Chapter - 1

Introduction
1.1 BIOMARKERS IN PANCREATIC CANCER

Pancreatic ductal adenocarcinoma (PDAC), whose nomenclature derives from its histological resemblance to ductal cells, is the most common pancreatic neoplasm. It accounts for >85% of pancreatic tumors with a median survival of <6 months and a dismal 5-year survival rate of 3% - 5% (Li et al., 2004; Warshaw AL, 1992). Pancreatic cancer is considered one of the most lethal forms of cancer in the United States and is the fourth and fifth leading cause of cancer deaths in men and women respectively. According to the American Cancer Society, approximately, 43,140 new PDAC cases and 36,800 PDAC-related deaths have been estimated in the year 2010 (Jemal et al., 2010). The incidence of pancreatic cancer in India is relatively low (0.5 - 2.4 per 100,000 men and 0.2 - 1.8 per 100,000 women) (Dhir V, 1999; SV Shrikhande, 2009). Various risk factor for pancreatic cancer include age, smoking, hereditary and chronic pancreatitis, increased body-mass index, late onset of diabetes mellitus without diabetes risk factors and inherited predisposition (Fuchs CS, 1996; Ghaneh et al., 2007; Li D, 2009; Michaud DS, 2001).

Pancreas consists of two major portions: (1) exocrine pancreas which has a branching network of acinar and ductal cells that secrete digestive enzymes into the pancreatic duct and (2) endocrine pancreas which regulates glucose homeostasis and metabolism, with a network of cells called islets of langerhans (Hezel et al., 2006). Cancers of endocrine pancreas, which make up only about 1% of the total cases, are far less common than those of exocrine pancreas (James C. Yao, 2007). PDAC commonly arises in the upper portion (head) of the pancreas and the resulting
cancerous cells and byproducts infiltrate into the surrounding tissues including lymphatics, peritoneum and blood stream with metastasis usually to the liver and lungs. Based on clinical and histopathological features, three types of precursor lesions of PDAC have been identified: PanIN (pancreatic intraepithelial neoplasm), MCN (mucinous cystic neoplasm) and IPMN (intraductal papillary mucinous neoplasm) (Brugge, 2004). Of these, PanINs are the most common neoplasms. PanINs have been further classified into three types: PanIN 1, 2 and 3. The high-grade PanINs ultimately transform into PDAC invading regions beyond the basement membrane. The progression of pancreatic cancer is shown in Figure-1.1 (modified from (Bardeesy and DePinho, 2002; Jörg et al., 2007)).

![Image of tumor progression]

**Figure-1.1: Course of pancreatic cancer:** Based on the clinical and histo-pathological features, PDAC precursor lesions have been classified into three groups: PanINs (pancreatic intra neoplasia which are sub grouped into PanIN 1, 2 and 3), MCN (mucinous cystic neoplasm) and intraductal papillary mucinous neoplasm (IPMN) (Hruban et al., 2000).
Although the biology of pancreatic cancer is well investigated, prognosis of the disease has not been satisfactory. In addition, due to the lack of effective early detection and screening methods, pancreatic tumors are usually recognized at very late stages, frequently after metastasis. Biomarkers that can indicate the prevalence of PDAC at an early stage such as PanIN 1 (Fig 1.1) would be very helpful in selecting proper treatment options. Serum tumor markers, including carbohydrate antigen (CA) 19-9, CA242, CA50 and carcinoembryonic antigen (CEA) have been widely used to detect pancreatic cancer. However, their diagnostic value is limited by their false positive rate (low sensitivity) in patients with benign pancreatic disorders such as pancreatitis (Akdoğan M, 2001; Goonetilleke and Siriwardena, 2007; Ni et al., 2005) as well as with other cancers (Ychou et al., 2000). Although imaging techniques such as CT scan are used for PDAC diagnosis, they cannot detect the cancer at early stages.

Since surgical biopsy of the pancreas is invasive, expensive and has a risk of infection, serum-based clinical tests may prove to be helpful and effective in the early diagnosis of PDAC. In the recent years, extensive proteomic profiling of serum samples from PDAC patients led to the identification of various novel potential biomarkers. (Faca et al., 2008; Gronborg et al., 2006; Honda et al., 2005; Sitek et al., 2009). However, a major rate limiting step that still remains a challenge in the early detection of PDAC is unavailability of effective reagents and screening assays to identify these markers. Despite the identification of several potential tumor markers, prognosis of PDAC has not changed substantially for the last few years (Fry et al., 2008; ZerĂłn et al., 2009). Therefore, one of the major goals in clinical pancreatology as well as the goal of the
present study is to develop assays to detect novel biomarkers that are specific and sensitive enough to make an early and accurate diagnosis of pancreatic cancer before it has disseminated and become untreatable.

One of the major challenges in the early detection of PDAC is distinguishing it from chronic pancreatitis (CP). CP is a persistent inflammation of the pancreas associated with structural damage with fibrosis and ductal structures. It usually occurs in the head of the pancreas which makes it very difficult to distinguish it from PDAC. Various types of CP that include alcoholic, autoimmune, and paraduodenal pancreatitis are often misdiagnosed as PDAC or vice-versa because both conditions are clinically and morphologically identical (Kloppel, 2007). Despite the availability of published literature for the differential diagnosis of PDAC versus CP (Klöppel G, 2009; Yian et al., 2007), sensitive and inexpensive assays are still in great demand. A recent study has shown that the clinical significance of CA-19-9 in PDAC patients with chronic pancreatitis is very low (Bedi et al., 2009) and it has a limited value in predicting early treatment in patients with advanced pancreatic cancer (Vormittag L, 2009).

1.2 HYPOTHESIS

Discovery of novel serum proteins by proteomic profiling in PDAC patients is only the initial step in the biomarker discovery pipeline. Mass Spectrometry (MS) analysis of patient samples is impractical on a large scale in clinical diagnosis and the specificity of this technique is very low especially when a complex clinical specimen such as serum is used. Therefore this method alone is not sufficient for the early detection of PDAC.
Potential biomarkers identified by MS analysis can be evaluated by readily available and highly sensitive assays to prove that these proteins can be used as reliable tumor markers. I hypothesize that this can be achieved by developing novel antibodies and sensitive sandwich Enzyme Linked Immunosorbent Assays (swELISA) that can detect these proteins in the serum, followed by retrospective evaluation of PDAC patient sera by swELISA. These assays could also serve as valuable tools in various epidemiological/clinical studies for the early detection of PDAC.

In order to attain diagnostic significance, every biomarker is expected to pass through various phases: 1) Target identification: This phase involves search for various targets at both mRNA and protein level in the cancer cell lines and clinical samples such as patient sera and tumor tissues. Extensive literature search of proteomic studies including MS analyses and genomic studies such as Microarrays are helpful in the identification of potential biomarker targets. 2) Assay development: Expression levels of the targets are verified in the cancer cell lines using various downstream assays such as ELISA, Western blot (WB), Immunohistochemistry (IHC) etc. 3) Assay validation: Evaluation of diagnostic value of the candidate markers in clinical samples including sera and tumor sections from cancer patients is performed using the same assays. During this phase, sensitivity and specificity of the assays is also verified. 4) Clinical evaluation: Once it is confirmed that the assay has diagnostic value, the targets are further validated by large scale epidemiological studies. 5) Biomarkers in clinics: Finally, the biomarkers are used in clinics for accurate diagnosis of the disease. The hypothesis
and focus of the current study is mainly based on Phases 2 and 3. Various phases of the biomarker discovery pipeline are illustrated in Figure 1.2.

**Figure 1.2: Biomarker discovery pipeline:** Initial step in the biomarker discovery is to identify potential biomarker targets by proteomic and genomic techniques. They need to be validated by various downstream assays. This is followed by clinical and epidemiological evaluation and finally used in the disease diagnosis.

### 1.3 SELECTION OF TARGETS FOR THE PDAC BIOMARKER STUDY

Harsha *et al.* have recently reported a comprehensive literature survey to systematically catalog overexpressed genes or proteins in PDAC to develop a list of potential biomarkers that could be validated by the pancreatic cancer research community (Harsha *et al.*, 2009). The approach was to identify the relevant publications and datasets (e.g. microarray data submitted to repositories such as GEO, Array Express and Oncomine) that contain information on over-expression of mRNAs or proteins in pancreatic cancer. Once a complete list of candidate molecules was generated, specific searches were carried out to identify the presence of these molecules in PDAC and then their status was determined in CP, which is an important
consideration in the differential diagnosis of pancreatic cancer. A brief overview of the selection strategy is shown in Figure 1.3 (modified from (Harsha et al., 2009)).

From a complete list of 60 potential protein markers generated by our collaborators as described above, our laboratory was allotted a task to develop swELISAs for a subset of 15 proteins. From this list, more stringent searches were performed in the NCBI journal database, PubMed and selected six targets for the current biomarker study. These targets include DKK1, CHI3L1, ITGB6, JUP, LAMB3 and LAD1. The selection was based mainly on the relative literature report of these targets in the PubMed. Since these targets were reported to be over-expressed in various tumors, they were expected to serve as reliable PDAC biomarkers. Targets were omitted from the list if no reports on their expression levels were available. This list, with related literature search, is shown in Table 1.1.
Figure-1.3: Selection of targets for the PDAC biomarker study. Extensive searches were carried out in PubMed and public repositories for potential biomarkers in PDAC. This data was compiled and the candidate markers were selected based on the over-expression at mRNA level (based on microarray and non-microarray techniques) and at protein level (based on MS and non-MS techniques).

Expression levels, both at the mRNA and protein level, for many of the targets selected for the current study, have been validated by investigators using various assays. However, the levels of these targets in human serum have not been tested so far, particularly in pancreatic cancer. This is due in part to the lack of effective antibodies and techniques such as swELISA that can detect the target in patient serum samples. Therefore, a sensitive swELISA is required to evaluate the expression levels of these targets in pancreatic cancer patients. Table 1.2 shows a brief summary that includes cellular localization, function and any available isoforms for each target.
Table 1.1: List of targets selected for the biomarker study and related literature

<table>
<thead>
<tr>
<th>Target</th>
<th>Available literature (categorized based on cancer type)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PDAC</td>
</tr>
<tr>
<td>CHI3L1</td>
<td>(Bi et al., 2009; Christina et al., 2002; Cintin C, 1999; Johansen JS, 2009)</td>
</tr>
<tr>
<td>DKK1</td>
<td>(Jun-Mei et al., 2010)</td>
</tr>
</tbody>
</table>
Table 1.1 continued

<table>
<thead>
<tr>
<th>Gene</th>
<th>References</th>
<th>References</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>JUP</td>
<td>(Karayiannakis AJ, 2001)</td>
<td>(Joo et al., 2002; Jose Luis et al., 2010; Nagel et al., 2010; Varis et al., 2002)</td>
<td>(Remmeli nk M, 2005)</td>
</tr>
<tr>
<td>LAMB3</td>
<td>(Logsdon et al., 2003)</td>
<td>(Isabelle et al., 2000; Mizushima et al., 1996)</td>
<td>(Carpenter PM, 2008)</td>
</tr>
<tr>
<td>LAD1</td>
<td>(Logsdon et al., 2003)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table-1.2: A brief summary of the targets selected for PDAC biomarker study

<table>
<thead>
<tr>
<th>Gene symbol</th>
<th>Protein name</th>
<th>Accession number</th>
<th>Potential function</th>
<th>Cellular localization</th>
<th>Isoforms, if any</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHI3L1</td>
<td>Chitinase 3 like 1</td>
<td>NM_001276</td>
<td>Role in the capacity of cells to respond to and cope with changes in their environment</td>
<td>Secreted, extra-cellular space</td>
<td>Not reported</td>
</tr>
<tr>
<td>DKK1</td>
<td>Dickkopf homolog 1 (Xenopus laevis)</td>
<td>NM_012242</td>
<td>Inhibitor of Wnt signaling pathway</td>
<td>Secreted</td>
<td>Not reported</td>
</tr>
<tr>
<td>ITGB6</td>
<td>Integrin beta 6</td>
<td>NM_000888</td>
<td>Receptor of fibronectin and cytoactin</td>
<td>Membrane-bound</td>
<td>Not reported</td>
</tr>
<tr>
<td>LAMB3</td>
<td>Laminin, beta 3</td>
<td>NM_000228</td>
<td>Mediate the attachment, migration and binding to cells via a high affinity receptor</td>
<td>Secreted, extra-cellular space</td>
<td>Not reported</td>
</tr>
<tr>
<td>LAD1</td>
<td>Ladinin 1</td>
<td>NM_005558</td>
<td>Anchoring filament protein which is a component of the basement membrane zone</td>
<td>Secreted, extra-cellular space, extra-cellular matrix, basement membrane</td>
<td>Not reported</td>
</tr>
</tbody>
</table>
1.4 OBJECTIVES OF THE STUDY

The major goal of the current study is to develop swELISA assays for potential biomarker targets and evaluate their expression in pancreatic cancer patients. To achieve this, the study was conducted with the following objectives:

1) Generation of screening reagents that are used to validate the antibodies and swELISA assays. These reagents include whole cell lysates over-expressing the target protein and lysates that express endogenous proteins.
2) Generation of at least two mouse monoclonal antibodies (mAbs) to different epitopes on each target or potentially one mAb if a commercially available antibody already exists followed by antibody validation.
3) Development of swELISA for each target using these antibodies followed by evaluation of specificity of the antibodies and ELISA by RNAi experiments.
4) Validation of the assays using serum samples followed by evaluation of expression of the targets by pilot epidemiological study in sera from patients with PDAC, pancreatitis and age and sex matched normal controls.

1.5 EXPERIMENTAL APPROACH AND STUDY LIMITATIONS

Once the targets were selected based on literature search, their expression levels need to be evaluated in PDAC serum samples. In order to perform this task, an assay that can detect the targets in serum samples is required. ELISA is a powerful tool used to detect soluble proteins in cell culture supernatants, whole cell lysates and various clinical specimens such as serum, urine and saliva. There are different types of ELISAs: Direct ELISA, Indirect ELISA, Competitive ELISA, Multiplex ELISA and
Sandwich or Capture ELISA. In a direct ELISA antigen or purified protein of interest is immobilized on a microtiter plate. The antigen is then detected using an enzyme-conjugated antibody. In an indirect ELISA, the immune complex (antigen-antibody complex) is detected using an enzyme-conjugated secondary antibody. In anti-sera or hybridoma culture supernatants antibody exists along with several other proteins. Hence it is not possible to conjugate the antibody Therefore, to detect the presence of an antibody-of-interest in such samples, indirect ELISA is used. This is one of the most widely used method for rapid screening on a large scale such as hybridoma fusion screen and subclone screen where hundreds of hybridoma culture supernatants are tested for a specific antibody. Therefore, this technique is used in the current study for screening anti-sera, fusions and hybridoma subclones. Competitive ELISA on the other hand is commonly used to evaluate the affinity of antibodies. Multiplex ELISA involves simultaneous detection of multiple analytes using secondary antibody conjugated to different enzyme labels or fluorochromes.

In a swELISA, free proteins or antigens in the solution are detected by two antibodies: 1. a capture antibody that is coated on to a stable surface and binds to the free antigen in the solution and 2. an enzyme-conjugated detection antibody that detects the protein bound to the capture antibody (Lequin, 2005; Van Weemen and Schuurs, 1971). The most important aspect of the swELISA is that each antibody should be able to recognize a distinct epitope on the target antigen with no cross-reactivity. This would eliminate any non-specific signal that might lead to misinterpretation of the data. One of the major advantages of this assay is that the proteins can be detected in
their native forms without denaturation. Whereas, other immune assays such as WB and Immunohistochemistry (IHC) techniques involve denaturation of the proteins that results in the loss of their secondary structure. Therefore swELISA is the technique of choice to detect proteins in PDAC sera in the current study. Based on clonality, antibodies are classified into two groups: monoclonal and polyclonal. Although polyclonal antibodies are relatively easy to generate, mAbs provide several advantages that include specificity of binding, homogeneity and ease with which they can be produced in unlimited quantities. Hence, a swELISA assay using mAbs is much more specific and sensitive than the same assay using polyclonal antibodies.

Extensive searches using the online database, Biocompare, revealed that pre-made ELISA kits were available commercially for DKK1 and CHI3L1. Since there was no evidence available on the expression levels of these targets in pancreatic cancer patients, these kits need to be validated on serum samples from PDAC patients. For ITGB6 and JUP, at least two mAbs were commercially available but no ELISA kits. Therefore, swELISA assays need to be developed for these two targets using the available antibodies. On the other hand, there were no effective mAbs available for LAD1 and there was only one mAb available for LAMB3. Hence mAbs need to be developed for these two targets (two mAbs for LAD1 and one mAb for LAMB3) in order to develop swELISA.

Generation of mAbs and development of swELISA assays require extensive validation to determine if they can specifically detect the protein of interest. Assay
validation is more efficient if the amount of target protein in the sample is in abundance. Since the biomarker targets used in this study are human, a human expression system such as Human Embryonic Kidney cell line, HEK 293T, would be ideal for the over-expression of targets. Whole cell lysates and tissue culture supernatants, derived from cell lines where the target-of-interest is over-expressed, are excellent sources of screening reagents. These reagents would not only serve as controls for the swELISA assay but also play a key role in the selection of hybridomas that produce target-specific antibodies. To over-express the biomarker target gene, the cDNA has to be cloned into a mammalian expression plasmid vector. Cloning by restriction digestion followed by ligation is the most widely used way of gene cloning. Recently, recombination cloning is gaining much importance as this process doesn’t involve the use of restriction enzymes. This not only speeds up the process but also enhances the efficiency of gene cloning. One of the most popular recombination techniques is Gateway Cloning system, a kind of Site-specific recombination cloning technology. This is a new, powerful methodology that greatly facilitates cloning of PCR products and protein expression. This technology replaces restriction endonucleases and ligases with a mechanism that is based on the site-specific recombination reactions of the bacteriophage λ in E. coli (Hartley et al., 2000; Walhout AJ, 2000). Therefore Gateway cloning technology was used in this study to generate the expression vectors containing biomarker genes that were over-expressed in HEK 293T cells.

For the generation of mAbs, antigens need to be generated for immunizing the host, preferably a rodent because they are easy to maintain. Synthetic peptides and
recombinant proteins are the most widely used antigens for mAb generation. In the current study, peptide antigens were used for immunizations since they are easy to generate and a specific epitope of the protein can be targeted. Hybridomas, formed by the fusion of splenocytes from immunized mice and mouse myeloma cells, produce target-specific antibodies which are purified by affinity chromatography technique.

Once the commercial or in-house generated antibodies are available, the next step would be to develop swELISA for each target. Validation of the swELISAs involves extensive evaluation of the assays on recombinant proteins (by testing the assays on whole cell lysates derived from 293T cells that were transiently transfected with target-specific expression vector) and endogenous proteins (by testing the assays on whole cell lysates derived from human PDAC cell lines). Finally, specificity of the assays needs to be verified using RNAi experiments where the target of interest is specifically down-regulated and the assays are evaluated for knock down of the target protein.

Differential diagnosis of PDAC from CP plays a key role in the early detection of PDAC. Therefore, all six targets used in this study need to be evaluated for differential expression in PDAC versus CP. Furthermore, verifying the expression levels of these targets in serum samples from normal healthy donors provides valuable information regarding their baseline expression in non-pathological conditions. An outline of the overall experimental design for the current study is shown in Figure 1.5.
Figure-1.4: Experimental design of the PDAC biomarker study: A list of six targets is generated based on the literature reports on their expression level and a recent study conducted by our collaborators. Targets were divided into three groups based on the availability of ELISA assays or antibodies. MAbs needed to be developed for LAMB3 and LAD1 as there were no effective antibodies available. ELISA assays were required for ITGB6 and JUP. Pre-validated ELISA assays were available for DKK1 and CHI3L1. Expression of these targets needs to be evaluated in serum samples from pancreatic cancer patients.
**Study limitations:** During the generation of screening reagents, epitope tags need to be inserted at the C-terminus of the plasmid. During this process, keeping the codons of all fragments in frame is critical. Verification of DNA sequence of the plasmids could overcome this limitation. Deletion of signal sequence is critical because secreted proteins lose N-terminal tag if the signal sequence is not deleted. This results in the non-expression of N-terminal tags which in turn makes it difficult to validate the swELISA. Signal sequence deletion for some relatively large genes such as LAMB3 could be challenging. During the antibody development process, selection of proper antibody that specifically works by swELISA is another challenge. Some antibodies that can recognize linear, denatured proteins might not detect soluble, non-denatured proteins due to the conformational changes between these two forms. Furthermore, wide range of samples such as tumor sections and PDAC cell lines would be required for extensive validation of the mAbs. Even though it doesn’t seem to be a major limitation, limited sample quantity is one of the most commonly encountered challenges in cohort studies. Therefore, the efficiency of any diagnostic assay depends on its sensitivity i.e. utilization of minimum sample quantities for the assay. Ability of the assay to detect the target at higher dilutions of the sample is very important. To address this challenge, different dilutions of the serum sample needs to be tested.