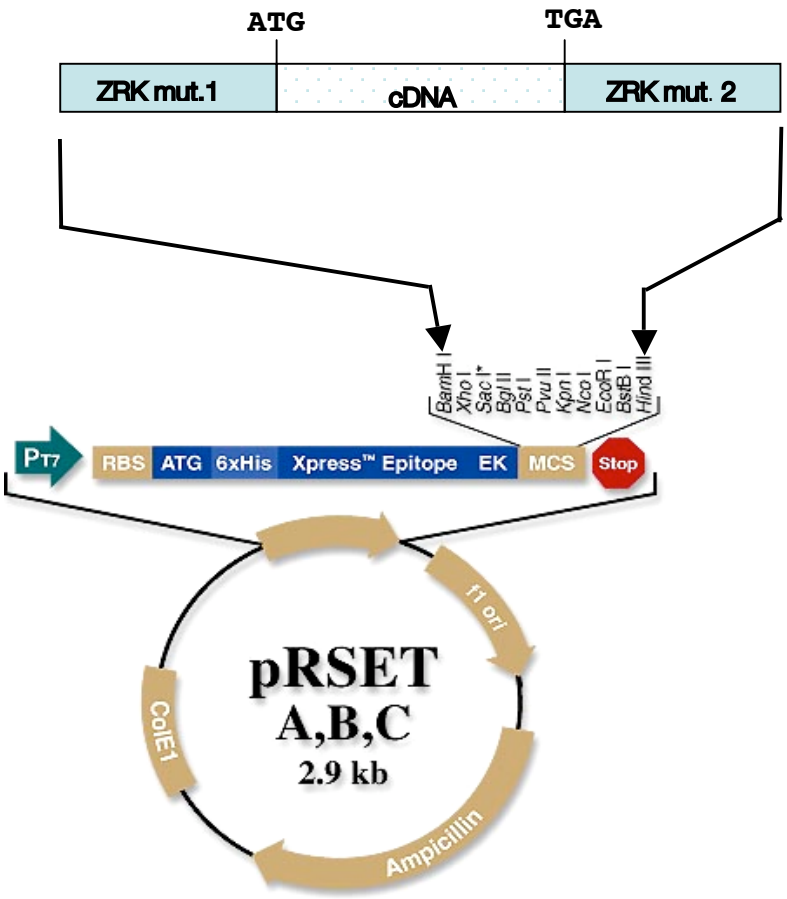

Figure 8 Schematic structure of pRSET expression vector (adapted from Invitrogen), showing the multiple cloning site. The cDNA fragment was ligated into the *Bam*HI and *Hind*III restriction sites as indicated by arrow. The sites *Bam*HI (**GGATCC**) and *Hind*III (**AAGCTT**) were introduced into the cDNA fragment by primers: ZRK-mut.1 for. 5'-CCA GT**G GAT CC**T TCT GTC GGC GAG CCA-3'; ZRK-mut.2 rev. 5'-C GCT GTA **AAG CTT** CTT AGA GAG GCC-3' respectively to facilitate cDNA fragment subcloning into the pRSET A vector.

3.3.2 Choice of expression system

A prokaryotic expression system was employed for expression. Several parameters where considered that may affect the expression of the protein in this system-BL21(DE3)pLysS, emphasizing on achieving a high level of expression and

rapid cell growth. The limitation of this *E. coli* strain, however, is that it cannot be used efficiently as a primary transformation hosts. The construct pRSET A vector- cDNA fragment therefore, was first transformed into another strain, the XL1-Blue *E. coli* bacteria. The clones were selected after *Bam*H1 and *Hin*dIII digestion (see figure 9) of each plasmid DNA obtained by mini-lysate preparation (see section 2.2.2.3). A single positive clone was chosen for expression – clone #2 see figure 9 – and was partially sequenced at both ends to confirm the presence of the insert and whether cloned in the correct frame.

3.3.3 Generation of plasmid constructs that expressed the recombinant fusion protein

The pRSET A vector was double digested using *Hin*dIII and *Bam*H1 restriction enzymes. Mutations were introduced into the fragment as explained in section 2.2.5.1.
The mutagenic PCR product was purified by agarose gel electrophoresis digested with *Hin*dIII and *Bam*H1, and purified by QIEAX purification kit. The resulting fragment was ligated into pRSET A vector and the ligation mixture 5 µl was transformed into 50 µl *E coli* BL21(DE3) bacteria for expression.

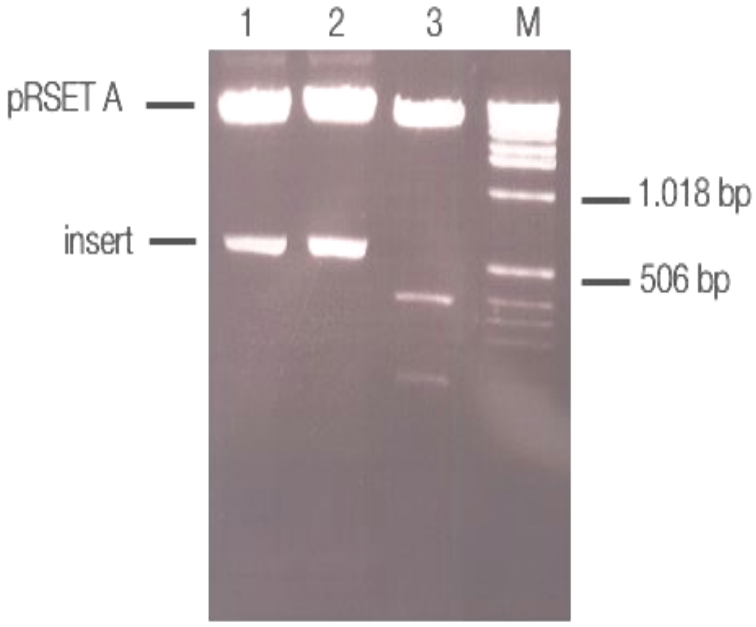

Figure 9 Gel electrophoresis of restriction digest of plasmid and cDNA insert run on a 1% agarose gel at 100 V for 40 minutes. Lane 1-2: pRSET A vector and insert (SPRMTK fragment). Lane 3: wrong insert size. M: 1 Kb molecular weight marker. The clone analysed in lane# 2 was used in expression.

3.3.4 Expression and analysis of the protein on a 12% polyacrylamide gel electrophoresis.

After identification of a positive clone, the isolated plasmid was retransformed into the BL21(DE3)pLysS. The bacteria now containing the pRSET A -derived expression vectors was grown overnight at 30°C in LB medium containing ampicillin and chloramphenicol. Subsequently, protein expression was induced by the addition of IPTG (see section 2.2.5.2) to a final concentration of 0.1 mM. After incubation at 30°C for 3h, the bacteria was harvested and 20 µl of the cell lysate was analysed on a 12% SDS-PAGE. The protein of expected size was expressed as indicated below (Figure 10). The purification of the 6H-tagged proteins from

the bacterial pellet was carried out see section 2.2.5.4. The purity of the recombinant proteins was checked by SDS PAGE; estimated to have a purity of 95%.

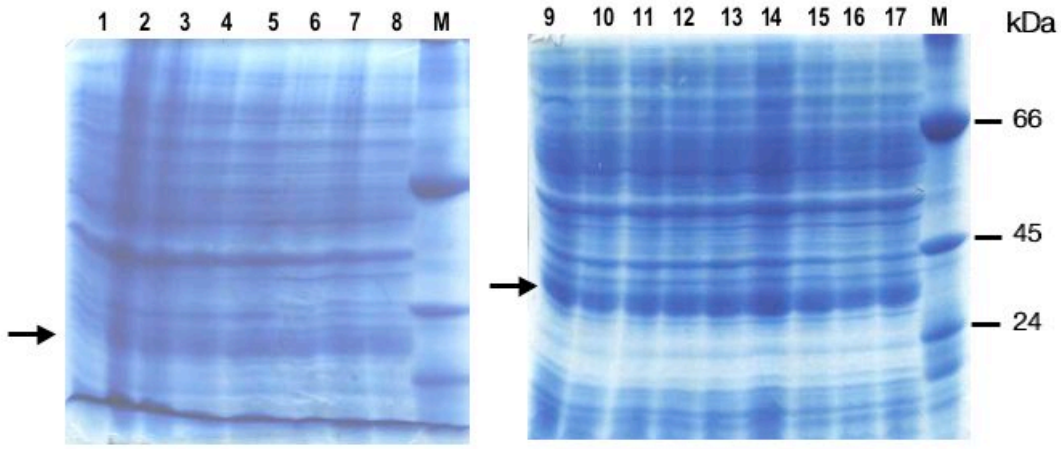

Figure 10 SDS/PAGE (12%) of the recombinant fusion protein produced in BL21(DE3)pLysS *E. coli* after IPTG induction. 20 µl of the soluble fraction of each cell lysate was analysed on a coomassie stained SDS-PAGE gel. Lane 1: total protein extract from non-induced cells. Lane 2-6: total protein extract from induced culture after 1h. Lane 7-11: total protein extract from induced culture after 2h. Lane 12-17: total protein extract from induced culture after 3h. Molecular size markers in kilo Daltons are indicated on the right. The arrows indicate the expressed recombinant fusion protein 27 kDa.

3.4 Catalytic Domain Phylogeny

Amino acid sequence alignments can be used to deduce phylogenetic relationships (Doolittle, 1979). We used the alignment data (see appendix 7.2) of fourteen-protein-kinase (see Tables 6 and 7) catalytic domains including the amplified catalytic domain of the *Sus scrofa* SPRMTK to construct a phylogenetic tree (see Figure 11). The proteins

were derived from both vertebrate and invertebrate sources and, in some cases, presumed functional homologs from both vertebrate and invertebrate sources are represented. The tree, therefore, reflects catalytic domain evolution stemming from gene duplication events, speciation events or both.

All protein kinases thus far characterized with regard to substrate specificity fall within one of two broad classes, serine/threonine-specific and tyrosine-specific. This phylogenic data here, adds to our knowledge, that the *Sus scrofa* SPRMTK shows the characteristic of the later, because in molecular sequence comparison high similarity usually implies significant functional or structural similarity.

The amino acid sequence comparison of the *porcine* SPRMTK versus *Homo sapiens* SPRMTK reveals the presence of the key residues; **DLAARN, RDL and DFG** (Hanks, 1987; Hunter, 1987) which are highly conserved through out the tyrosine-specific protein family.

Introduction of the software package used in construction of the phylogenetic tree

Clustal W: Clustal W is a general purpose multiple alignment program for DNA or proteins. In this case, all pairs of sequences were aligned separately in order to calculate a distance matrix giving the divergence of each pair of sequences. A guide tree was then calculated from the distance matrix and the sequences progressively aligned to the branching order in the guide tree. The bottom line is a series of pairwise alignments (explained in Phylip-format dendrogram) is used to align the rest of the amino acid sequence, following the branching order in the guide tree.

Phylip-format dendrogram: Briefly, similarities scores were obtained for all possible pairwise comparisons and transformed into a different matrix from which branch order and length were determined. The individual abbreviations are indicated in a separate list; see tables 6 and 7 for abbreviation. The position of the protein-tyrosine kinase cluster was determined by including a protein-tyrosine kinase (RET) together with that of the *porcine* and *Homo sapiens* SPRMTK. The other three clusters represent the protein –serine/threonine kinase.

Table 5 Determination of the branch order and pairwise distance matrix applying the Phylip-format dendrogram:

The Kinase proteins plotted into four different clusters:

Pairwise distance matrix	**Branch order**
1). KIN2:0.39753,	
CDK4:0.37475)	:0.01921
2). PKA2:0.00000,	
TPK2P:0.00000)	:0.32400
PCK-53E:0.35232)	:0.00735
PKC-1:0.35594)	:0.03535
PSK-H1:0.34590):0.03276)	:0.01053
PKU:0.42059)	:0.00564
3). STE7:0.39635,	
Protein-Serine/threonine-Ki:0.37262)	:0.02430
Gag-raf:0.38278)	:0.01089
4). RET:0.33830,	
SSPRMTK:0.33329)	:0.07328
HUMSPRMTK:0.55594)	:0.01023

The protein kinases as listed under serine/threonine and tyrosine-kinase in tables 6 and 7 respectively, were selected for alignment. All protein kinases included have either been characterized as distinct by complete or partial protein purification, or have been identified as unique based on nucleotide sequencing. The *porcine* SPRMTK not listed.

Table 6 Protein-serine/threonine kinase family members used in the construction of the phylogenetic tree.

Name of Protein	**reference**
PKC-1: Protein kinase C-like	Palmer *et al.*, 1994
PSK-HI: Protein serine kinase	Brede *et al.*, 2002
PCK-53E: Protein C Kinase 53E	Adams *et al.*, 2000
TpK2P: Tyrosine protein Kinase 2P	Bussey *et al.*, 1997
PKA2: cAMP-dependent protein kinase type 2	Toda *et al.*, 1987
CDK4: Cyclin-dependent kinase 4	Strausberg *et al.*, 2002
PKU: Protein kinase tousled-like 1	Yamakawa *et al.*, 1997
STE7: Serine/threonine protein kinase	Teague *et al.*, 1986
Gag-raf: gag-raf polyprotein	Mark *et al.*, 1984
KIN2: protein kinase KIN2	Levin *et al.*, 1987
Protein serine-threonine kinase	Lohia *et al*, 1994

Table 7 Protein–tyrosine Kinase family members used in the construction of the phylogenetic tree.

Name of Protein	**reference**
HUMSPRMTK: Homo sapiens tyrosine kinase	Burks *et al.*, 1995
RET:Protooncogene tyrosine-protein kinase receptor	Moreau *et al.*, 1998

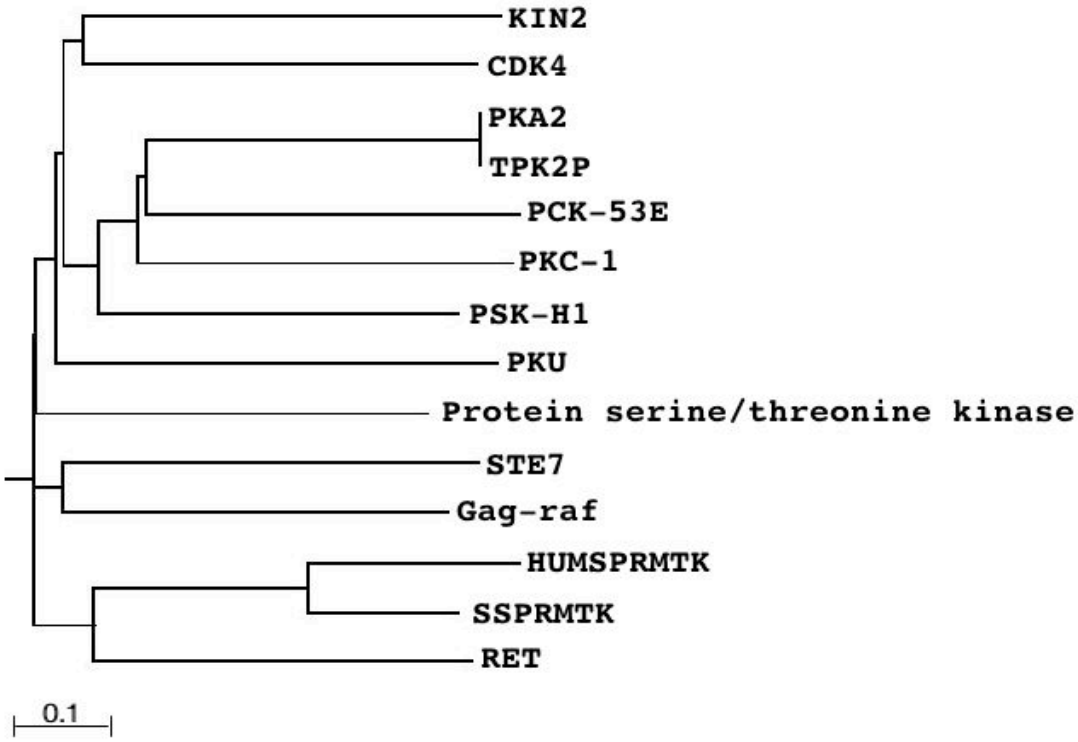

Figure 11 Deduced phylogeny of the protein kinase catalytic domains. The phylogenetic tree was constructed from the multiple alignment of the amino acid sequence as shown in appendix. The tree-building concept of Fitch and Margoliash (1967) was used as implemented by Feng and Doolittle (1987). Scale bar 0,1 mm.

3.5 Genetic analysis of the *porcine* SPRMTK gene

3.5.1 Screening the PAC library using PCR

In order to analyse the SPRMTK gene the *porcine* genomic PAC-library (IVMp 714) (Al-Bayati *et al.*, 1999) was screened using oligonucleotides designed from the published HUMSPRMTK-mRNA sequence (Accession number: L08961). The primer pair: SP1f 5'-GCG AGA TGA CAT GAC TGT CT-3' and

SP1r 5'-CTG TAC CCA TGG CAC GCA CA-3' amplified a fragment of about 150 bp long. Initially, the super-super pool (SSP) was screened to identify potential positive pools. PCR products were obtained for SSP# 5,8,13 with SSP# 13 amplifying the brightest band as illustrated in figure 14. However, to rule out false positives all three SSP# 5,8,13 were selected for a second rounds of screening. The subsequent PCR amplified a positive signal at SP# 62 (SP of SSP# 13). Further screening identified a signal in plate pool# 51. To determine the correct clone, a two-dimensional PCR screening of the rows and columns was performed and a positive signal at position E11 was identified (see figure 12).

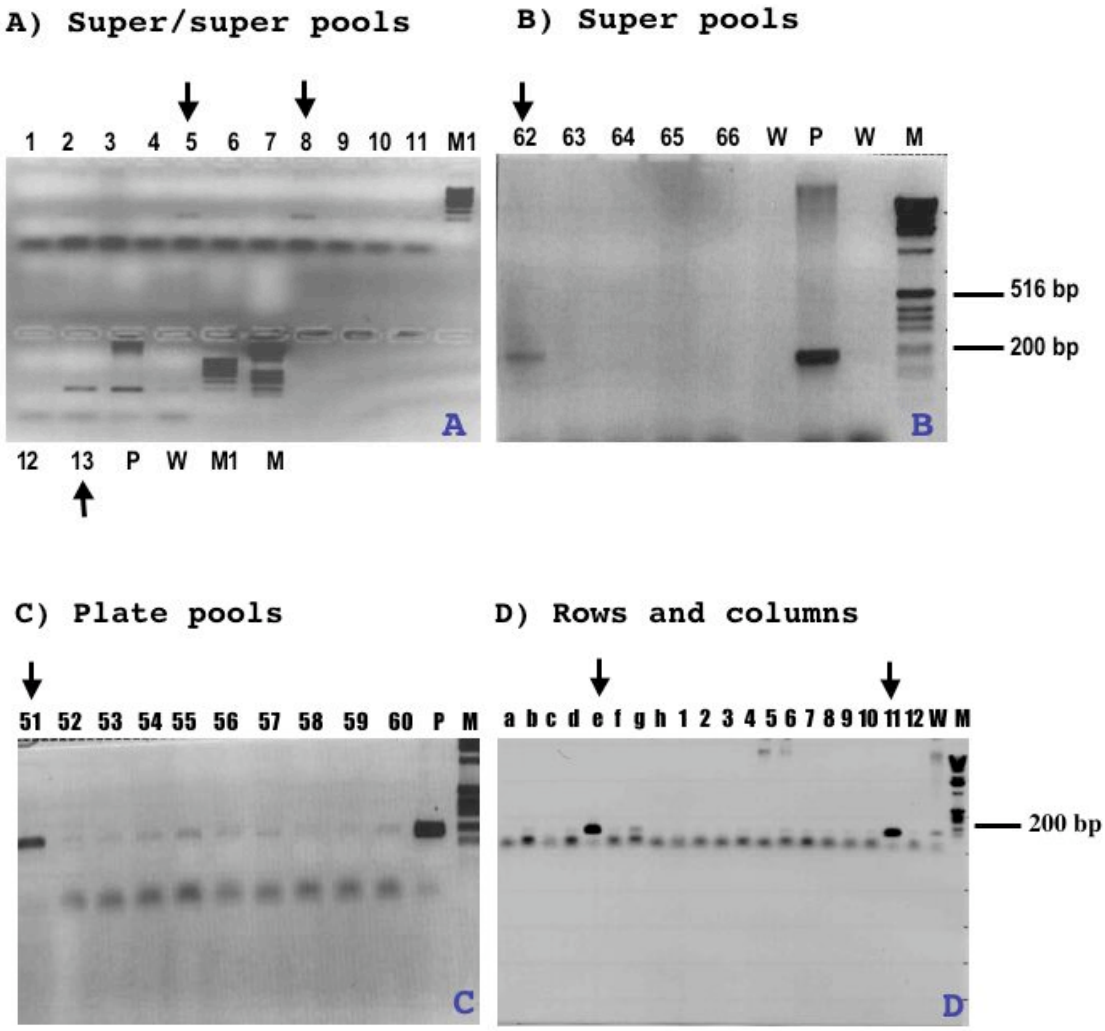

Figure 12 Screening of the *porcine* genomic PAC Library by PCR of hierarchical DNA pools; **A)** super/super pools, **B)** super pools, **C)** plate pools and **D)** subsequently the rows and columns, where the clone E11 was identified. Arrows indicate the positive pools and plate that were screened in successive rounds. The products were analysed on a 1.2% agarose gel. P: pig genomic DNA as positive control, W: H_20 as negative control, M & M1: 1000 bp & 100 bp molecular weight markers

Pulsed field gel electrophoresis of the PAC clone

Once the PAC clone was isolated, it was probed using PFGE technique. The direction of current flow in the electrophoresis chamber is periodically altered; this allows

fractionation of pieces of DNA ranging from 50.000 bp to 5 million bp, which is much larger than can be resolved on standard gels. The PAC Clone was digested using *Not*1; 2 µg DNA was digested using 10U enzyme for 1h incubation at 37°C (see section 2.2.3.3). The sample was loaded on the gel. The size of the PAC clone is approximately 150 kb as indicated in figure 13.

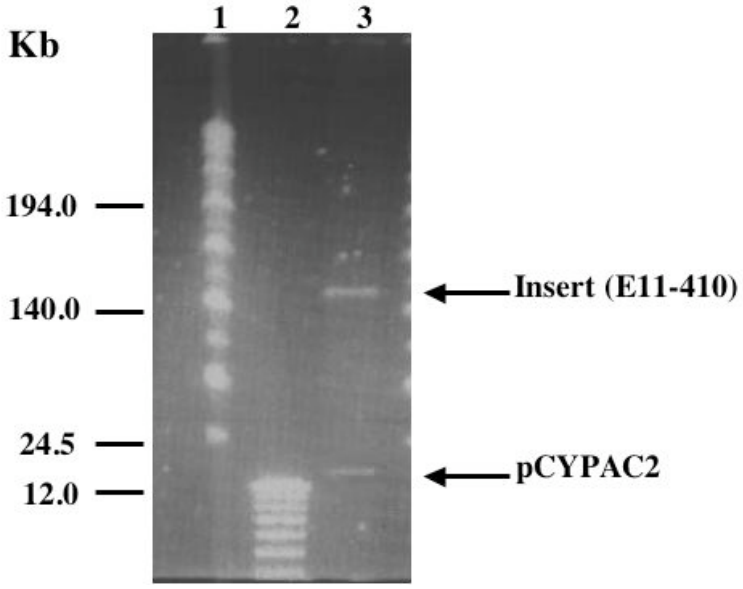

Figure 13 Autoradiogram of the PFGE gel probed with the PAC clone E11-410. Arrows indicate the faint bands of the vector pCYPAC2 and the insert (E11-410 PAC clone) of approx. 150 kb. Lane# 1 = mid range marker II; lane# 2 = 1 Kb maker and lane# 3 = vector and insert.

3.5.2 Southern Blot hybridization and subcloning of the PAC clone E11-410

The PAC clone E11-410 was used to inoculate 500 ml LB-medium supplemented with 25 µg/ml kanamycin. After 16h of growth shaking at 225 rpm at 37°C, plasmid DNA was isolated as explained in section 2.2.2.2. The positive PAC plasmid DNA was digested using the following enzymes; *Kpn*I, *Hin*dIII,

*Bam*H1, *Sph*1, *Pst*1, *Xba*1, *Eco*R1 and was size-separated on a 0.8% agarose gel (See figure 14). This was then blotted onto a nylon membrane using Southern Blotting technique. The hybridization probe was isolated by PCR and ECL labelled. Next, the labelled probe was added to the hybridization solution and membrane for hybridization at 42°C. After over-night hybridization, filters were washed at 54°C (high stringency) and then the hybridized fragments were detected using X-ray film. The corresponding fragments were excised for subcloning into the pGEM-4Z vector for sequence analysis.

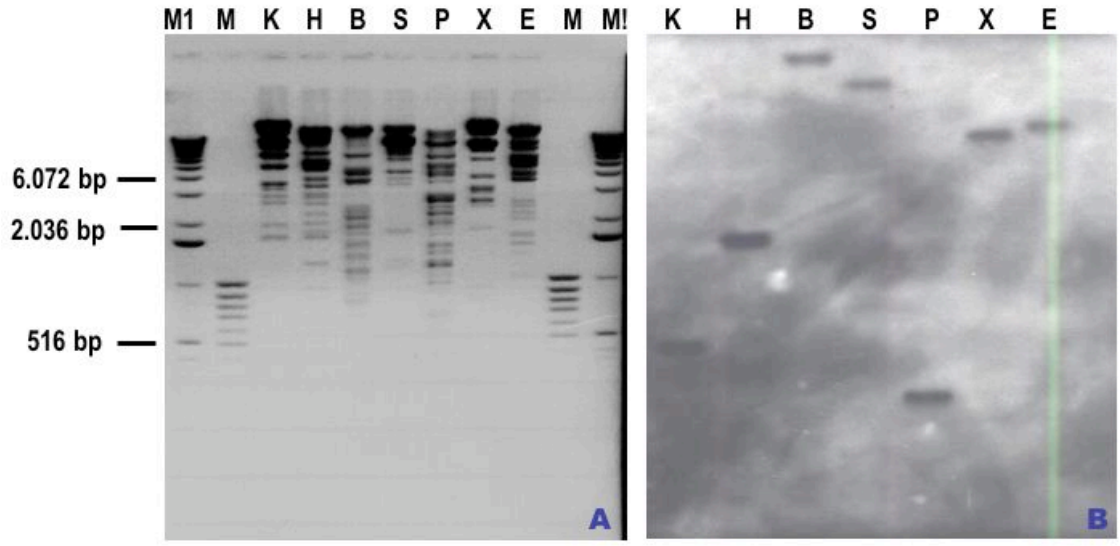

Figure 14 Isolation of the PAC DNA E11-410 fragments
A) 8 µg of the PAC DNA E11-410 was digested using the following restriction enzymes *Kpn*1 (K), *Hind*111 (H), *Bam*H1 (B), *Sph*1 (S), *Pst*1 (P), *Xba*1 (X), *Eco*R1 (E). Analysed on a 0.8% agarose gel.
B) Positive fragments, as indicated in autoradiogram B, were excised, subcloned in their respective pGEM-4Z vectors and sequenced. Exons 1 & 2 not sequenced.

Table 8 Primer pairs used to amplify the respective Exons.

Pimer pairs 5'-3'	Size (bp)	Position	Annealing temperature(°C)
1. cctattgccagttctagtccag 2. ccaagggtctacctggtccca	202	Exon 3	55
3. cccagtggactgactgggag 4. aactgaagtaaaccagcaag	98	Exon 4	55
5. aggaccactctaagtgccatgc 6. tggacttgaacagagatccg	131	Exon 5	52
7. ggcgcctggcaacgcagatc 8. attcctgtcaatcacaacat	756	Exon 6	54
9. cggggacctgcatacttacttac 10. caggtcctctcactaaccctgc	270	Exon 7	58

Table 9 Nucleotide sequences of Exon/Intron junctions of *porcine* SPRMTK gene

Exon (bp)	3' Splice donor[a]	Intron (bp)	5' Splice acceptor[a]
3 202	TTCCAG**gt**aagtccga...	1.552	...gccccatc**ag**TAGCAC
4 98	GAACTG**gt**gggctacc...	322	...atggcctg**ag**TGACAA
5 131	CTCTCT**gt**ctccagaa...	535	...tttgggta**ag**TCTCCC
6 756	TTTTCA**gt**acatagct...	1.876	...gtttgcca**ag**AAGTTT
7 270	CTATTGAagttcatgg....		

[a] Intron sequence is shown in lowercase and exon sequence in uppercase

3.6 SPRMTK cDNA synthesis

3.6.1 Rapid amplification of cDNA Ends

Homo sapiens SPRMTK-mRNA was used to design *porcine* specific primers to obtain he upstream and downstream regions of the gene using Rapid Amplification of cDNA Ends (RACE) strategy. Amplification of the 3′ region utilised the tail-introduced smart primer (3′-smart primer) together with a *porcine* specific primer (ZRK3-for primer see appendix section 7.3). All PCR reactions were carried out at 94°C for 60s, 56°C for 60s, 72°C for 90s. Amplification of the 5′ region was not achieved. These might possibly be due to a premature termination of the first-strand cDNA synthesis caused by pausing of RT generally.

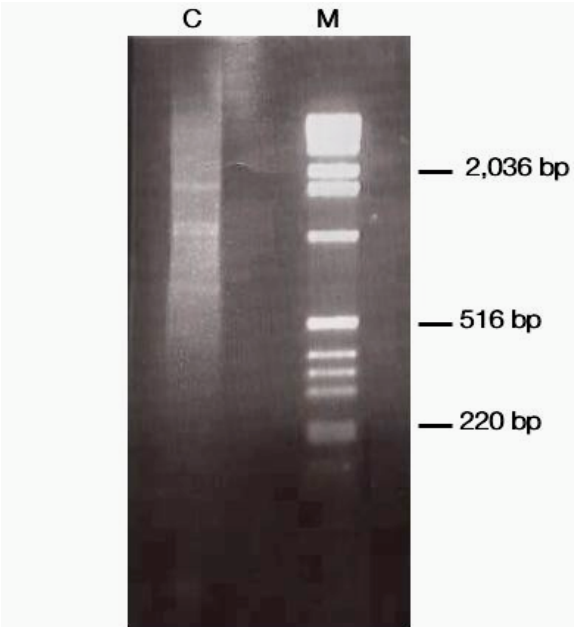

Figure 15 Analysis of SPRMTK ds cDNA synthesis
Lane 1: 5 µl ds CDNA loaded on gel showing a size distribution of about 0.5-10 kb on an agarose/EtBr gel. M: 1 kb molecular weight marker.

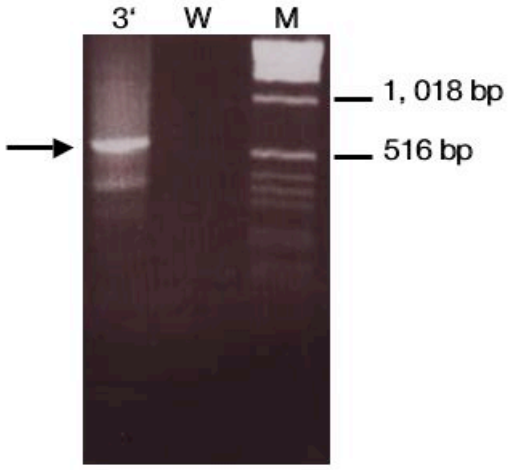

Figure 16 3'-RACE results analysed on a 1% gel. Arrow indicate the amplified 3'-end PCR product. W: negative control reaction primed with 3' oligonucleotide only. M: 1 kb molecular weight marker.

4 Discussion

4.1 Characterization of the 95 kDa SPRMTK protein

In recent years, there has been a veritable explosion in the number of identified protein kinases. Most were first described, not through the traditional biochemical approach of protein purification and enzyme assay, but as putative protein kinase amino acid sequences deduced from the nucleotide sequences of molecularly cloned genes or complementary DNAs. In this project, we characterized the 95 kDa sperm adhesion protein in *porcine*, using the later strategy. The 95 kDa protein, with the characteristics of a protein tyrosine kinase (Burks *et al*., 1995), possesses a catalytic domain in which phosphotyrosine content increases after sperm-zona pellucida binding (Leyton *et al.,* 1992). An important factor in this has been the recognition that all protein kinases have striking sequence similarity in their catalytic domain. Hence the phylogenetic mapping of this highly conserved protein kinase catalytic subunit can serve as a useful first step in the characterization of novel protein kinase.

4.1.1 Conserved landmarks of the catalytic domain.

Amino acid sequences deduced from nucleotide sequences are considered to represent protein tyrosine-specific kinase if they reveal a series of short sequence motifs that are highly conserved in their catalytic domain; amino acid single codes; DFG, GMEY, RDL, DLAARN, VADFG, MPVKWIAIES, DVWAFGV (Hanks *et al*., 1995; Hunter, 1987). The nucleotide sequence of the pig-SPRMTK (95 kDa) catalytic domain amplified using PCR (see section 3.2.2) indicates stretches

of these highly conserved sequence motifs of a protein kinase (see figure 17). Thus the pig-SPRMTK can be classified as a member of the protein tyrosine-kinase (PTK) family.

In general, protein kinase catalytic domains range approximately from 250 to 300 residues, corresponding to about 30 kDa. Fairly precise boundaries for the catalytic domains have been defined through analysis of conserved sequences (Barker and Dayhoff, 1982). The location of the catalytic domain is not fixed throughout the family and in the case of pig-SPRMTK it is located near the carboxylic terminus (see figure 17).

All protein kinases so far characterized with regard to substrate specificity fall within one of two broad classes: serine/threonine-specific and tyrosine-specific. Although these specific classes of protein kinase have similar catalytic domain primary structure, certain short amino acid stretches appear to characterize each class (Hanks, 1987) and those regions can be used to predict whether a putative protein kinase will phosphorylate tyrosine or serine/threonine.

Several lines of evidence strongly support the human 95 kDa protein (SPRMTK) as a member of the tyrosine-specific kinase family (Burks *et al.*, 1995):

1). The presence of the consensus sequence: **DLAARN**, **RDL**, **DFG** (Hunter *et al.*, 1986; Hunter, 1987).
2). It contains phosphotyrosine, the amount of which increases on exposure to ZP proteins (Leyton and Saling, 1989a).
3). PTK activity is stimulated by direct exposure of the isolated receptor to ligand (ZP3) (Leyton *et al.*, 1992).
4). ZP3 receptor aggregation is an initiating signal in the cascade leading to acrosomal exocytosis (Leyton and Saling, 1989a).

Considering these characteristics one can refer to the 95 kDa protein (SPRMTK) as a tyrosine-specific kinase.
On these grounds, applying these evidences to *Sus scrofa* as a case study we suggested initially the existence of a SPRMTK homolog in *Sus scrofa* spermatozoa. The key residues indicated in figure 17 can undoubtedly underscore and reinforce the suggestion that *Sus scrofa* SPRMTK is a protein tyrosine-specific kinase. Additionally, consistent with observations in mice (Leyton and Saling, 1989a), the 95 kDa protein is the major phosphotyrosine-containing protein identified in pig sperm. Furthermore, there is an evolutionary conservation of receptor PTK (Pawson, 1992) in mammals.

ctgtcatcacatcatctntgtacgacgnccagtgaaatccttctgtcggcgagccatcgaa
C H H I I X V R X P V K S F C R R A I E
ctcactttacagagcttgggagtcagtgaggaattgcagaataaactagaagatgttgtg
L T L Q S L G V S E E L Q N K L E D V V
attgacaggaaccttctaattcttggaaaaattctgggcgaaggagagttcgggtctgtc
I D R N L L I L G K I **L G E G E F G S V**
atggaaggaaatcttaatcagcaagatgggacctctcagaaggtggcagtgaaaaccatg
M E G N L N Q Q D G T S Q K V A V K T M
aagttggacaacttttcgcatagagaaattgaggagttcctcagtgaggccgcatgcatg
K L D N F S H R E I E E F L S E A A C M
aaagacttcaaccacccgaatgtcatccgactcttaggcgtatgtatagaaatgagccct
K D F N H P N V I R L L G V C I E M S P
caaggcatccccaagcccatggtgattttacccttcatgaaatacggggacctgcacacc
Q G I P K P M V I L P F M K Y G D L H T
tacttgctctactcccgactggacacaggaccaaagcatattcctctgcagacactgttg
Y L L Y S R L D T G P K H I P L Q T L L
aagttcatggtggacattgcccagggaatggagtatctgagcaacaggaattttcttcat
K F M V D I A Q **G M E Y** L S N R N F L H
cgagatttagctgctcggaactgcatgttgcgagatgacatgaccgtctgcgtggcggac
R D L A A R N C M L R D D M T V C **V A D**
tttggcctctctaagaagatttacagcggcgattactaccgccaaggtcgcatcgctaag
F G L S K K I **Y** S G D **Y Y** R Q G R I A K
atgcctgtgaaatggatcgccatagagagtctcgcagaccgagtctacaccagtaaaagt
M P V K W I A I E S L A D R V Y T S K S
gacgtgtgggcatttggcgtgaccatgtgggaaatagccaaacgaggaatgacccctatc
D V W A F G V T M W E I A K R G M T P I
ccggagttcagaaccatggtcatagctgtttctg
P E F R T M V I A V S

Figure 17 The nucleotide and deduced amino-acid of the pig-SPRMTK catalytic domain. The highly conserved residues are written in bold and underlined. Y predicted autophosphorylation sites.

```
Human   209  KSFCRRAIELTHSLGVSEELQNKLEDVVIDRNLLILGKILGEGEKGTVYEGLWNIPEGK
pig     12   KSFCRRAIELTLQSLGVSEELQNKLEDVVIDRNLLILGKILGEGEFGSVMEGNLNQQDGT

Human   268  EVKIPVAIKTLKLDTMANKEI---LDEASVMKGFGNPHVVRLLGICMTSTIYVITE-YCL
pig     72   SQK--VAVKTMKLDNFSHREIEEFLSEAACMKDFNHPNVIRLLGVCIEMSPQGIPKPMVI

Human   324  LVYRRNKDKAEQHRSNCAELNP---PLQTLLKFMVDIALGMEYLSNRNFLHRDLAARNCM
pig     130  LPFMYGDLHTYLLYSRLDTGPKHIPLQTLLKFMVDIAQGMEYLSNRNFLHRDLAARNCM

Human   381  LRDDMTVCVADFGLSKKIYSGDYYRQGRIAKMPVKWIAIESLADRVYT-KSDVWAFGVTM
pig     190   LRDDTVCVADFGLSKKIYSGDYYRQGRIAKMPVKWIAIESLADRVYTSKSDVWAFGVTM

Human   440  WEIATTLRGMTPYP
pig     250  WEIA--RGMTPIP
```

Figure 18 Amino-acid sequence comparison of the catalytic domains of *porcine* versus *Homo sapiens* reveals a high homology of 74%. Y predicted autophosphorylation sites. **RDL, DFG DLAARN** consensus sequences are highly conserved in protein tyrosine kinase. Dashes represent gaps introduced to optimize the alignment

4.1.2 Phylogeny of the catalytic domain.

Phylogenetic mapping of the conserved protein kinase catalytic domain (see section 3.4) can serve as an additional information tool in the functional classification of *Sus scrofa*-SPRMTK protein.

The protein kinase forms a large family of enzymes, many of which mediate the response of eukaryotic cells to external stimuli. In the case of SPRMTK, the external stimuli will be the ZP3 ligand that activates tyrosine phosphorylation reaction in the sperm and it is very important in initiating the early steps of fertilization (Leyton and Saling, 1989b).

The inhibition of the protein tyrosine kinase activity prevents the stimulation of several intracellular signal transduction pathways (Leyton and Saling, 1989b) and subsequently prevents fertilization.

To compare primary structures, we aligned catalytic domains of fourteen protein tyrosine kinases. The alignment was performed using the clustal W multiple sequence alignment (Algorithm Citation: Higgins *et al.*, 1992; Thompson *et al.*, 1994; Felsenstein, 1989). The alignment clearly demonstrates the overall similarity among the catalytic domains. The domains are not conserved uniformly but, rather, consists of alternating regions of high and low conservation. Features of the alignment are indicated as follows: the three single fully conserved residues are indicated by an asterisk (*), seven strong and seven weak groups conservation are indicated by a semicolon (:) and a point (.) respectively (see appendix 7.2). These conserved domains must be important for catalytic function, either directly as components of the active sites or indirectly by contributing to the formation of the active site through constraints imposed on secondary structure. The non-conserved regions, on the other hand, are likely to occur in loop structures, where folding allows the essential conserved regions to come together. The central core most characteristic of the tyrosine-specific kinase family as already mentioned is the amino acid sequence: Asp-Leu-Ala-Ala-Arg-Asn, Arg-Asp-Leu, Asp-Phe-Gly (single letter codes: **DLAARN, RDL, DFG** respectively) (Hanks, 1987; Hunter, 1987).

The HUMSPRMTK (95 kDa), SSPRMTK (95 kDA) and RET form a cluster as indicated in the phylogenetic tree (see figure 11). Each of these operational taxonomic units (OTUs) displays the short sequence motifs characteristic (see figure 19) of tyrosine-specific protein kinases. This relatedness, by having the same divergence point or ancestral unit, implies that they descended from the same ancestral protein kinase. Additionally, it is evident to suggest

from the branch lengths, that the *Sus scrofa* SPRMTK (/SSPRMTK) evolved earlier than that of *Homo sapiens.* Also the homology suggests a shared cellular signalling pathway.
It is interesting to note that many protein-tyrosine kinases harbor a latent oncogenic potential and hence the cluster of SPRMTKs with the RET (protooncogene tyrosine-protein kinase receptor-in Humans). Such oncogenic homologs provide invaluable clues to the mechanisms of normal growth factors, in addition to enhancing our understanding of cellular transformation. The STE7, protein serine/threonine kinase and gag-raf have least divergence from this cluster, because predicted amino acid sequences shows homology to several members of the protein tyrosine-specific kinase family. The gag-raf primary structure is closely related to the scr family of oncogenes (Mark *et. al.*, 1984), and as such can be described as possessing oncogenic potential.
The amino acid sequence: Asp-Leu-Lys-Pro-Glu-Asn (single letter code, **DLKPEN)** (Hanks, 1987) is highly conserved in serine/threonine-specific kinase proteins. Hence the serine/threonine proteins all cluster together and probably descended from the same ancestral protein kinase. Nevertheless, based on sequence similarities and the phylogenetic tree it appears that all of the fourteen protein aligned, serine/threonine and tyrosine kinase genes arose from a single archetypal gene.

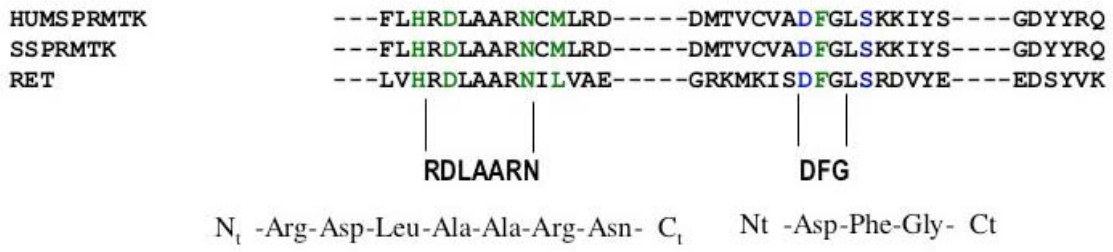

Figure 19 Alignment of the amino acid sequences of the catalytic domains of tyrosine-specific kinase reveals two separate highly conserved segments. Sequences aligned are as follows: HUMSPRMTK- *Homo sapiens* sperm protein tyrosine kinase, SSPRMTK- *Sus scrofa* sperm protein tyrosine kinase and RET- Protooncogene tyrosine-protein kinase receptor (in Humans). **RDL, DLAARN and DFG** consensus sequences characteristic of protein tyrosine-specific kinase family indicated.

The sites of autophosphorylation found in the protein-tyrosine kinase lie in the carboxylic terminus positions for both *Homo sapiens* and *Sus scrofa* (see figure 18). The role of this autophosphorylation site is not entirely settled, but for several protein tyrosine kinase there is evidence that phosphorylation of this site leads to increased catalytic activity (Leyton and Saling, 1989a). One school of thought claims that autophosphorylation may result in a conformational change that allows better access of exogenous substrates to the active site.

The phylogenetic tree reveals not only a relationship between catalytic domain sequences but also certain biochemical properties; catalytic domains from protein kinases having similar modes of regulation or substrate specificities tend also to have similar primary structure and thus cluster together within the tree. This structural relatedness does not only predict similar divergence point but also their function for instance, synthetic peptides, as shown by Burks *et al.*, 1995, corresponding to regions of the

predicted extracellular domain of the HUMSPRMTK inhibits sperm binding to human zona pellucida and subsequently prevents the biochemical events.

Indeed, hints to function may come through the analysis of catalytic domain primary structure and subsequent phylogenetic mapping. A catalytic domain that has only limited divergence from another, well characterized member of the family can be expected to play a similar role in cellular physiology. However, a leap in our understanding of the functional roles of the conserved catalytic domain residues will come with the solution of a crystal structure for one of the protein tyrosine-specific kinase catalytic domains. Hence, similarities in primary structure should carry over to the higher-order structure and catalytic mechanism as well. Further clues are most likely to come from characterization of amino acid sequences lying outside the catalytic domain sequences where residues involved in enzyme regulation may be found.

4.1.3 Expression of the catalytic domain

The first question here was what type of protein is SPRMTK? Expression of protein of prokaryotic origin, the obvious choice of expression system would be *Escherichia coli.* Although SPRMTK protein is of eukaryotic origin our first method of choice was the prokaryotic expression system, BL21(DE3)pLysS. Jansen *et al.,* (1998) successfully used the prokaryotic system BL21(DE3)pLysS in expressing recombinant fragments of the acrosin protein, which is also a member of the sperm adhesion protein. However, many eukaryotic proteins do not fold properly in *E. coli* and form insoluble aggregates (inclusion bodies). To circumvent this problem we expressed at a low temperature and the expression vector pRSET A vector used provides stability to the fusion protein.

The BL21(DE3)pLysS bacteria strain are derivatives of *E. coli* B and naturally lack the Lon protease, which may affect the stability of our recombinant protein (pRSET A-SPRMTK).

The BL21 cell series is tailored solely for protein expression; however, these strains have two significant problems that limit their use as primary transformation hosts. The BL21 cells are derived from the *E.coli* B cell line rather than *E.coli* K-12 cell line, and as such, expression efficiency is relatively low. The other problem causing BL21 competent cell lines to be inadequate for initial transformation experiments is that they contain the EndA gene, which encodes an endonuclease that rapidly degrades miniprep DNA, leading to the so called ”plasmid loss“.

Hence, to retrieve and confirm our plasmid before expression studies we first transformed the ligation mix of the protein expression construct pRSET A-SPRMTK into a high—efficiency, XL1-Blue competent cells (EndA-cell line). Then a re-transformation into BL21 cells was performed for expression. Analysis of the protein on a 12% agarose gel gave the expected size of 27 kDA (see figure 10), which is the approximate size of all known protein tyrosine kinase catalytic domains (Barker and Dayhoff, 1982). The putative protein has an enterokinase cleavage site between the recombinant protein and histidine peptide (Figure 8), to isolate the protein after purification.

The reason for expressing the *porcine* SPRMTK protein, is to explore the biological functions of the tyrosine kinase domain in phosphorylated and dephosphorylated conditions. It has already been shown that in the kinase negative mutant receptors are not able to phosphorylate and as such cannot bind. The key amino acids involved in this interaction can be determined by site-directed mutagenesis as illustrated in proacrosin by Jansen *et al.,* 1998.

4.1.4 Expression analysis of SPRMTK-gene

To investigate the expression of the SPRMTK gene in *Sus scrofa,* we performed RT-PCR using RNAs from five different adult *porcine*-tissues: brain, testis, kidney, spleen and liver. The data obtained provided an accurate identification of the SPRMTK expression in the testicular germ cell, no bands (signals) where amplified in the other tissues (see section 3.2.2). This is consistent to the observations of Burks *et al.,* 1995. The authors analysed the expression pattern of the SPRMTK within the testis using Northern blotting and *in situ* hybridization; localization was confined to the postmeiototic germ cells.
The *porcine* tissues in the Northern Blot analysis included all tissues mentioned above. Hybridizing mRNA with a cDNA probe detected no significant band. Possible explanation for this discrepancy could be that the SPRMTK –mRNA amount in testis is relatively low.
However, the RT-PCR result confirms the testicular germ cell expression, where it plays a key role in the early stages of fertilization.

4.2 Genomic analysis of SPRMTK-gene

4.2.1 Chromosomal localization of the SPRMTK Gene.

The cytogenetic mapping of the PAC clone E11-410 is in agreement with our anticipation and shows that the *porcine* homolog of the *Homo sapiens*-SPRMTK is located on SSC 3q11-12.
In the somatic analysis, the chromosome probability of 96% allocated the gene to chromosome 3 and with 88% on the q-arm region E (see figure 2).

In the radiation hybrid panel we choose to use the LOD> 6 criterion to select what LOD cut-off to use for significant linkage, we want to rule out LODs that might associate with incorrect chromosomal location and to include LODS that identify the correct one. The Two-point LOD scores indicate that the gene is linked to the markers SWR978 (Rohrer *et al.*, 1994) and IL1B (Mellink *et al.*, 1994) The most significantly linked marker (two-point-analysis) is SWR978 on SSC3 (39cR; LOD = 11.08). The marker ILIB has a LOD score of 9.82; one can infer that the correct chromosomal location of SPRMTK is close to the two markers as supported by the multipoint analysis.

The radiation and somatic hybrid results reveal an insight as to the location of SPRMTK-gene. In order to confirm this results a high-resolution mapping was carried out; Flourescense *in situ* hybridization (FISH). The FISH analysis on metaphase cells confirmed the assignment of the gene to chromosome 3q11-12 and the absence of additional hybridization signals demonstrated a signal to noise ratio appropriate for FISH experiments (see figure 3). The human homolog of this gene was assigned to HSA2q13. It appears there is no apparent synteny between *Homo sapiens* and *porcine* at this segment. The reason could be that the boundaries of the conserved synteny between HSA2 and SSC3 still need to be extended and refined.

To the best of our knowledge, this is the first time this gene has been mapped in swine.

4.2.2 Structural features of receptor tyrosine kinases

The family of receptor tyrosine-specific kinases (see figure 20) shares a similar topology. A binding domain is placed on the outside of the cell where it can specifically recognise its ligand. The link to the cytoplasm is provided

by a stretch of hydrophobic amino acids, long enough to cross the plasma membrane once.

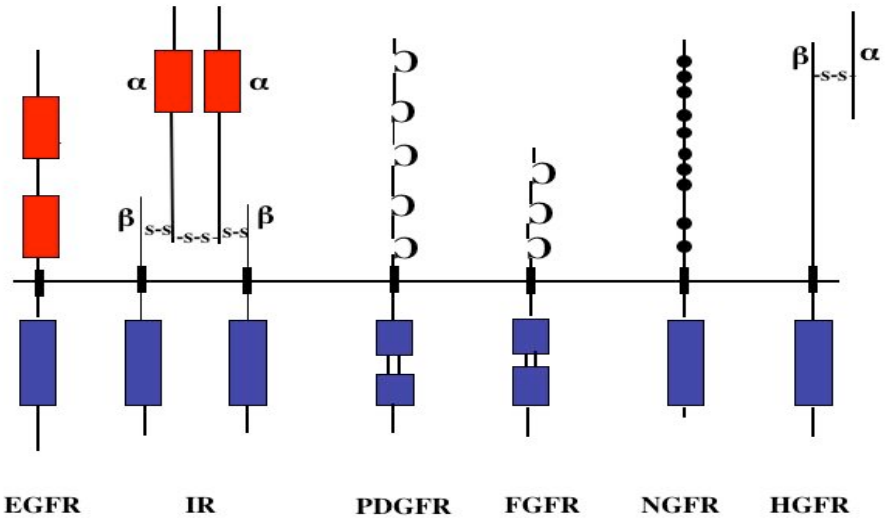

Figure 20 Structural organization of receptor tyrosine kinases. A representative example is shown; the different subclasses are: **EGFR** =epidermal growth factor **IR** =insulin and insulin like growth factor **PDGF** =platelet-derived growth factor **FGFR** =fibroblast growth factor **HGF** =hepatocyte growth factor **NGF** =nerve growth factor. The blue rectangle =Kinase domain; red rectangle =cysteine-rich region; ● =conserved cysteine, ▯ =Kinase insert; ▮ =Transmembrane region; ⊃ =IgG-like domain. The SPRMTK has a single pass transmembrane structure with a catalytic tyrosine kinase domain intracellularly, similar to that of EGFR (Saling *et al.*, 1991).

4.2.3 Genomic organization of the SPRMTK gene

The genomic DNA sequences of the *porcine* SPRMTK gene was determined from the PAC clone E11-410 (see section 3.5.1). The genomic sequences of the human SPRMTK were available from the human genome draft sequence. A comparison between the genomic sequences and the cDNA sequence revealed the presence of 7 exons.

We assigned this gene to chromosome SSC3q11-12 in *porcine* (see section 3.1) and in *Homo sapiens* to HSA2q13 (NIH data).

We isolated and sequenced all but 5 exons: Exon 3,4,5,6,7. The total exon size is 1457 bp, size ranging from exon 4 (98 bp) to exon 6 (756 bp). The location of the exon/intron boundaries in the SPRMTK gene with respect to human is conserved. The gene introns have the canonical GT-AG dinucleotides at the splice junctions.

The exon 6 in *porcine* shows a very high homology to human; region of highly conserved catalytic domain.

Derived RT-PCR product, exon regions were used to design *porcine* specific primers to obtain the upstream and downstream regions of the gene using Rapid Amplification of cDNA Ends (RACE) strategy. Amplification of the 3′ region utilised the smart modified introduced tail primer, together with a *porcine* specific primer (ZRK3-for primer see appendix section 7.3) from exon 6.

Amplification of the 5′ region was not achieved. These might possibly be due to a premature termination of the first-strand cDNA synthesis caused by pausing of RT generally. This is a common problem with larger RNAs, and it is a difficult problem to overcome, since it is due to an intrinsic limitation of RT.

5 BIBLIOGRAPHY:

Adams, M.D., S.E. Celniker, R.A. Holt, C.A. Evans *et al.*, (2000). The genome sequence of Drosophila melanogaster. Science *287*: 2185-2195.

Aitken, J. (1995). The complexities of conception. Science *269*: 39-40.

Al-Bayati, H.K., S. Duscher, S. Kollers, G. Rettenberger, R. Fries, B. Brenig (1999). Construction and characterization of a *porcine* P1-derived artificial chromosome (PAC) library covering 3.2 genome equivalents and cytogenetical assignment of six type I and type II loci. Mammal Gen *10*: 569-572.

Alexander, L.J., D.L. Troyer, G.A. Rohrer, T.P.L. Smith, L.B. Schook, C.W. Beattie (1996). Physical assignments of 68 *porcine* cosmid and lambda clones containing polymorphic microsatellites. Mammal Gen *7*: 368-372.

Barker, W.C., M.O. Dayhoff (1982). Viral *src* gene products are related to the catalytic chain of mammalian cAMP-dependent protein kinase. Proc Natl Acad Sci *79*: 2836-2839.

Bedford, J.M. (1977). Sperm/egg interaction: The specificity of human spermatozoa. Anat Rec *188*: 477-483.

Benoff, S. (1997). Carbohydrates and fertilization: an Overview. Mol Hum Rep *3*: 599-637.

Bercegeay, S., M. Jean, H. Lucas, P. Barriere (1995). Composition of human zona pellucida as revealed by SDS-PAGE after silver staining. Mol Reprod Dev *41*: 355-359.

Berger, T., K.O. Turner, S. Meizel, J.L. Hedrick (1989). Zona pellucida induced acrosome reaction in boar sperm. Biol Reprod *40*: 525-530.

Birnboim, H.C., J. Doly (1979). A rapid alkaline extraction procedure for screening recombinant plasmid. Nucl Acids Res *7*: 1513.

Birren, B., E.D. Green, S. Klapholz, R.M. Myers, J. Roskams (1997). Genome analysis: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, USA.

Bleil, J.D., P.M. Wassarman (1980). Structure and function of the zona pellucida: Identification and characterization of the proteins of the mouse oocyte's zona pellucida. Dev Biol *76*: 185-202.

Bleil, J.D., P.M. Wassarman (1983). Sperm-egg interaction in the mouse: Sequence of events and induction of the acrosome reaction by a zona pellucida glycoprotein. Dev Biol *95*: 317-324.

Bleil, J.D., P.M. Wassarman (1988). Galactose at the non-reducing terminus of O-linked oligosaccharides of mouse oocyte zona pellucida glycoprotein ZP3 is essential for the glycoprotein's sperm receptor activity. Proc Natl Acad Sci USA *58*: 6778-6782.

Bleil, J.D., P.M. Wassarman (1990). Identification of a ZP3-binding protein on acrosome-intact mouse sperm by photoaffinity crosslinking. Proc Natl Acad Sci U.S.A *87*: 5563-5567.

Bookbinder, L.H., H. Cheng, J.D. Bleil (1995). Tissue and species-specific expression of sp56, a mouse sperm fertilization protein. Science *269*: 86-89.

Brede, G., J. Solheim, H. Prydz (2002). PSKH1, a novel splice factor compartment-associated serine kinase. Nucleic Acids Res. *30*: 5301-5309.

Brewis, I.A., H.D. Moore (1997). Molecular mechanisms of gamete recognition and fusion at fertilization. Mol Hum Reprod *12*: 156-165.

Brewis, I.A., R. Clayton, M. Martin, C.E. Browes, C.L.R. Barrat, D.P. Hornby, H.D. Moore (1998). Tyrosine phosphorylation of 95 kDa protein and induction of the acrosome reaction in human spermatozoa in response to recombinant human ZP3. Mol Hum Reprod *12*: 1136-1144.

Bruce, E.K., R.B. Pearson (1990). Protein kinase recognition motifs. TIBs *15*: 342-343.

Burks, D.J., R. Carballada, H.D.M. Moore, P.M. Saling (1995). Interaction of a tyrosine kinase from human sperm with the zona pellucida at fertilisation. Science *269*: 83-86.

Bussey, H., R.K. Storms, A. Ahmed, K. Albermann, E. Allen *et al*., (1997). The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI. Nature *387*: 103-105.

Chapmann, N.R., C.L.R. Barratt (1996). The role of carbohydrate in sperm-ZP3 adhesion. Mol Hum Reprod *2*: 767-774.

Cheng, A., T. Le, M. Palacios, L.H. Bookbinder, P.M. Wassarman, F. Suzuki, J.D. Bleil (1994). Sperm-egg recognition in the mouse: Characterization of sp56, a sperm protein having specific affinity for ZP3. Cell Biol *125*: 867-878.

Cherr, G.N., H. Lambert, S. Meizel, D.F. Katz (1986). In vitro studies of the golden hamster acrosome reaction. Completion on the zona pellucida and induction by homologous soluble zona pellucida. Dev Biol *114:* 119-113.

Doolittle, R.F. (1979). The proteins. H. Neurath and R.L. Hill (eds). Academic press, New York, ed. 3, *4*: 1-118.

Dunbar, B.S., G.A. Maresh, K. Washenik (1989). Ovarian development and formation of the mammalian zona pellucida; in Dietl, J.J. (ed): The Mammalian Egg Coat: Structure and function. Berlin, Springer, pp 38-48.

Dunbar, B.S., S. Avery, V. Lee, S. Prasad, D. Schwahn, E. Schwoebel, S. Skinner, B. Wilkins (1994). The mammalian zona pellucida: its biochemistry, immunochemistry, molecular biology and development expression. Reprod Fertil Dev *6*: 331-347.

Felsenstein, J. (1989). PHYLIP -- Phylogeny Inference Package (Version 3.2). Cladistics *5*: 164-166.

Feng, D.F., R.F. Doolittle (1987). Progressive sequence alignment as a prerequisite to correct phylogenetic tree. Mol Evol *25*: 351.

Fitch, W.M., E. Margoliash (1967). Construction of a phylogenetic tree. Science *155*: 279.

Florman, H.M., B.T. Storey (1982). Mouse gamete interactions: The zona pellucida is the site of the acrosome reaction leading to fertilization in vitro. Dev Biol *91*: 121-130.

Francavilla, S., A. Gabriele, R. Romano, L. Gianaroli, A.P. Ferraretti, F. Francavilla (1994). Sperm-zona pellucida binding of human sperm is correlated with the immunocytochemical presence of proacrosin and acrosin in the sperm heads but not with the proteolytic activity of acrosin. Fertil Steril *62*: 1226-1233.

Fries, A.E., E. Töpfer-Petersen, H. Nguyen, W.B. Schill (1987). Electron microscope localization of a fucose-binding protein in acrosome reacted boar spermatozoa by the fucosylperoxidase-gold method. Histochemistry *86*: 297-303.

Gabius, H.J., (1998). The how and why of protein-carbohydrate interaction: A primer to the theoretical concept and guide to application in drug design. Pharmacol Res *15*: 23-30.

Gabius, H.J., F. Sinowatz (1998). Glycosciences. Acta Anat *161*: 1-276.

Gustavsson, I., (1988). Karyotype of the domestic pig. Heriditas *109*: 151-157.

Gwatkin, R.B.L., D.T. Williams (1977). Receptor activity of the hamster and mouse solubilized ZP before and after the zona reaction. Reprod Fertl *49*: 55-59.

Hanks, S.K. (1987). Homology probing: Identification of cDNA clones encoding members of the protein-serine kinase family. Proc Natl Acad Sci *84*: 388-392.

Hanks, S.K., T. Hunter (1995). Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. FASEB J. *8*: 576-596.

Hardy, D.M., M.N. Oda, D.S. Friend, T.T.F. Huang (1991). A mechanism for differential release of acrosomal enzymes during the acrosome reaction. Biochem *275*: 759-766.

Hardy, D.M., D.L. Garbers (1995). A sperm membrane protein that binds in a species-specific manner to the egg extracellular matrix is homologous to von Willebrand factor. Biol Chem *270*: 26025-26028.

Harris, J.D., D.W. Hibler, G.K. Fontenot, K.T. Hsu, E.C. Yurewicz, A.G. Sacco (1994). Cloning and characterization of zona pellucida genes and cDNAs from a variety of mammalian species: The ZPA, ZPB, and ZPC gene families. DNA Seq *4*: 361-393.

Hedrick, J.L., N.J. Wardrip (1986). Isolation of the zona pellucida and purification of its glycoprotein families from pig oocytes. Dev Biol *121*: 478-488.

Higgins, D.G., A.J. Bleasby, R. Fuchs (1992). CLUSTAL V: improved software for multiple sequence alignment. Computer applications in the Biosciences (CABIOS), *8*: 189-191.

Hinsch, E., S. Oehninger, K.D. Hinsch (1998). ZP3-6 peptide and acrosome reaction; in Vereecken, A., C. De Jonge, D. Franken, S. Oehninger, W. Ombelet (eds): Semen Evaluation, Testing and Selection: A New Look in the 2000s, Pre-Congress training course, Andrology in the Nineties. Modern ART in the 2000s, Genk, pp 25-29.

Hoshiba, H., F. Sinowatz (1998). Immunohistochemical localisation of the spermadhesin AWN-1 in the equine male genital tract. Anat histol Embryol *27*: 351-353.

Hunnicutt, G.R., P. Primakoff, D.G. Myles (1996). Sperm surface protein bifunctional-One activity is a hyaluronidase and the second, distinct activity is required in secondary sperm-zona binding. Biol Reprod *55*: 80-86.

Hunter, T., (1987). A Thousand and One Protein Kinases. Cell *50*: 823-829.

Hunter, T., J.A. Cooper (1986). In the Enzymes. Boyer, P.D. and E.G. Krebs (eds.). Academic Orlando, FL, pp. 191-246.

Ikari, N., H. Yoshino, A.C. Moses, J.S. Flier (1988). Evidence that receptor aggregation may play a role in transmembrane signalling through the insulin-like growth factor-1. Mol Endocrinol *2*: 831-837.

Jansen, S., R. Jones, J. Coadwell (1993). Analysis of poly-sulphate binding domains in boar proacrosin. Biochem Soc Trans *21*: 470.

Jansen, S., M. Quigley, W. Reik, R. Jones (1995). Analysis of polysulphate-binding domains in *porcine* proacrosin, a putative zona adhesion protein from mammalian spermatozoa. Int J Dev Biol *39*: 501.

Jansen, S., R. Jones, I. Jenneckens, B. Marschall, B. Kriegesmann, J. Coadwell, B. Brenig (1998). Site-directed-mutagenesis of boar proacrosin reveals residues involved in binding of zona pellucida glycoproteins. Mol Reprod Dev *51*: 184-192.

Jansen, S., M. Ekhlasi-Hundrieser, E. Töpfer-Peterson (2001). Cell Tissues Organs *168*: 82-92.

Jones, R., S. Jansen (1993). Mechanism of recognition and adhesion during fertilization in mammals. Repro Dev *39*: 47-48.

Kalab, P., P. Visconti, P. Leclera, G.S. Kopf (1994). P95, the major phosphotyrosine-containing protein in mouse spermatozoa, is a hexokinase with unique properties. Biol Chem *269*: 3810-3817.

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

Levin, D.E., C.I. Hammond, R.O. Ralston, J.M. Bishop (1987). Two yeast genes that encode unusual protein kinases. Proc Natl Acad Sci U.S.A. *84*: 6035-6039.

Leyton, L., P.M. Saling (1989a). Evidence that aggregation of mouse sperm receptors by ZP3 triggers the acrosome reaction. Cell Biol *108*: 2163-2168.

Leyton, L., P.M. Saling (1989b). 95KD sperm protein bind ZP3 and serve as tyrosine kinase substrates in response to zona binding. Cell *57*: 1123-1130.

Leyton, L., P. LeGuen, D. Bunch, P.M. Saling (1992). Regulation of mouse gamete interaction by sperm tyrosine kinase. Proc Natl Acad Sci U.S.A. *89*: 11692-11.

Litscher, E.S., P.M. Wassarmann (1993). Carbohydrate-mediated adhesion of eggs and sperm during mammalian fertilization. Trends Glycosci Glycotechnol *5*: 369-388.

Lohia, A., J. Samuelson, (1994). Molecular cloning of an Entamoeba histolytica gene encoding a putative mos family serine/threonine-kinase. Biochim Biophys Acta *1222*: 122-124.

Lui, C., E.S. Litscher, P.M. Wassarman (1997). Zona pellucida glycoprotein mZP3 bioactivity is not dependent on the extent of glycosylation of its polypeptide or on sulfation and sialylation of its oligosaccharides. Cell Sci *110*: 745-752.

Mark, G.E., U.R. Rapp (1984). Primary structure of v-raf: relatedness to the src family of oncogenes. Science *224*: 285-289.

McLeskey, S.B., C. Dowds, R. Carballada, R.R. White, P.M. Saling (1998). Molecules involved in mammalian sperm-egg interaction. Int Rev Cytol *177*: 57-113.

Mellink C.H.M., Y. Lahbib-Mansais, M. Yerle, J. Gellin (1994). Mapping of the regulatory type I alpha and catalytic beta subunits of cAMP-dependent protein kinase and interleukin 1 alpha and 1 beta in the pig. Mammal Gen *5*: 298-302.

Miller, D.J., M.B., Macek, B.D. Shur (1992). Complementarity between sperm surface beta-1,4-galactosyltransferase and egg-coat ZP3 mediates sperm-egg binding. Nature *357*: 589-593.

Miller, D.J., B.D. Shur (1994). Molecular basis of fertilization in the mouse. Semin Dev Biol *5*: 255-264.

Moore H.D.M., T.D. Hartman, A.P. Bye, P. Lutjen, M. de Witt, A.O. Trounson (1987). Monoclonal antibody against a sperm antigen Mr 95 000 inhibits attachment of human spermatozoa to the zona pellucida. Reprod Immun *11*: 157-166.

Moreau E., J. Vilar, M. Lelièvre-Pégorier, C. Merlet, Bénichou, T. Gilbert (1998). Regulation of c-ret expression by retinoic acid in rat metanephros: implication in nephron mass control. J Physiol Renal Physiol *275*: 938-945.

Nakano, M., N. Yonezawa (2001). Localization of sperm-ligand carbohydrate chains in pig zona pellucida glycoproteins. Cells Tissues Organs *168*: 65-75.

Nixon, B., L. Qingxian, M.J. Wassler, C.I. Foote, M.A., Ensslin, B.D. Shur (2001). Galactosyltransferase function during mammalian fertilization. Cells Tissues Organs *168*: 58-64.

Noguchi, S., M. Nakano, (1992). Structure of the acidic N-linked carbohydrate chains of the 55-kDa glycoprotein family (pZP3) from *porcine* zona pellicida. Eur J Biochem *209*: 883-894.

Noguchi, S., N. Yonezawa, T. Katsumata, K. Hasizume, M. Kuwayama, S. Hamano, S. Watanabe, M. Nakano (1994). Characterization of the zona pellucida glycoproteins from bovine ovarian and fertilized eggs. Biochim Biophys Acta *1201*: 7-14.

Notland, T.D., L.S. Davis, G.E. Olsen (1989). Regulation of proacrosin conversion in isolated guinea pig sperm acrosomal apical segments. J Biol Chem *264*: 13586-13590.

O'Brien, R.M., M.A. Soos, K. Siddle (1987). Monoclonal antibodies to the insulin receptor stimulate the intrinsic tyrosine kinase activity by cross-linking receptor molecules. J EMBO *6*: 4003-4010.

Palmer, R.H., J. Ridden, P.J. Parker (1994). Identification of multiple novel, protein kinase C-related gene products. FEBS Lett. *356*: 5-8.

Pawson, T. (1992). Tyrosine kinases and their interaction with signaling proteins. Curr. Opin. Genet. Dev *2*: 4-12.

Phelps, B.M., P. Primakoff, D.E. Koppel, M.G. Low, D.G. Myles (1988). Restricted lateral diffusion of PH-20, a P1-anchored sperm membrane protein. Science *240*: 1780-1782.

Prasad, S.V., S.M. Skinner, C. Carino, N. Wang, J. Cartwright, B.S. Dunbar (2000). Structure and function of the proteins of the mammalian zona pellucida. Cells Tissues Organs *166*: 148-164.

Reed, K.C., Mann, D.A. (1985). Rapid transfer of DNA from agarose gels to nylon membranes. Nucl Acids Res *13*: 7207-7221.

Rohrer, G.A., L.J. Alexander, J.W. Keele, T.P. Smith, C.W. Beattie (1994). A microsatellite linkage map of the *porcine* genome. Genetics *136*: 231.

Saiki, R.K., D.H. Gelfand, S. Stoffel, S.J. Scharf *et al.*, (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science *239*: 487-491.

Saling, P.M. (1991). How the egg regulates sperm function during gamete interaction: facts and fantasies. Biol Reprod *11*: 339-388.

Sambrook, J.E., E.F. Fritsch, T. Maniatis (1989). Molecular cloning: A laboratory manual. Cold Spring Harbour Laboratory Press, Cold Spring Harbor, NY, USA.

Sinowatz, F., J. Plendl, S. Kölle (1998). Protein carbohydrate interactions during fertilization. Acta Anat *161*: 196-205.

Snell, W.J., J.M. White (1996). The molecules of mammalian fertilization. Cell *85*: 629-637.

Solinas-Toldo, S., A. Mezzelani, G.A. Hawkins, M.D. Bishop, I. Olsaker, A. Mackinley, L. Ferretti, R. Fries (1995). Combined Q-banding and fluorescence in situ hybridization for the identification of bovine chromosomes 1 to 7. Cytogenet Cell Genet *69*: 1—6.

Solis, D., J. Jimenez-barbero, H. Kaltmer, A. Romero, H.-C. Siebert, C.-W. von der Lieth, H.-J. Gabius (2001). Towards defining the role of glycans as hardware in information storage and transfer: Basic principles, experimental approaches and recent progress. Cells Tissues Organs *168*: 5-23.

Strausberg, R.L., E.A. Feingold, L.H. Grouse, J.G. Derge (2002). Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. Proc Natl Acad Sci U.S.A. *99*: 16899-16903.

Teague, M.A., D.T. Chaleff, B. Errede (1986). Nucleotide sequence of the yeast regulatory gene STE7 predicts a protein homologous to protein kinases. Proc Natl Acad Sci. U.S.A *83*: 7371-7375.

Thompson J.D., D.G. Higgins, T.J. Gibson (1994). "CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice." Nucl Acids Res. *22*: 4673-4680.

Toda, T., S. Cameron, P. Sass, M. Zoller, M. Wigler (1987). Three different genes in S. cerevisiae encode the catalytic subunits of the cAMP-dependent protein kinase. Cell *50*: 277-287.

Toldo, S.S., R. Fries, P. Steffen, H.L. Neibergs, W. Barendse, J.F. Womack, D.J. Hetzel, G. Stranzinger (1993). Physically mapped, cosmid-derived microsatellite markers as anchor loci on bovine chromosomes. Mammal Gen *4*: 720—727.

Töpfer-Peterson, E., A.E. Fries, H. Nguyen, W.B. Schill (1985). Evidence for a fucose-binding protein in boar spermatozoa. J Histochem *83*: 139-145.

Töpfer-Peterson, E., A. Henschen (1987). Acrosin shows zona and fucose binding, novel properties for a serine protease. FEBS lett *226*: 38-42.

Töpfer-Peterson, E., J.J. Clavete, W. Schäfer, A. Henschen, (1990a). Complete localization of the disulfide bridges and glycosylation sites in boar sperm acrosin. FEBS Lett *275*: 139-142.

Töpfer-Peterson, E., D. Cechova, A. Henschen, M. Steinberger, A.E. Fries, A. Zucker (1990b). Cell biology of acrosomal proteins. Andrologia *22 (suppl 1)*: 110-121.

Töpfer-Peterson, E., A. Romero, P.F. Varela, M. Ekhlasi-Hundrieser, Z. Dostalova, L. Sanz, J.J. Calvete (1998). Spermadhesins: A new protein family. Facts hypotheses and perspectives. Andrologia *30*: 217-224.

Töpfer-Peterson, E. (1999a). Carbohydrate based interaction on the route of a spermatozoon to fertilization. Hum Reprod Update *5*: 314-329.

Töpfer-Peterson, E. (1999b). Molecules on the sperm's route to fertilization. J Exp Zool *285*: 314-329.

Tulsiani, D.R.P., H. Yoshiba-Komiya, Y. Araki (1997). Mammalian fertilization: A carbohydrate-mediated event. Biol Reprod *57*: 487-494.

Ullrich, A., J. Schlessinger (1990). Signal transduction by receptors with tyrosine kinase activity. Cell *61*: 203-212.

Urch, U.A., H. Patel (1991). The interaction of boar sperm proacrosin with its natural substrate, the zona pellucida and with its natural substrate, the zona pellucida, and with polysulfated polysaccharides. Development *111*: 1165-1172.

Wassarman, P.M. (1982). Fertilization; in Yamada, K. (ed): Cell Interactions and Development. Molecular Mechanims. New York, Wiley, pp 1-27.

Wassarman, P.M. (1987). The biology and chemistry of fertilization. Science *235*: 553-560.

Wassarman, P.M. (1990). Profile of a mammalian sperm receptor. Develop *108*: 1-17.

Wassarman, P.M., S. Mortillo (1991). Structure of the mouse egg extracellular coat, the zona pellucida. Int Rev Cytol *130*: 85-110.

Wassarman, P.M. (1992). Regulation of mammalian fertilization by gamete adhesion molecules; in Nieschlag, E., U.F. Habenicht 8eds): Spermatogenesis, Fertilization, contraceptive. Molecular, Cellular and Endocrine Events in Male Reproduction. Berlin, Springer, pp 345-366.

Wassarman, P.M., C. Liu, J. Chen, H. Qi, E.S. Litscher (1998). Ovarian development in mice bearing homozygous or heterozygous null mutations in zona pellucida glycoprotein gene mZP3. Histol Histopathal *13*: 293-300.

Wassarman, P.M. (1999a). Mammalian fertilization: Molecular aspects of gamete adhesion, exocytosis, and fusion. Cell *96*: 175-183.

Wassarman, P.M., E.S Litscher (2001). Towards the molecular basis of sperm and egg interaction during mammalian fertilization. Cells Tissues organs *168*: 36-45.

Westbrook-case, V.A., V.P. Winfrey, G.E. Olsen (1994). A domain-specific 50-kilodalton structural protein of the acrosomal matrix is processed and released during the acrosome reaction in the guinea pig. Biol Reprod *51*: 1-13.

Yamakawa, A., Y. Kameoka, K. Hashimoto, Y. Yoshitake, K. Nishikawa, K. Tanihara, T. Date (1997). cDNA cloning and chromosomal mapping of genes encoding novel protein kinases termed PKU-alpha and PKU-beta, which have nuclear localization signal. Gene *202*: 193-201.

Yanagimachi, R. (1988). Mammalian fertilization; in Knobil, E., J.D. Neill (eds): The Physiology of Reproduction. New York, Raven Press, pp 135-185.

Yanagimachi, R. (1994). Mammalian fertilization; in Knobil, N., J.D. Neill (eds): The physiology of Reproduction. New York, Raven Press, pp 189-317.

Yerle, M., G. Echard, A. Robic, A. Mairal, C. Dubut-Fontana, J. Riquet, P. Pinton, D. Milan, Y. Lahib-Mansais, J. Gellin (1996). A somatic cell hybrid panel for pig regional gene mapping characterized by molecular cytogenetics. Cytogenet Cell Genet *73*: 194–202.

Yerle, M., P. Pinton, A. Robic, A. Alfonso, Y. Palvadeau, C. Delcros, R. Hawken, L. Alexander, C. Beattie, I. Schook, D. Milan, J. Gellin (1998). Construction of a whole genome radiation hybrid panel for high resolution gene mapping in pigs. Cytogenet Cell Genet *82*: 182–188.

6 SUMMARY

There are still many unanswered questions regarding mammalian sperm interaction with the oocyte. Although efforts have been made to establish the characteristics of the protein present in the sperm surface, which act as the receptor for ZP3 glycoprotein, the exact function for these proteins in the mechanism associated to the ZP3-induced acrosome reaction has not yet been elucidated. An understanding of the connection between protein phosphorylation and activation of specific genes seems distant, at best.

One class of molecule that accounts for ZP3 recognition, docking and signal transduction is the 95 kDa protein tyrosine kinase (SPRMTK).

To set the framework of this project, a gene expression studies was pursued. Using the conventional RT-PCR method, five *porcine* tissues were PCR-screened. The expression studies deemed sufficient to confirm the expression of the molecule in the *porcine* testicular germ cells; expressed at postmeiototic stage. The amplified RT-PCR product was verified as a SPRMTK by sequence alignment of the nucleotide sequence with that of *murine* and *Homo sapiens* revealing a sequence homology of 88.5% and 74% respectively. The homology between *murine* and *Homo sapiens* is 74.5%

The 95 kDa protein can be described as having a typical single pass transmembrane structure with a catalytic tyrosine domain intracellularly, which activates the acrosome reaction. To characterize the protein, the catalytic domain was amplified using primers designed from the published human SPRMTK sequence (Accession number: L08961). Alignment of the amino acid sequences of *Homo*

sapiens and *porcine* for maximum homology revealed several short stretches within the catalytic domain where amino acids are highly conserved throughout the protein kinase family: (single amino acid codes: DFG, GMEY, RDL, DLAARN, VADFG, MPVKWIAIES, DVWAFGV. These sequences classified the 95 kDa protein to the protein kinase family. Furthermore, some residues (DFG, RDL and DLAARN) within this stretches distinguished the 95 kDa into a subfamily, the tyrosine-specific kinase. The amino acid comparison to human 95 kDa revealed also the predicted position of the phosphorylated tyrosine residues near the carboxylic terminal end.

To characterize the evolution and predict the function of the protein tyrosine-specific kinase a phylogenetic tree (rooted tree) of the conserved catalytic domain of fourteen protein kinase was plotted using the clustal W alignment and the PHYLIP- format dendrogram program. The two protein tyrosine kinase subfamilies were included in this tree: serine/threonine kinase and tyrosine kinase. The 95 KDa protein formed a cluster with tyrosine-specific kinase proteins. The data suggest that the *porcine* 95 kDa protein has a signal transduction function and an oncogenic potential likewise the other members of the cluster. However, despite their diverse origin these proteins may have evolved from a common ancestral protein.

To an effort of better understanding of the mechanism of binding, the catabolic domain was expressed in a prokaryotic expression system BL21(DE3)pLysS using the pRSET A vector. The size of the expressed protein was confirmed to have the expected molecular weight of a catalytic domain of protein kinases, 27 kDa.

The *porcine* 95 kDa gene was isolated from a *porcine* PAC library using PCR screening. The PCR primers were designed from the human SPRMTK sequence and a positive clone was

identified. Subsequent hybridization was performed with an ECL labelled 150 bp PCR-fragment. The SPRMTK gene covers seven exons. All five exons sequenced confirm the consensus exon/intron boundaries of other eukaryotic genes. The partial cDNA (3'-end) sequence of the gene was also isolated.

The isolation of the gene allowed the physical mapping of the gene. Screening of the somatic hybrid and radiation hybrid panel assigned the gene to chromosome SSC3q11-12 with 96% concordance. This localization was confirmed using the high resolution Fluorescence *in situ* hybridization. Comparative mapping confirms the gene to human HSA2q13.

7 APPENDIX:

7.1 **Table 10** Abbreviations for Amino Acids

Amino Acid	Three-letter Abbreviation	One-letter Symbol
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Asparagine	Asx	B
Cysteine	Cys	C
Glutamine	Gln	Q
Glutamic acid	Glu	E
Glutamine	Glx	Z
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

7.2 Figure 21 Multiple amino acid sequence alignment of 14 protein kinase catalytic domains. See table number for classification of the proteins. HUMSPRMTK is chosen as prototype protein kinase; the numbers indicate residue position from the polypeptide amino terminus. Gaps represented by dashes were introduced into the sequences to optimize the alignment. Features of the alignment indicated are as follows: the three single fully conserved residues are indicated by an asterisk (*), seven strong and seven weak groups conservation are indicated by a semicolon (green :) and a point (blue .) respectively.

```
                                224
PKA2                            SQLPQK-----SLVS KGKYTLHDFQIMRTLGTGSFGRVHLVRSVHNG--
TPK2P                           SQLPQK-----SLVS-KGKYTLHDFQIMRTLGTGSFGRVHLVRSVHNG--
PCK-53E                         SDTNTH-----TSSK-KDMIRATDFNFIKVLGKGSFGKVLLAERKGSE--
PKC-1                           ETQETPGPALCSPLR-KSPLTLEDFKFLAVLGRGHFGKVLLSEFRPSG--
HUMSPRMTK                       SEELQN------KLE-DVVIDRNLLILGKILGEGEKGTVYEGLWNIPEGK
SSPRMTK                         SEELQN------KLE-DVVIDRNLLILGKILGEGEFGSVMEGNLNQQDG-
RET                             MENQVSVDAFKILEDPKWEFPRKNLVLGKTLGEGEFGKVVKATAFHLKGR
PSK-H1                          RAKFDP-------------RVTAKYDIKALIGRGSFSRVVRVEHRATR--
Gag-raf                         QRDSSY-----------YWKMEASEVMLSTRIGSGSFGTVYKGKWHG----
Protein_Serine/threonine_Ki     ETEQTC------------ILDPDELIQKKKIGEGTFGVVYKGEFKGNS--
STE7                            DLKDTLSG-----TSNGNYIQLQDLVQLGKIGAGNSGTVVKALHVPDS--
CDK4                            -------------------MATSRYEPVAEIGVGAYGTVYKARDPHSG--
KIN2                            PTYMSKSNEISIKVPKSHSRTISDYIPSARRYPSYVPNSVDVKQKPAKN-
PKU                             IRELKRINNEDNSQFKDHPTLNERYLLLHLLGRGGFSEVYKAFDLYEQR-
                                                                  .
                                268
PKA2                            ---RYYAIKVLKKQQVVK------------MKQVEHTNDERRMLKLVEHP
TPK2P                           ---RYYAIKVLKKQQVVK------------MKQVEHTNDERRMLKLVEHP
PCK-53E                         ---ELYAIKILKKDVIIQ------------DDDVECTMIEKRVLALGEKP
PKC-1                           ---ELFAIKALKKGDIVA------------RDEVESLMCEKRILAAVTSA
HUMSPRMTK                       EVKIPVAIKTLKLDTMAN----------------KEILDEASVMKGFGNP
SSPRMTK                         -TSQKVAVKTMKLDNFSH------------R-EIEEFLSEAACMKDFNHP
RET                             AGYTTVAVKMLKEN--AS------------PSELRDLLSEFNVLKQVNHP
PSK-H1                          ---QPYAIKMIETKYREG------------R---EVCESELRVLRRVRHA
Gag-raf                         ----DVAVKILKVVDP-T------------PEQLQAFRNEVAVLRKTRHV
Protein_Serine/threonine_Ki     -----VAIKRMKPKINDN------------SSEIEFR-KEVEMLEKFRCN
STE7                            ----KIVAKKTIPVEQNN------------STIINQLVRELSIVKNVKPH
CDK4                            ---HFVALKSVRVPNGGGG----------GGGLPISTVREVALLRRLEAF
KIN2                            -TTIAPPIRSVSQKQNSDLPALPQNAELIVQKQRQKLLQENLDKLQINDN
PKU                             --YAAVKIHQLNKSWRDEK----------KENYHKHACREYRIHKELDHP
                                       :                          *
```

```
                                  299
PKA2                              ---FLIRMWGTFQDARN--------IFMVMDYIEGGEL----FSLLRKSQ
TPK2P                             ---FLIRMWGTFQDARN--------IFMVMDYIEGGEL----FSLLRKSQ
PCK-53E                           --PFLVQLHSCFQTMDR--------LFFVMEYVNGGDL----MFQIQQFG
PKC-1                             GHPFLVNLFGCFQTPEH--------VCFVMEYSAGGDL----MLHIHS-D
HUMSPRMTK                         ---HVVRLLGICMTSTIYVITEYCLLV----YRRNKDK----AEQHRSNC
SSPRMTK                           ---NVIRLLGVCIEMSPQGIPKPMVILPFMKYGDLHTY----LLYSRLDT
RET                               ---HVIKLYGACSQDGP--------LLLIVEYAKYGSL----RGFLRESR
PSK-H1                            N---IIQLVEVFETQER--------VYMVMELATGGEL----FDRIIAKG
Gag-raf                           N----ILLFMGYMTKDN--------LAIVTQWCEGSSL----YKHLHVQE
Protein_Serine/threonine_Ki       ---YIIHFYGAVIIQDN--------KCMVTEYAKYGSV----QKMIESKP
STE7                              EN--IITFYGAYYNQHINN-----EIIILMEYSDCGSL----DKILSVYK
CDK4                              EHPNVVRLMDVCATSRTDR---EIKVTLVFEHVDQDLR----TYLDKAPP
KIN2                              DNNNVNAVVDGINNDNSDHYLSVPKGRKLHPSARAKSVGHARRESLKFTR
PKU                               R---IVKLYDYFSLDTDT-------FCTVLEYCEGNDLDFYLKQHKLMSE
                                        .
                                  319
PKA2                              ----------------------RFPNP---VAKFYAAEVILALEYLH-AHN
TPK2P                             ----------------------RFPNP---VAKFYAAEVILALEYLH-AHN
PCK-53E                           ----------------------KFKEP---VAVFYAAEIAAGLFFLH-TKG
PKC-1                             ----------------------VFSEP---RAIFYSACVVLGLQFLH-EHK
HUMSPRMTK                         ----------------------AELNPPLQTLLKFMVDIALGMEYLS-NRN
SSPRMTK                           ----------------------GPKHIPLQTLLKFMVDIAQGMEYLS-NRN
RET                               KVGPGYLGSGGSRNSSSLDHPDERALTMGDLISFAWQISQGMQYLA-EMK
PSK-H1                            ----------------------SFTER---DATRVLQMVLDGVRYLH-ALG
Gag-raf                           ----------------------TKFQMFQLIDIARQTAQGMDYLH-AKN
Protein_Serine/threonine_Ki       ----------------------SNSLSKSIKIKMLLDIARGIEYLH-NNG
STE7                              RFVQRG-----------TVSSKKTWFNELTISKIAYGVLNGLDHLYRQYK
CDK4                              ----------------------PGLPAETIKDLMRQFLRGLDFLH-ANC
KIN2                              PPIPAALPPSDMTNDNGFLGEANKERYNPVSSNFSTVPEDSTTYSNDTNN
PKU                               ----------------------------KEARSIVMQIVNALRYLNEIKP
                                                                    .  .
                                  366
PKA2                              ---IIYRDLKPENILLDR-----NGHIKITDFGFAKEVQT-------VTW
TPK2P                             ---IIYRDLKPENILLDR-----NGHIKITDFGFAKEVQT-------VTW
PCK-53E                           ---ILYRDLKLDNVLLDA-----DGHVKIADFGMCKENIV----GDKTTK
PKC-1                             ---IVYRDLKLDNLLLDT-----EGYVKIADFGLCKEGMG----YGDRTS
HUMSPRMTK                         ---FLHRDLAARNCMLRD-----DMTVCVADFGLSKKIYS----GDYYRQ
SSPRMTK                           ---FLHRDLAARNCMLRD-----DMTVCVADFGLSKKIYS----GDYYRQ
RET                               ---LVHRDLAARNILVAE-----GRKMKISDFGLSRDVYE----EDSYVK
PSK-H1                            ---ITHRDLKPENLLYYHPG--TDSKIIITDFGLASARKKG---DDCLMK
Gag-raf                           ---IIHRDMKSNNIFLHEG-----LTVKIGDFGLATVKSR----WSGSQQ
Protein_Serine/threonine_Ki       ---ILHRDIKPDNMLITSLDNDIPVNAKLTDFGSARNINS----LMTNMT
STE7                              ---IIHRDIKPSNVLINS-----KGQIKLCDFGVSKKLIN------SIAD
CDK4                              ---IVHRDLKPENILVTS-----GGTVKLADFGLARIYSY-------QMA
KIN2                              RLTSVYSQELTEKQILEEASKAPPGSMPSIDYPKSMFLKGF---FSVQTT
PKU                               --PIIHYDLKPGNILLVDG--TACGEIKITDFGLSKIMDDDSYGVDGMDL
                                       : :    : :              *:  .
```

```
                                  407
PKA2                              TLCGTP--DYIAPEVITTK--------------PYNKSVDWWSLGVLIYE
TPK2P                             TLCGTP--DYIAPEVITTK--------------PYNKSVDWWSLGVLIYE
PCK-53E                           TFCGTP--DYIAPEIILYQ--------------PYGKSVDWWAYGVLLYE
PKC-1                             TFCGTP--EFLAPEVLTDT--------------SYTRAVDWWGLGVLLYE
HUMSPRMTK                         GRIAKMPVKWIAIESLADR--------------VYT-KSDVWAFGVTMWE
SSPRMTK                           GRIAKMPVKWIAIESLADR--------------VYTSKSDVWAFGVTMWE
RET                               RSQGRIPVKWMAIESLFDH--------------IYTTQSDVWSFGVLLWE
PSK-H1                            TTCGTP--EYIAPEVLVRK--------------PYTNSVDMWALGVIAYI
Gag-raf                           VEQPTGSVLWMAPEVIRMQDD-----------NPFSFQSDVYSYGIVLYE
Protein_Serine/threonine_Ki       FTKGVGTPSFMAPEILKRK--------------KYKTAADIYSFAISITL
STE7                              TFVGTS--TYMSPERIQGN--------------VYSIKGDVWSLGLMIIE
CDK4                              LTPVVVTLWYRAPEVLLQS--------------TYATPVDMWSVGCIFAE
KIN2                              SSKPLPIVRHNIISVLTRMNIDFKEVKGGFICVQQRPSIETAAVPVITTT
PKU                               TSQGAGTYWYLPPECFVVGKEP----------PKISNKVDVWSVGVIFFQ
                                             . :                       :  .
```

7.3 Table 11 Sequence information of the primers used in this study

Name	Sequence
Primers used for screening:	
HUMSPRMTK$_1$ forward	5´- GCG AGA TGA CAT GAC TGT CT -3´
HUMSPRMTK$_2$ reverse	5´- CTG TAC CCA TGG CAC GCA CA -3´
RT-PCR:	
SP1f	5'-GCG AGA TGA CAT GAC TGT CT-3'
SP1r	5'-CTG TAC CCA TGG CAC GCA CA-3'
SP2f	5'-TGC TCA TCA TCX TXG GCT GC -3'
SP3f	5'-GAA ATC CTT CTG TCG GCG -3'
SP4r	5'-ATG GTT ATG AAC TCC GGG -3'
Expression primers:	
ZRK-mut.1 f	5'-CCA GT**G GAT CC**T TCT GTC GGC GAG CCA-3'
ZRK-mut.2 r	5'-C GCT GTA **AAG CTT** CTT AGA GAG GCC-3'
Exon primers:	
Ex3f	5'-CCTATTGCCAGTTCTAGTCCAG-3'
Ex3r	5'-CCAAGGGTCTACCTGGTCCCA-3'
Ex4f	5'-CCCAGTGGACTGACTGGGAG-3'
Ex4r	5'-AACTGAAGTAAACCAGCAAG-3'
Ex5f	5'-AGGACCACTCTAAGTCCATGC-3'
Ex5r	5'-TGGACTTGAACAGAGATCCG-3'
Ex6f	5'-GGCGCCTGGCAACGCAGATC-3'
Ex6r	5'-ATTCCTGTCAATCACAACAT-3'
Ex7f	5'-CGGGGACCTGCATACTTACTTAC-3'
Ex7r	5'-CAGGTCCTCTCACTAACCCTGC-3'
cDNA primers:	
3'-Smart primer	5'-TGGTAACAACGCAGAGTACGCG-3'
5'-Smart primer	5'-AGTGGTAACAACGCAGAGTACTTTTT-3'
ZRK3-for	5'-AGTTCATGGTGGATATTGCCCTGGGAA-3'

ACKNOWLEGEMENTS

This research was conducted in the Institute of Vertinary Medicine, Goettingen, Germany.

During my doctorate thesis, a lot of people encouraged me to finish my research in various ways, which I will gratefully acknowledge.

The first person to direct my appreciation to is my principal supervisor PD Dr. Stephan Jansen, for his support, encouragement and invaluable guidance throughout my research. It is a great fortune for me in having had the opportunity to work under his supervision. Also, he devoted a lot of time and care to critically reading all drafts of this thesis and gave expert editorial advice for completion of this thesis. Thanks also for organising my finances and *Aufenthaltungsbewillingung.*

I am extremely grateful to Prof. Dr. H.-J. Fritz for accepting to be my examiner, likewise Prof. Dr. med. W. Engel my co-examiner from the biological and medical faculties respectively.

My sincere gratitude goes to Prof. Dr. Dr. Bertram Brenig, CEO of this research group, who provided stimulating suggestions throughout this research and his ‘open door’ for all kind of problems. I will always bear his suggestions in my mind as a scientific researcher for the rest of my life. Many thanks!

I am also indebted to Dr. Christoph Knorr for having an open mind, patience and enthusiasm in suggesting new strategies. Also his warm hospitality in the Laboratory.

Thanks to Dr. med. Felix A. Habermann for performing the FISH analysis and Dr. med. vet. Wilhelm Wemheuer for his never-ending support in providing tissue samples.

I have much pleasure and delight in acknowledging all of my colleagues in the Vertinary Institute including past and present: Jule Beck, Monika Mokisz, Tanja Dziomba, Monique Germerodt, Kirsten Bornemann-Kolatzki, Annette Müller, Dr. Ina Pfeiffer, Kesinee Gatphayak, Kiefer Chen, Jun Ren, Viola Raupach, Christian Beuermann, Dr. med. Ekkehard Schütz, Alexandra Siebels, Uli Peters, Sonja and Gerold Kierstein, Brigitte Dierkes, all you guys helped me in one way or the other. *Ich Danke euch*!

Last but not the least, my thanks to Dr. Reza Nezazamdeh who shared most of my time in this laboratory and shared meaningful insight in molecular biology and all.

All the good friends that I met here in Goettingen: This thesis will not only remind me of the hours that I spend in the lab., but as well of the fun that we had with the basketball competition in and around Goettingen.

Finally, I would like to gratefully acknowledge all the support of my parents. Without their help, I could not start this research, either.

This Thesis was written on Prof. Dr. Dr. Bertram Brenig's power book (Mac OS X system) using Microsoft word 98.

Declaration

This work presented in this thesis is original and was undertaken in the Institute of Vertinary Medicine while I was enrolled as a doctorate student in the faculty of biology, university of Göttingen.
I certify that this thesis has not been submitted for any other degree, and that all sources of information and assistance during the experimental work and preparation of the thesis are duly acknowledged.

Curriculum Vitae:

Name	Leonard Bull
Date of Birth	1.12.68
Place of birth	Yengema/ Freetown-Sierra Leone
Nationality	Sierra Leonean
Profession	Diploma in Biology (/Msc. in Biology.)

72-79	Primary school (United Methodist Church School)
80-87	Secondary school (Y.S.S. and the Prince of Wales School Freetown)
88-89	Fourah Bay College in Freetown
91-98	Graduation 1998, Title: Diploma in Biology (/Msc. in Biology) Master thesis (Diplomarbiet): Max Planck Institute for Experimental Medicine in Goettingen
99-03	PhD studies and research at the Institute of Veterinary Medicine Goettingen