CHARPTER 2

LITERATURE REVIEWS

1. Shrimp diseases and defense mechanism

1.1 Penaeid shrimp diseases

Penaeid shrimp production is a worldwide economic activity primarily important for intertropical developing countries. Since its start in the 1970′s it has developed into a multi-billion of dollar industry. The intensification of shrimp farming over last few decades has been accompanied by development of infectious diseases from viral, bacterial, and in some cases fungal origin (Destoumieux-Grazon et al., 2001). Several outbreaks of shrimp diseases have caused major problem in countries such as China, Thailand, Indonesia, Taiwan and Ecuador (Garbriel and Filipe, 2000).

1.1.1 Bacterial diseases

The primary bacterial diseases of penaeid shrimp are vibriosis, filamentous bacterial disease, necrotizing hepatopancreatitis, chitinolytic bacterial shell disease, red disease, mycobacteriosis and rickettsial infection (Gabriel and Felipe, 2000). Both gram-positive and gram-negative bacteria, are etiological agents responsible for severe diseases in crustaceans, examples of infectious gram-positive bacteria include: *Aerococcus viridans*, and for gram-negative bacteria: *Vibrio harveyi*, *V. vulnifius*, *V. damsela* and *V. penaeicida* (Bachere, 2000; Lightner and Redman, 1998).

1.1.2 Fungal diseases

Fungi which may form potential pathogens for stressed or immunodeficient crustaceans is groups of *Lagenidium* sp, *Sirolpidium* sp and *Fusarium* sp (Bachere, 2000; Lightner and Redman, 1998). The main fungal diseases are larval mycosis and Fusarium disease. Larval mycosis caused by *Lagenidium* sp, and *Sirolpidium* sp, which are typical fungi attaching shrimp larvae results in fungal spores in gills and appendages, and high

shrimp larvae mortality (90%) in 2-3 days. Fusarium disease is a common disease that affects all developmental stages of penaeid shrimp in aquaculture. This disease generated from *Fusarium solani*, *F. oxysporum* and other *Fusarium* sp causes gill fouling (black gill), tissue lesions with melanization of appendages, fungal spores in gills and high mortality (Gabriel and Felipe, 2000).

1.1.3 Viral diseases

The various viruses of penaeid shrimp, White Spot Syndrome Virus (WSSV), Taura Syndrome Virus (TSV), Infectious Hypodermal and Haematopoietic Necrosis Virus (IHHNV), Yellow Head Virus (YHV), MBV (Monodon baculovirus) and HPV (Hepatopancreatic parvo-like virus) are responsible for major diseases problem in commercial shrimp ponds that have resulted in high mortality and economic losses. All of them are readily transmitted by ingestion of infected tissue from dead or moribund animals (Lightner, 1996; Lightner and Redman; 1995, 1981).

The prevention and control of diseases afflicting shrimp is now a high priority research topic in Thailand. The research currently focuses on three areas: genetics-development of diseases resistant shrimp broodstock, immunology and microbiology for development of effective shrimp probiotics (Rengpipat et al., 2000) and aid in the design of efficient strategies for disease control. In this study, we focus on the viral infection of shrimp, in particular, by the WSSV as well as the prevention and control of WSSV infection with respect to the three indicated areas of current research.

1.2 White spot syndrome virus (WSSV)

White spot syndrome virus is a pathogen of major economic importance in cultured penaeid shrimp, which was first discovered in Southeast Asia around 1992, is currently the most serious viral pathogen of shrimp worldwide. WSSV was reported in the United States in 1995 (Rosenberry, 1996) and from Central America and South America in early 1996 (Rosenberry, 2000). In 2002, WSSV was detected in Europe (France) and the Middle-East (Iran) (Rosenberry, 2002). It has spread rapidly to shrimp-farming areas all over the world and become one of the major pathogen in cultured shrimp. It causes up to 100% mortality within 3 to 10 days in commercial shrimp farms, resulting in large

economics losses to the shrimp farm industry (Lightner, 1996). It has abroad host range, infection all cultured shrimp as well as other invertebrate aquatic organisms such as crab and crayfish (Wang et al, 1998; Zhang et al., 2005). This virus has infected a number of other crustacean species such as amphipods, ostracods, swimming crabs, crayfish, and copepods. In the case of shrimp, WSSV infects and causes disease in many species of shrimp worldwide, including *P. monodon*, *P. semisulcatus*, *P. merguiensis*, *P. indicus*, *P. chinensis*, *P. penicillatus*, and *P. japonicus*. Clinical signs of the disease are a rapid reduction in food consumption, lethargy, a loose cuticle, a pink to reddish-brown discoloration and white spots on the carapace of the cephalothorax (Chou et al., 1995; Lightner, 1996). The virus severely damages the stomach, gills, subcuticular epithelial cells, lymphoid organ, antennal gland, and haemocyte (Chang et al., 1996; Lightner, 1996).

WSSV belongs to a new virus family, the *Nimaviridae*, and contains a large circular double stranded DNA genome size ranging from 292 to 305 kb in isolate from Thailand and China (van Hulten et al., 2001; Huang et al., 2005), but isolation of larger genomes have also been identified (Yang et al., 2001). The double-stranded viral DNA is 305 kb with 181 open reading frames (ORFs) (van Hulten et al., 2001; Yang et al., 2001. One Chinese WSSV genome isolate listed at Genbank (accession no. AF332093) has 305 kb of dsDNA genome (Yang et al., 2001), and another from *P. monodon* in Thailand has a 293 kb genome (Genbank accession no. AF369029) (van Hulten et al., 2001). Huang et al. (2002) were able to identify 18 proteins of WSSV by one-demensional SDS-PAGE followed by mass spectrometry and transcriptional analysis, and major virion protein genes (vp28, vp26, vp24, vp19, and vp15) have been identified (Marks et al., 2003). WSSV virions are ovoid to bacilliform in shape, approximately 250 nm in length and 120 nm in width, with tail-like appendage at one end of the virion (Wang et al., 2000; Wongteerasupaya et al., 1995). The viral nucleocapsid has a pattern of opaque and transparent striations arranged or composed of subunits in ring formation stacked in a series perpendicular to the long axis of the capsids (Figure 1) (Huang et al., 2005).

Because of the devastation effects on shrimp farms by WSSV various diagnostic methods with high efficiency, simplicity of use, and accuracy have been developed in many different laboratories throughout the world to help monitor and control the spread of WSSV. Several diagnostic methods have been described, such as polymerase Chain Reaction (PCR) (Kimura et al., 1996; Lo et al., 1996b; Nunan and Lightner,

1997; Kim et al., 1998; Tapay et al., 1999; Yoganandhan et al., 2003; Galaviz-Silva et al., 2004), in situ hybridization (Chang et al., 1996; Durand et al., 1996; Wongteerasupaya et al., 1996; Wang et al., 1998b), miniarray (Quere et al., 2002), observation of tissues subjected to fixation or negative staining (Inouye et al., 1993), detection of virus in the haemolymph or gastric epithelium by dark field microscopy (Momoyama et al., 1995) and immunological methods using monoclonal and polyclonal antibodies to WSSV or their component proteins (Huang et al., 1995; Nadala et al., 1997; Sahul-Hameed et al., 1998; Nadala and Loh, 2000; Poulos et al., 2001). Most recently, the new research technique, Reverse Passive Latex Agglutination (RPLA) is being developed to detect WSSV in the haemolymph of infected *P. japonicus*. The assay is sensitized with anti-WSSV-IgG and the WSSV agglutinated specifically by antigen-antibody reaction. However, in its initial trial the haemolymph formed nonspecific agglutination with a negative control reagent (sensitized with IgG refined from a nonimmunized rabbit) (Okumura et al., 2005). Furthermore, RT-PCR and western blot assays for vp19 have also been used to study the differential viral expression amongst different tissues of infected shrimp (Rout et al., 2005)

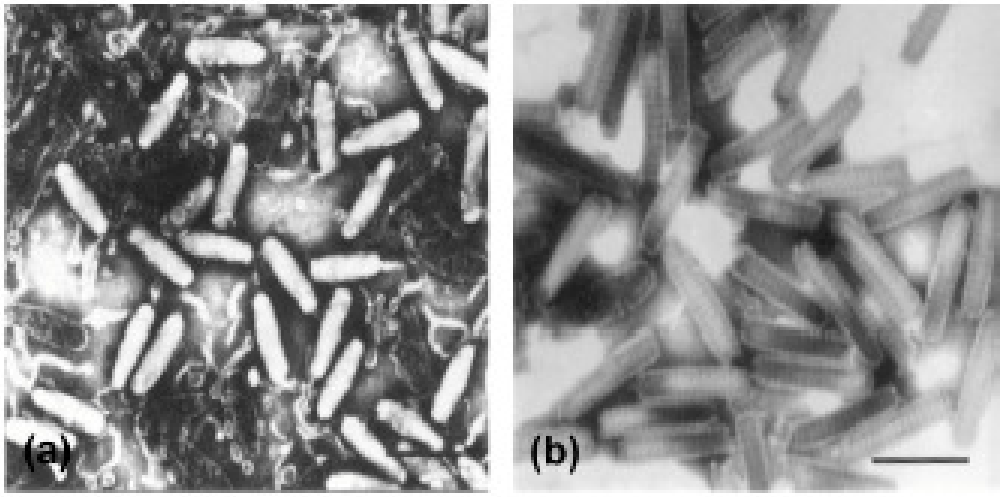

Figure 1. Transmission electron micrograph of White Spot Syndrome Virus (WSSV)

Transmission electron micrographs of purified WSSV virions (A) and nucleocapsids (B) from the haemolymph of WSSV- infected crayfish. Bars, 300 nm (Huang et al., 2005)

1.3 Shrimp immune response

Shrimp, like other crustacean have non-self recognition comprising the recognition of invading microorganisms (Vargas-Albores and Yepiz-Plascencia, 2000), the interaction of haemocytes with foreign molecules (Johansson, et al., 2000), and the activation of immediate defense systems (Srituyalucksana and Söderhäll, 2000). These processes may consist of both humoral and cellular response (Söderhäll and Cerenius, 1992). Crustacean haemocytes have different functions depending on haemocytes types (Sö derhäll and Cerenius, 1992). Based on the presence of cytoplasmic granules in the cells, crustacean haemocytes can be divided by light or electron microscopy into hyaline (the smallest haemocytes with minute cytoplasmic granules), semi-granulars (cell with small granules in cytoplasm) and granulars (the biggest haemocytes with large eosinophilic cytoplasmic granules) (Tsing, et al., 1989; Rodrigues, et al., 1995; Van de Braak, et al., 1996). Fifty to eighty percents of the circulating haemocytes are hyaline cells (Tsing, et al., 1989; Le Moullzc, et al., 1997; Sung, et al., 1999) involved in encapsulation, degranulation, and the prophenoloxidase (proPO) activating system (Johasson, et al., 2000). Between four to twenty five percents of the circulating haemocytes are granulars cells (Tsing, et al., 1989; Le Moullzc, et al., 1997; Sung, et al., 1999) participating in storage, cytotoxicity and release of proPO (Johasson, et al., 2000). Despite the use of similar criteria for haemocyte classification, different proportions of circulating haemocyte types have been reported, according to species, mounting stage and physiological condition of the animals (Tsing, et al., 1989; Le Moullac, et al., 1997; Sung, et al., 1999; Martin and Graves, 1985; Owens and O'Neill, 1997; Van de Braak, et al., 2000)

Several genes involved in immune response in *P. monodon* have been cloned and characterized, important examples include gene encoding the proteins including prophenoloxidase is an enzyme of the proPO system and is localized in the semigranular and granular cells (Sritunyalucksana et al., 1999). Peroxinectin is a multifunctional proteins involved in cell adhesion activity, degranulation activity, encapsulation-promoting activity, opsonic activity and peroxidase activity in immune responses (Sritunyalucksana et al., 2001). β-1, 3-glucan binding protein (BGBP) binds to β-glucans and the glucan-BGBP complexes inducing degranulation and activation of the proPO system (Sritunyalucksana et al., 2002). Syntenin is an adapter like protein involved in signaling pathway in response to WSSV infected shrimp (Bangrak et al, 2002). Furthermore, Somboonwiwat et al. (2005)

identified anti-lipopolysaccharide factor (ALF) from haemocyte of *P. monodon*. ALP has a function of anti-fungal properties against filamentous fungi, and anti-bacterial activities against both Gram-positive and Gram-negative bacteria especially *Vibrio* species including strains pathogenic for shrimp.

Recently, immune response molecules in other organisms have also been identified. Penaeidins are antimicrobial peptides constitutively produced and stored in the haemocyte of penaeid shrimp (Bachere et al., 2002). Lysozymes, an antibacterial protein has been found in white shrimp *P. vannamei* (Sotelo-Mundo et al., 2003). α_2M of kuruma shrimp (*Masupenaeus japonicus*) is mainly found in haemocyte and expression was found to be significantly induced by administration of the immunostimulant, peptidoglycan (PG) (Rattanachai et al., 2004), Thrombospondins (TSPs) from Chinese shrimp (*Fenneropenaeus chinensis*) was found to be upregulated in the hemocytes, heart, intestine and stomach of challenged shrimp with *Staphylococcus aureus* and *Vibrio anguillarum*. Additionally, TSP-like protein may be involved in the defense responses of shrimp (Sun et al., 2005). It has also been discovered that peritrophin-like protein obtained from Chinese shrimp (*Fenneropenaeus chinensis*) binds Gram-negative bacteria and has strong binding activity to chitin. Thus, it may play a role in defense system and other physiological responses (Du et al., 2006).

Interestingly, Witteveldt et al. (2004) developed two WSSV envelope proteins V19 and V28. They investigated the potential of oral vaccination of shrimp (*P. monodon*) using food pellet coated with inactivated bacteria over-expressing of V19 and V28. Only shrimp vaccinated with the V28 showed a significant lower cumulative mortality compare with the control. In addition, siRNA injection technique was developed as anti-viral using vp15 siRNA and vp28 siRNA by Westernberg et al. (2005). Shrimps injected with vp15 siRNA and vp28 siRNA before challenge gave a significantly lower mortality rate. These findings provide the new ways to stimulate immune response system of shrimp and as a result are benefit to the WSSV-hampered shrimp farming industry.

2. Using the expressed sequence tags library to identify immune gene

Expressed Sequence Tags (ESTs) generated from partial sequencing randomly selected cDNA clones have proven to be an effective approach for investigating gene expression in various texa where knowledge about the genome under investigation is not available or rather limited (Aaronson et al., 1996). This is a relatively simple and powerful technique for identification of useful genes in the particular application. Analysis of ESTs have been published for a number of fish species, amongst which are the Zebrafish, *Danio rerio* (Geisler et al., 1999) and Japanese flounder, *Paralichthys olivaceus* (Inoue et al., 1997).

The ESTs technique has also been employed for identification of immune gene in penaeid shrimp. In 1999, Lehnert et al. reported the EST sequence analysis from cephalothorax, eyestalk and pleopods tissues of *P. monodon*. Forty-nine newly isolated genes that had not been previously identified in crustaceans were reported. Several immune system genes have also been discovered in other litopenaeid shrimps. The ESTs results of the Pacific white shrimp (*Litopenaeus vannamei*) and the Atlantic white shrimp (*L. setiferus*) were determined from the haemocytes and hepatopancreas of a single individual from each species. The most common immune-function ESTs found corresponded to antimicrobial peptides, which were restricted to the haemocytes library (Gross et al., 2001)

P. japonicus (Rojtinakorn et al., 2002) and *P. monodon* (Supungul et al., 2002) ESTs were revealed in 2002. Gene expression in haemocytes of normal and WSSV infected Kuruma prawn (*P. japonicus*) were identified. The biodefense molecules were found from the library that involved in proPO system, basic protease inhibitor, antimicrobial peptides and tumor-related proteins (Table 1) (Rojtinakorn et al., 2002). On the other hand, the ESTs library of haemocyte of the Black tiger shrimp (*P. monodon*) was identified by Supungul et al. in 2002. The isolated genes were composed of coding enzymes and proteins involved in the clotting system and the prophenol oxidase-activating system, antioxidative enzymes, antimicrobial peptides and serine protease inhibitors (Table 1).

Other techniques, such as suppression subtractive hybridization and differential hybridization have been employed to study the differential profiles of genes expression in haemocytes of WSSV infected shrimp (*P. japonicus*). One study identified 30 genes as both involved in the anti-viral process as defense-relevant. Among them 22 were found for the first time in penaeid shrimp (He et al., 2005). To date, ESTs cDNA clones

have been applied with the suppression subtractive hybridization technique for gene expression profile analysis of the Pacific blue shrimp (*Litopenaeus stylirostris*) both infected and non-infected with *Vibrio penaeicida*. The results have identified genes directly involved in immune and hematopoietic processes from shrimp (de Lorgeril et al., 2005).

Recently, microarray technology has been added to the list of methods used to study differential expression profiles of genes in WSSV-infected shrimp (*Fenneropenaeus chinensis*) (Wang et al., 2006). The cDNA microarray using ESTs has proved to be a powerful tool to investigate the change in gene expression profiles of many genes simultaneously. However, this technique has been limited by a lack of cDNA microarray probes available for marine organisms (Chen et al., 2004).

Table 1. The putative biodefense peptides of EST in penaeid shrimps

Organism	Putative identification	Closet species	Data base accession
***P. japonicus*: non challenged (Rojtinnakorn et al., 2002)**			
	Prophenoloxidase	*Penaeus monodon*	AAD45201
	Maquerde-like protein	*Pacifastacus leniusculus*	CAA72032
	Alpha-2- macroglobulin	*Limus* sp.	T18544
	Clotting protein	*Penaeus monodon*	AF089867
	Hemocyte TGase	*Tachypleus tridentatus*	Q05187
	Lysozyme c type	*Mus musculus*	NP038618
	Penaeidin-2-precursor (P2)	*Penaeus vannamei*	GI3024356
	Translationally controlled tumor	*Homo sapiens*	NP004858
	Beta-integrin	*Pacifastacus leniusculus*	CAA67357
	Cell adhesion molecule (CAM)	Mus musculus	P02463
	Ubiquinol-cytochrome c reductase complex	*Bos taurus*	P13271
	Programed cell death-6-interacting	*Homo sapiens*	NP037506
***P. japonicus*: WSSV challenged (Rojtinnakorn et al., 2002)**			
	Coagulation factor	*Tachypleus tridentatus*	B49878
	Factor D	*Tachypleus tridentatus*	BAA13312
	Antileukoprotease	*Gallus gallus*	S36170
	Chelonianin	*Caretta caretta*	P00993
	Elastase inhibitor	*Aemonia sulcata*	P16895
	Kasal inhibitor type 1	*Pacifastacus leniusculus*	S45677
	Kasal inhibitor type 2	*Pacifastacus leniusculus*	S45677
	Kunitz-type inhibitor	*Ipera ammodytes*	P00991
	Hemocyte TGase	*Tachypleus tridentatus*	Q05187
	Bactinecin 11	*Ovis aries*	AAB62000
	Translationally controlled tumor	*Homo sapiens*	NP004858
	CD34+ (integral membrane protein 2A)	*Homo sapiens*	NP004858
	Drac-Ras-like GTP-binding protein	*Drosophila melanogaster*	BAA87881
	ESM-1 protein	*Homo sapiens*	CAB94771
	Neural killer cell-enhancing factor	*Trypanasomabrucei Rhodesiense*	Q26695
	Notch homolog	*Drosophila melanogaster*	NP036045
	Collagen alpha 3 (IV) precursor	*Mus musculus*	P02463
	Collagen alpha 4	*Homo sapiens*	NP000083
	Ubiquinol-cytochrome c reductase complex	*Bos taurus*	P13271
	Proteasome 26S subunit	*Homo sapiens*	NP002807

Table 1. Continued

Organism	Putative identification	Closet species	Data base accession
***P. japonicus*: WSSV challenged (Rojtinnakorn et al., 2002) (continued)**			
	Programed cell death-6-interacting	*Homo sapiens*	NP037506
	KE-3 (ribosomal protein S18)	*Homo sapiens*	P25232
	Rat insulinoma gene-Rig	*Homo sapiens*	NP001009
	Wilm's tumor-related protein-QM	*Bombyx mandarina*	GI4063389
***P. monodon*: no challenged (Supungul et al., 2002)**			
	Antilipopolysaccharide factor	Atlantic horseshoe crab	A23931
	11.5 kDa antibacterial protein	*Carcinus maenas*	CAB51030
	Protease inhibitor-signal crayfish	*Pacifastacus leniusculus*	S45677
	Penaedine-3c precursor	*Penaeus vannamei*	P81060
	P lysozyme structural	*Mus musculus*	NPO38618
	Hemocyte protein-glutamine γ -glutamyltransferase	*Tachypleus tridentitus*	Q05187
	Thymosin β-11	*Oncorrhynchus mykiss*	CAA21832
	Thymosin β-9 and β-8	Bovine	P21752
	Pro-phenoloxidase actvating enzyme I-precursor	*Holotrichia diomphalia*	BAA34642
	Protein kinase c inhibitor	Bovine	A35350
	Penaeid-2 precursor	*Penaeus vannamei*	P81057
	Prophenoloxidase	*Penaeus monodon*	AAD45201
	Glutathione peroxidase	*Homo sapiens*	A45207
	Peroxidase	*Aedes aegypti*	AAC97504
	Heat shock cognate protein 70 kDa	*Trichoplusia ni*	1495233
	Heat shock protein 90	Chicken	HHCH90
	Heat shock protein 70	*Hydra magripapillata*	Q05944
	Hemocyte protease-1	*Manduca sexta*	AAB94557
	Protein c	*Mus musculus*	NP032960
	Heat shock protein 10	*Gallus gallus*	AAB86581
***P. monodon*: *Vibrio harveyi* challenged (Supungul et al., 2002)**			
	P lysozyme structural	*Mus musculus*	NPO38618
	Penaedine-3k precursor	*Litopenaeus setiferus*	AAK83450
	Fc fragment of IgE	*Homo sapiens*	NP001993
	11.5 kDa antibacterial protein	*Carcinus maenas*	CAB51030
	Antilipolysaccharide factor	Atlantic horseshoe crab	A23931
	Heat shock-like protein (HSP70)	*Ceratitis capitata*	AAC23392

Table 1. continued

Organism	Putative identification	Closet species	Data base accession
***P. monodon*: *Vibrio harveyi* challenged (Supungul et al., 2002) (continued)**			
	Protease inhibitor-signal crayfish	*Pacifastacus leniusculus*	S45677
	Cytosolic manganese superoxidase dismutase precursor	*Callinectes sapidus*	AAF74771
	Prophenol oxidase activating factor	*Holotrichia diomphalia*	CAC12665
	Gene MAC25 protein (Kasal proteinase inhibitor homology)	*Homo sapiens*	152825
	Whey acidic protein (Putative protease inhibitor)	*Trichosurus vulpecula*	AAK69407
	Heat shock cognate 70 kD protein	*Oncorhynchus mykiss*	P0818
	Penaeidin-3c precursor	*Litopenaeus vannamei*	P81060
	Peptidyl-prolyl cis-trans isomerase 5	*Caenorhabditis elegans*	P52013

3. Comparative genomics analysis

Comparative genomics is the analysis and comparison of genomes from different species. Once a genome sequence is available, a primary goal is to identify functional regions in the sequences, including genes and regulatory sequences. Comparative genomics involves the use of computer programs that can line up multiple genomes and look for the regions of similarity among them. Some of these sequence-similarity tools are accessible to the public over the internet. There are a number of analysis programs available that are applicable to comparative sequencing analysis and genome sequence analysis in general. One of the most widely used programs is BLAST, which is available from the National Center of Biotechnology Information (NCBI). BLAST is a set of programs designed to perform similarity searches on all available sequence data in GenBank (International Human Genome Sequencing Consortium, 2001).

Comparative sequence analysis is becoming increasingly popular, and in recent years, a number of comparative studies of large genomic regions have been done, mainly between human and mouse (Hardison et al. 1997). Many functions of human genes have been determined by examining their counter parts in simpler model organisms such as the mouse. Genome researchers look at many different features when comparing genomes: sequencing similarity, gene location, the length and number of noncoding DNA in each genome, and highly conserved regions which may be maintained in organisms as simple as bacteria and as complex as humans. Model organisms offer a cost-effective way to follow the inheritance of genes through many generations in a relatively short time. Some model organism studies in the human genome project involved the bacterium *Escherichia coli*, yeast *Saccharomyces cerevisiae*, round worm *Caenorhabditis elegans*, fruit fly *Drosophila melanogaster* and laboratory mouse (International Human Genome Sequencing Consortium, 2001). The comparative genome sizes of humans and these model organisms being studies are shown in Table 2.

Thus, to begin studying the shrimp genome, comparative genomics is the best strategy to determine the functions of genes using data from other organisms such as *E. coli*, yeast, insect, crustacean, etc. Moreover, the application of comparative genomics with genome sequencing, protein-protein interaction, systemic localization, systemic deletion (RNA interference analysis), biochemical genomics and structural genomics will provide

useful information that may be applied to functional genomics of other organisms in future studies.

Table 2. The comparative genome sizes of human and other organisms

Organism	estimated sizes (million bases)	estimated gene number	chromosome number	References
Homo sapiens (Human)	2,900	~30,000	46	International Human Genome Sequencing Consortium, 2001
Rattus norvegicus (Rat)	2,750	~30,000	42	Rat Genome Sequencing Project Consortium, 2004
Mus muscular (mouse)	2,500	~30,000	40	Mouse Genome Sequencing Consortium, 2002
Drosophila melanogaster (fruit fly)	180	13,600	8	Adam et al., 2000
Arabidopsis thaliana (plant)	125	25,500	10	The Arabidopsis Genome Initiative, 2000
Caenorhabditis elegans (round worm)	97	19,100	12	The *C. elegans* Sequencing Consortium, 1998
Saccharomyces cerevisiae (yeast)	12	6,300	32	Goffeau et al., 1996
Escherichia coli (Bacteria)	4.7	3,200	1	Blattner et al., 1997

4. Shrimp syntenin (*Penaeus monodon*)

Previously, in order to determine the immune effectors in shrimp infected with WSSV, PCR-coupled subtractive hybridization was applied to cDNA from haemolymph of WSSV-infected shrimp (*Penaeus monodon*) and non-infected shrimp (Bangrak et al., 2002). One cDNA derived from up-regulated mRNA was identified. A homology search indicated similarity to the putative protein syntenin and it was name Pm-syntenin. Reports from studies of other organisms indicate that this protein is involved in signal transduction pathway. Therefore, Pm-syntenin may function as an adapter that couples PDZ-binding proteins in a signaling pathway involved in the shrimp defense response to WSSV infection. The Pm-syntenin was obtained from a subtractive cDNA library, it consisted of 969 nucleotides. It encodes 322 amino acid protein containing two PDZ domains in tandem. These two domains span amino acids 135-218 (Pm-PDZ1) and 219-299 (Pm-PDZ2), respectively. Sequence analysis of Pm-PDZ1 and Pm-PDZ2 using the BLAST showed the amino acid sequences had high similar to human syntenin (57% identity) and mouse syntenin (58% identity). Two putative postsynaptic density protein (PDZ) domains of Pm-syntenin were identified by an 83 residues stretch of amino acids of PDZ1 followed by another stretch of another 81 amino acids. The alignment of Pm-PDZ1 and Pm-PDZ2 with other known PDZ proteins is shown in Figure 2 (Bangrak, et al., 2002). Nevertheless, Pm-syntenin function in crustaceans is still unclear and further study needs to be carried out to identify its functions.

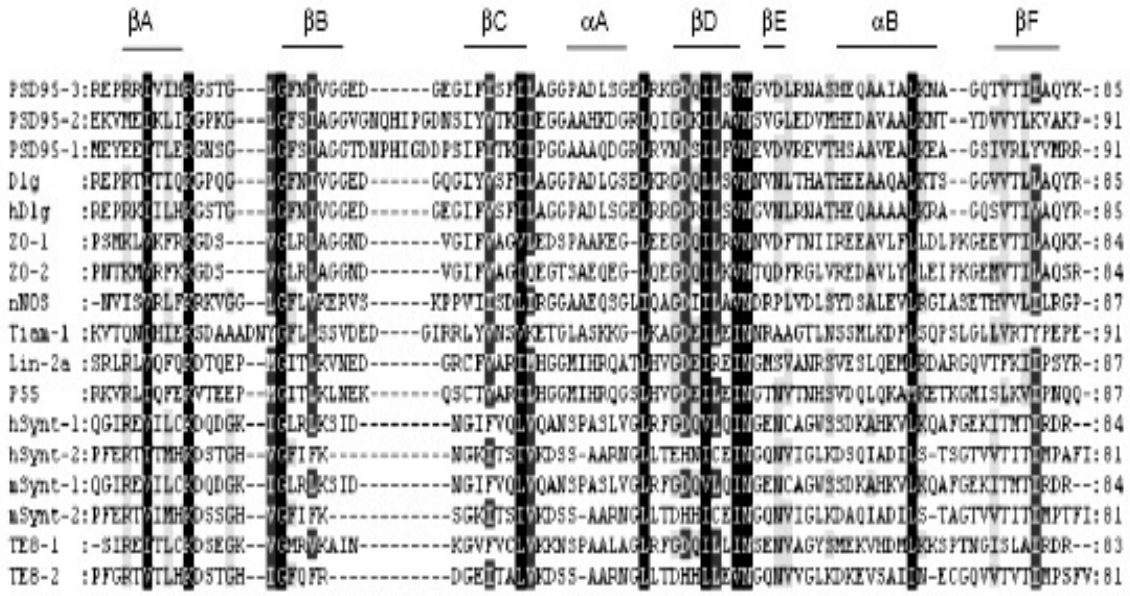

Figure 2. Sequence alignment of the Pm-PDZ1 and Pm-PDZ2 of Pm-syntenin with the PDZ homologs others organisms

The first PDZ (TE8-1)and second PDZ (TE8-2) domains of TE8 were aligned to the 3 PDZ domains of the postsynaptic density protein PSD-95 of rat brain (M96853), PDZ-3 of the *Drosophila* discs-large (Dlg) protein (M73529), human homolog hDlg (U49089), PDZ-2 of the human tight junction protein ZO-1 (L14837), PDZ-3 of human ZO-2 (L27152), the PDZ domain of rat neuronal nitric oxide synthase or nNOS (X59949), the PDZ domain of murine Tiam-1 (U05245), the PDZ domain of the *Caenorhabditis celegans* protein LIN-2a (X92564), the PDZ domain of human p55 erythrocyte membrane protein (M64925), the PDZ-1 (hSynt-1) and PDZ-2 (hSynt-2) domains of human syntenin (5032083), and the PDZ-1 (mSynt-1) and PDZ-2 (mSynt-2) domains of *Mus musculus* syntenin (AAC27646). The sequences were aligned by CLUSTAL X algorithm. The dashes represent gaps at the indicated proteins. Positions where the chemical character of residues is conserved in 100, 80 and 60% of sequences are highlighted in black, dark gray and light gray respectively (Bangrak, et al., 2001).

5. Human syntenin and syntenin binding proteins

Syntenin was originally identified as an adapter protein binding to the cytoplasmic domain of syndecans, which plays a major role in cell-to-cell and cell matrix interaction (Grootjan et al., 1997). Syntenin has no obviously catalytic domain and therefore is unlikely to have a signaling function by itself, but could serve as an adaptor or scaffolding protein to attach other protein. As a member of the PDZ protein family, syntenin functions in the cell. Adapter molecules function as intracellular scaffolds around which signaling complexes are formed (Pawson and Scott, 1997; Norian and Koretzky, 2000; Yaffe and Elia, 2001). The integration or connection between initial signaling molecules and the down stream second messengers are facilitate by group of proteins that mediates interaction with other proteins through their specific domains. These adapter molecules possess no enzymatic or transcriptional activity but rather act solely to mediate protein-protein interaction through a variety of functional domains thus coordinating interactions among the signal transduction effecter molecules (Pawson and Scott, 1997; Norian and Koretzky, 2000). Based on their interacting domains adapter molecules can be grouped into the following categories with syntenin, a PDZ protein being classified into group (4).

(1) *src* homology 2 (SH2) domains that bind phosphotyrosine (pTyr) residues followed by three to five COOH-terminal residues, such as those generated by autophosphorylation of activated receptor tyrosine kinases (Songyang et al., 1993; Waksman et al., 1993; Pascal et al., 1994).

(2) *src* homology 3 (SH3) domains that bind to polyproline motifs with the consensus PXXP domain that form a left-handed polyproline type II helix (Feng et al., 1994).

(3) pTyr-binding (PTB) domains that recognize phosphopeptide motifs in which pTyr is preceded by residues that form a β-turn, usually with the consensus NPXpY domain (Kavanaugh et al., 1995; van der Geer et al., 1995).

(4) PDZ domains that usually recognize peptides with a COOH-terminal hydrophobic residue and a free carboxylate group such as an E(S/T)DV motif (Kim et al., 1995; Songyang et al., 1997).

(5) WW domains that bind to proline-rich motifs, commonly with the consensus PPXY or PPLP domains (Sudol, 1996; Macias et al., 1996).

(6) pleckstrin homology (PH) domains that bind to phospholipids and thus regulate the targeting of proteins to the plasma membrane (Harlan et al., 1994).

5.1 Structure of human syntenin

The human cDNA of syntenin is ~ 2.1 kb with the open reading frame of 894 bp that codes for a protein of 298 amino acid residues with the predicted molecular mass of ~33 kDa (Lin et al., 1996; Grootjan et al., 1997). The molecule has four domains: an N-terminal domain (aa 1-109) that shows no striking homology to any structural motifs, the first PDZ (PDZ1; aa 110-193), the second PDZ (PDZ2; aa 194-274), and a COOH-terminal domain (Grootjan et al, 1997). The term PDZ is an acronym that represents three proteins; **P**ostsynaptic density protein PSD95/SAP90, **D**rosophila tumor suppressor DLGA, and tight junction protein **Z**O-1, in which this conserved sequence element was first identified (Harris and Lim, 2001; Hung and Sheng, 2002).

PDZ domains are ubiquitous signaling domains with over 400 distinct copies in the human genome (Sheng and Sala, 2001). This domain is composed of 80-90 amino acids with a distinct fold of six β-strands and two α-helices (Figure 3) and may occur in proteins harboring other anchoring domains, but is also found in proteins that contain no other domains (Doyle et al., 1996). Multiple PDZ domains might be present in a single protein. An extreme example is MUPP, a scaffolding protein containing 13 PDZ domains (Mancini et al., 2000). PDZ domain containing proteins have broad species distribution and can be found in plants, prokaryotes, and eukaryotes (Ponting et al., 1997). The majority of the PDZ domain containing proteins are associated with the plasma membrane. They are generally restricted to specific subcellular domains such as synapses, cell-cell contacts or the apical, basal, or lateral cell surface, thereby supporting a potential role as central organizers of protein complexes at the plasma membrane (Fanning & Anderson, 1999). In general, PDZ domains bind preferentially to peptides that terminate in a hydrophobic amino acid, usually valine or isoleucine, and with either serine, threonine, or tyrosine located two residues from the COOH-terminus (-2 position). PDZ domains are typically grouped into three classes depending on their target peptides: class I (X-S/T-X-Ø), class II (X-Ø -X- Ø), and class III (X-D-X-V), where Ø is a hydrophobic residue and X is any amino acid residue (Table 3) (Sheng and Sala, 2001). However, this simplified model is not adequate and an increasing number of PDZ-domain interactions that

do not conform to the canonical type of recognition are being identified and one PDZ domain may recognize more than one class of C-terminal sequence motifs (Bezprozvanny and Maximov, 2001)

The crystal structure of syntenin has been elucidated and contains two PDZ domains conjoined by a short linker (Kang et al., 2003a, 2004a, 2004b). The PDZ domains show a typical fold with two opposing anti-parallel β-sheet capped by two α-helices as shown in Figure 3 (Huang and Sheng, 2002). Each domain has at least one β-strand is partly contained in both sheets. In the crystal form, the two PDZ domains are arranged in a head-to-tail fashion. Despite the fact that the two PDZ domains have only 26% amino acid identity, they are structurally very similar to each other (Kang et al., 2003b). In both domains, the fragment equivalent to the signature GLGF loop, involved in the terminal carboxylate binding (Kang et al., 2003b).

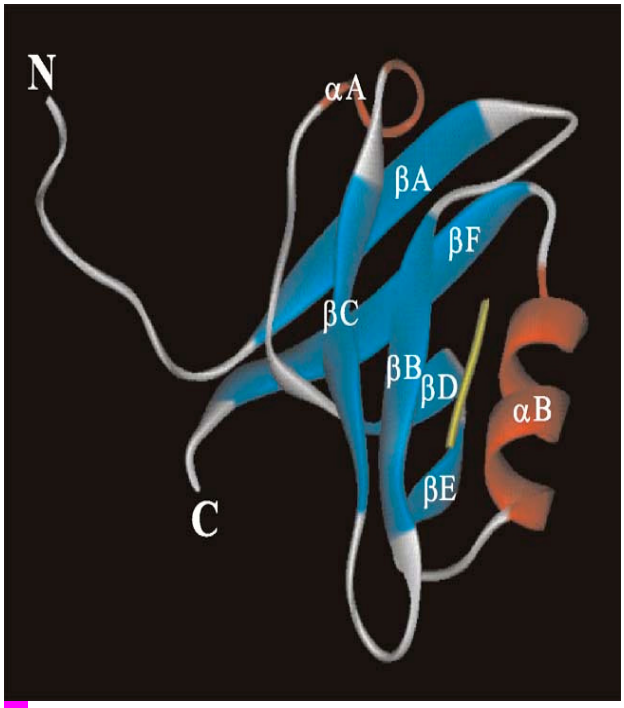

Figure 3. Structure of PDZ-like domain

The structure demonstrates the six β-strands (turquoise) and two α-helices (red) with the peptide (yellow) binding as a β-strand between the αB helix and βB strands. The N- and C-termini are labeled (Huang and Sheng, 2002).

Table 3. Classification of the PDZ domains according to specificity for C-terminal peptide (Sheng and Sala, 2001)

Class	C-terminal Sequences	Interacting protein	PDZ domain	References
Class I				
X-S/T-X-V	E-S-D-V	NMDAR2A, B	PSD-95 (PDZ2)	Kornau et al., 1995
	E-T-D-V	Shaker channel		Kim et al., 1995
	Q-S-S-V	Citron	PSD95 (PDZ3)	Zhang et al., 1999
	Q-T-S-V	CRIPT		Niethammer et al., 1998
	T-T-R-V	Neurogilin		Irie et al., 1997
	E-T-S-V	PMCA4b	PSD-95(PDZ1/2/3)	Kim et al., 1998
	E-S-L-V	Voltage-gated Sodium channel	Syntrophin	Gee et al., 1998
X-S/T-X-L	Q-T-R-L	GKAP	Shank	Naisbitt et al., 1999
	S-S-T-L	mGluR5		Tu et al., 1999
	D-S-S-L	β2-adrenergic Receptor	NHERF (PDZ1)	Hall et al., 1998
	P-T-R-L	GRK6A		Hall et al., 1999
	D-T-R-L	CFTR		Wang et al., 1998
Class II				
X-Ø-X-Ø	E-Y-Y-V	Neuroxin	CASK	Hata et al., 1996
	E-F-Y-A	Syndecan	CASK, Syntenin	Hsueh et al., 1997
	E-Y-F-I	Glycophorin C	p55	Grootjans et al., 1997; Marfatia et al., 1997
	S-V-K-I	GluR2	GRIP (PDZ5)	Dong et al., 1997
			PICK-1	Xia et al., 1999
	S-V-E-V	EphB2	GRIP (PDZ6), PICK-1	Torres et al., 1998
	G-I-Q-V	EphA7	GRIP (PDZ6), PICK-1, Syntenin	
	Y-Y-K-V	EhrinB1	GRIP (PDZ6), PICK-1, Syntenin	
Class III				
X-D-X-V	V-D-S-V	Melatonin recepter	nNOS	Stricker et al., 1997

5.2. Syntenin binding proteins and functions

Most molecules interacting with syntenin have been identified by yeast two hybrid assays. A list of these molecules and their potential functional significance are enumerated in Table 4. At this point we will review theses molecules and their interactions with syntenin in greater detail.

5.2.1 Syndecan

The syndecans are the transmembrane proteoglycan that place structurally heterogenous heparin sulfate chains at the cell surface and a highly conserved polypeptide in the cytoplasm. Syntenin was identified as a syndecan bindind protein by use of the yeast two hybrid screening system, surface Plasmon resonance experiments and ligand-overlay assays. Syntenin interacts with the C-terminal FYA sequence of syndecans. All four syndecans and syntenin colocalized to the plasma membrane. Overexpression of syntenin resulted in large, flat, vesiculated cell with numerous membrane extensions of the plasma membrane. However, if the N-terminal of syntenin is deleted, only PDZ domains containing the mutation showed increased plasma membrane modifications in comparison with the wild type molecules, thus indicating that the only PDZ domains of syntenin have an inhibitory effect on the plasma membrane modifications (Zimmermann, et al. 2001). This implied a potential role for syntenin in the positioning of syndecan-associated receptor and involved in cytoskeleton-membrane organization (Grootjan et al., 1997; Zimmermann, et al. 2001).

5.2.2 Pro-TGFα

Transforming growth factor subunit α (TNFα) is a ligand for epidermal growth factor receptor and is synthesized as a part of precursor, transmembrane and cell surface molecule (Derynck et al., 1984; Lee et al., 1985). Utilization of the yeast two-hybrid assay with the C-terminal of pro-TGFα as bait identified two PDZ domain containing proteins as a pro-TGFα interacting proteins, one of which was syntenin (Fernandez-Larrea et al., 1999). Both PDZ domains of syntenin are required for this interaction. Syntenin colocalized with pro-TGFα in the early stages of the secretory

pathway such as within the endoplasmic reticulum, intermediate compartment and golgi. Suggesting that syntenin might function in regulating the correct localization of transmembrane receptors (Fernandez-Larrea et al., 1999).

5.2.3 Ephrin B

Ephrin B protein function as ligands for B class Eph receptor tyrosine kinase and postulated possess an intrinsic signaling function (Lin et al., 1999). Yeast two-hybrid assays have identified ephrin B as interacting with protein kinase C (PICK1), Glutamate receptor-interacting protein (GRIP), syntenin and EphA7 (Torres et al., 1998). These finding indicated that the carboxy terminal motif of ephrin B provides a binding site for specific PDZ domain-containing proteins, which might localize the transmembrane ligands for interaction with Eph receptor or participate in signaling within ephrin B expressing cell (Torres et al., 1998; Lin et al., 1999).

5.2.4 Interleukin 5 receptor α subunit (IL-5Rα)

Cytokine receptors consist of multiple subunits, which are often shared between different receptors, resulting in the functional redundancy sometimes observed between cytokines. The interleukin 5 receptor (IL-5R) consists of an IL-5-specific α-subunit (IL-5Rα) and a signal-transducing β-subunit (βc) shared with IL-3 and granulocyte colony-stimulating factor (GM-CSF) receptors. The IL-5Rα is expressed on B cell and on eosinophilic and basophilic granulocytes (Travernier et al., 1991; Murata et al., 1992; Hitoshi et al., 1990). Using yeast two-hybrid screening from human granulocyte cDNA library, syntenin was identified as IL-5Rα binding protein (Geijsen et al., 2001). Moreover, syntenin was found to directly associate with the transcription factor Sox4. The association of syntenin with IL-5Rα was required for IL-5-meadiated activation of Sox4. These findings imply that is syntenin involved in a mechanism of transcriptional activation by cytokine-specific receptor subunit (Geijsen et al., 2001).

5.2.5 Neurofascin

Neurofascin is a member of the L1 subgroup of the immunoglobulin superfamily of cell adhesion molecules and is implicated in axonal growth and fasciculation. Syntenin has been found as a neurofascin binding protein by yeast two-hybrid and ligand overlay assay. This interaction promotes neuronal migration, axonal growth and fasciculation in the developing nervous system (Koroll et al., 2001).

5.2.6 Other syntenin binding proteins

There are several more syntenin binding proteins that had been found and characterized in Table 4 (Sarkar, et al., 2004). In humans, syntenin function is involved in signaling transduction in neural cells and plasma membrane localization. Recently, new syntenin binding proteins were found to include; the glycine transporter subtype 2, eukaryotic translation initiation factor 5A, Lymphocyte receptor CD6.

The glycine transporter subtype 2 (GlyT2) is localized at glycinergic axon terminal where it mediates the re-uptake of glycine from extracellular space. Thus syntenin may be an in vivo binding partner of GlyT2 that regulates its trafficking and/or presynaptic localization in glycinergic neuron (Ohno et al., 2004). The eukaryotic translation initiation factor 5A (eIF5A) was originally designed as an initiation factor of protein translation. It has been found that syntenin interacts with eIF5A both in vitro and in vivo and functions collaboratively to regulate p53 activity. Therefore, the new finding suggests syntenin might regulate p53 by balancing the regulation of eIF5A signaling to p53 for apoptosis (Li et al., 2004). Lymphocyte receptor CD6 was identified as interacting with syntenin through use of the yeast two-hybrid, co-immunoprecipitation and GST-pull down assays (Gimferrer et al., 2005). Therefore, syntenin may function as a scaffolding protein coupling CD6 and most likely other lymphocyte receptors to skeleton and/or signaling factor (Gimferrer et al., 2005).

Table 4. Molecules interacting with syntenin

(Sarkar, et al., 2004)

Interaction molecule	Type of of interaction	Binding sequence	Possible functions
IL-5Rα	Type I	-D-S-V-F	Sox4 activation; B cell development and Differentiation
ProTGFα	Type I	-E-T-V-V	Intracellular trafficking
Neurofascin	Type I	-Y-S-L-A	Synaptic transmission
Neuroglian	Type I	-Y-S-L-A	Not identified
Syndecans	Type II	-E-F-Y-A	Cell polarization; Plasma membrane integrity
Ephrin-B	Type II	-Y-Y-K-V	Eph receptor signaling
EphA7	Type II	-G-I-Q-V	Eph receptor signaling
r-PTPη	Type II	-G-Y-I-A	Plasma membrane localization
Neurexin-I	Type II	-E-Y-Y-V	Not identified
Unc51.1	Type II	-V-Y-A	Axon formation
Rab5 GTPase		Not identified	Axon formation
Schwannomin	Type III	-F-E-E-L	Plasma membrane localization
Ionotropic and metabotropic glutamate receptors	Type I/II/III	Variable	Glutamate signaling
PIP_2			Plasma membrane anchoring

6. Protein-protein interaction technologies

Protein-protein interactions have attracted much attention because they form the basis of a wide variety of biochemical reactions. The identification of proteins that interact with the know protein is an essential aspect of the elucidation of the regulation and function of that protein. The result of two or more proteins interacting with a specific functional objective can be demonstrated in several different ways. The measurable effects of protein interactions have been outlined by Phizicky and Fields (1995). Protein interactions can:

- Alter the kinetic properties of enzymes. This may be the result of subtle changes at the level of substrate binding or at the level of an allosteric effect.

- Allow for substrate channeling by moving a substrate between or among subunits, resulting ultimately in an intended end product.

- Create a new binding site, typically for small effecter molecules.

- Inactivate or destroy a protein.

- Change the specificity of a protein for its substrate through interaction with different binding partners; e.g., demonstrate a new function that neither protein can exhibit alone.

- Serve a regulatory role in either an upstream or a downstream action

Interest has stimulated the development of the number of biochemical and genetics approach to identify and clone genes encoding interacting proteins, this includes; yeast two-hybrid system, co-immunoprecipitation, GST-pull down assay, surface plasmon resonance and direct expression library screening using protein as probes etc. However, the development of the yeast two-hybrid system and GST-pull down appear to have the greatest impact on the interaction methodology.

6.1 Yeast two-hybrid system

The yeast two-hybrid system was devised to identify gene encoding protein that physically associate with a given protein in vivo. This is a versatile and powerful method that is applicable to most, if not all, proteins once their genes have been isolated. In contrast to biochemical methods detecting protein-protein interaction, this system is based on yeast genetic assay in which the interaction of two proteins is measured by reconstitution of functional transcription activator in yeast (Field and Song, 1989; Chien, et al., 1991). This method not only allows identification of proteins that interact, but also can be used to define and/or test the domains/residues necessary for the interaction of two proteins (Li and Field, 1993). Since its development, a large number of genes from a variety of studies have been identified using this method, including many cell cycle regulators that have contributed significantly to our understanding of the eukaryotic cell cycle (Harper et al., 1993; Toyoshima and Hunter, 1994; Bai et al., 1996).

The basis of the yeast two-hybrid system relies on the structure of particular transcription factors that have two physically separable domains: a DNA-binding domain and transcription activator domain (Figure 4). The DNA binding domain serves to target the transcription factor to specific promoter sequence (designated UAS for upstream activation domain sequence) whereas the activating domain serves to facilitate assembly of the transcription complex allowing the initiation of transcription. The fact that a functional transcription factor can be reconstituted through non-covalent interaction of two independent hybrid proteins containing either a DNA-binding domain or an activation domain constitutes the basis of the two-hybrid approach (Fields and Song, 1989). The hybrid proteins are normally transcriptionally inactive alone or when co-expressed with a non-interaction hybrid protein. However, when co-expressed they associate via interaction between the two fusion protein partners. Their association results in activation and thus expression of the reporter gene driven by the specific UAS for the DNA-binding domain (Fields and Song, 1989). A schematic representation of the two-hybrid system's genetic selection by activation of transcription of *HIS3* (encoding histidine), *ADE2* (encoding adenine), *MEL1* (encoding α-galactosidase) and *lacZ* (encoding β-galactosidase) genes is shown in Figure 5.

High throughput yeast two-hybrid systems is the system using for large scale protein interaction analysis, which can enable both rapid study of many different

proteins and the mapping of complex interaction networks (Buckholz et al., 1999; Uetz et al., 2000; Ito et al., 2001). Interaction-mating-based yeast two-hybrid systems have also been used to study protein interactions on a proteomic scale. In 2000 two groups published the first large-scale interaction maps for an eukaryotic organism using yeast two-hybrid interaction mating. It is fitting that the proteome studied was derived from the yeast *Saccharomyces* itself (Uetz et al., 2000; Ito et al., 2001). More recently, the system has been used for the rapid identification of the interactions of greater than 4,000 *C. elegans* proteins (Giot et al., 2003) and of >10,000 *Drosophila* proteins (Li et al., 2004). Such highly efficient, high-throughput approaches now make large-scale protein interaction mapping a practical proposition for any organism, including humans.

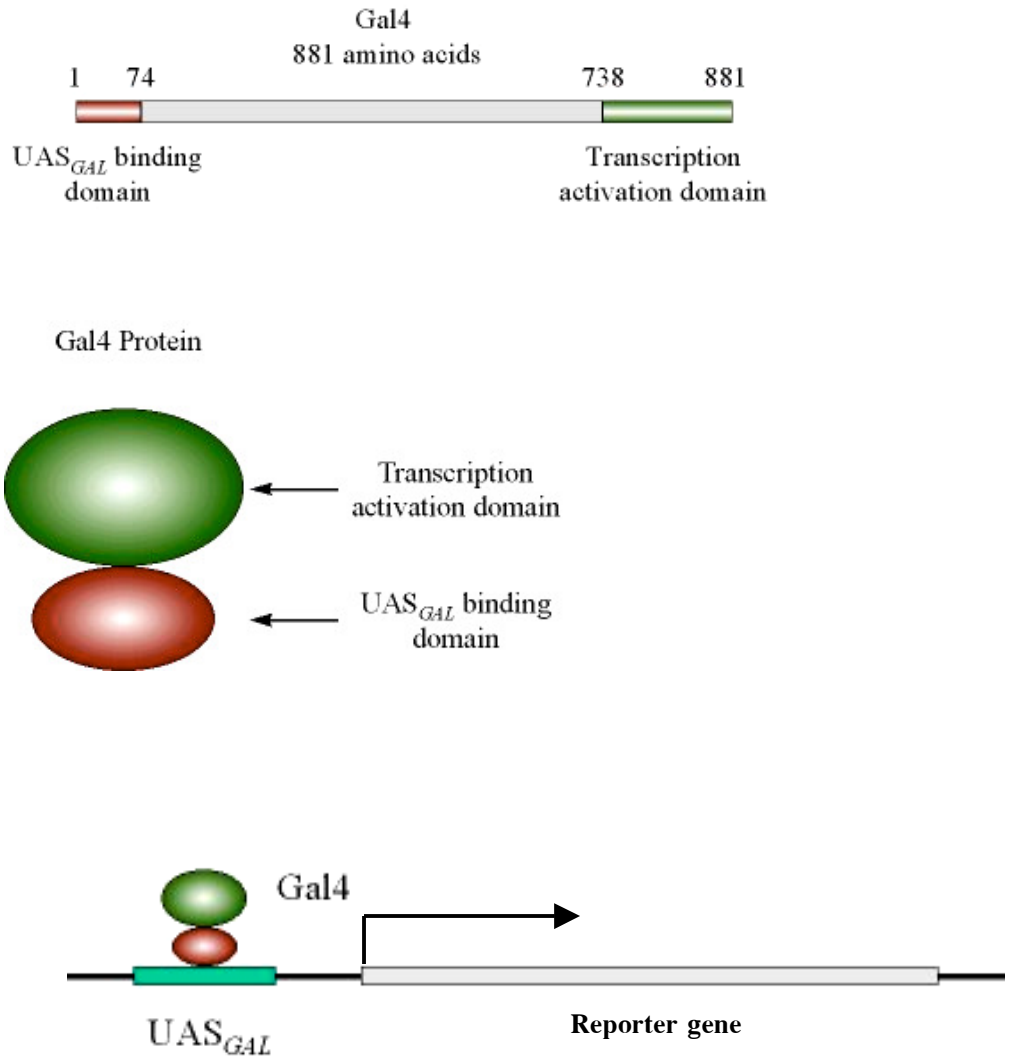

Figure 4. A schematic representation of the Gal4 protein

The Gal4 protein is a transcriptional activator protein which specifically activates the transcription of genes involved in the galactose metabolic pathway. Gal4 protein binds to a specific 17 nucleotide long region of DNA located in the 5′ region of gene involved in galactose metabolism. These region termed upstream activator sequences (UAS$_{GAL}$). The binding of Gal4 protein and increases UAS$_{GAL}$ the transcription of the reporter gene (Fields and Ok-kyu, 1989).

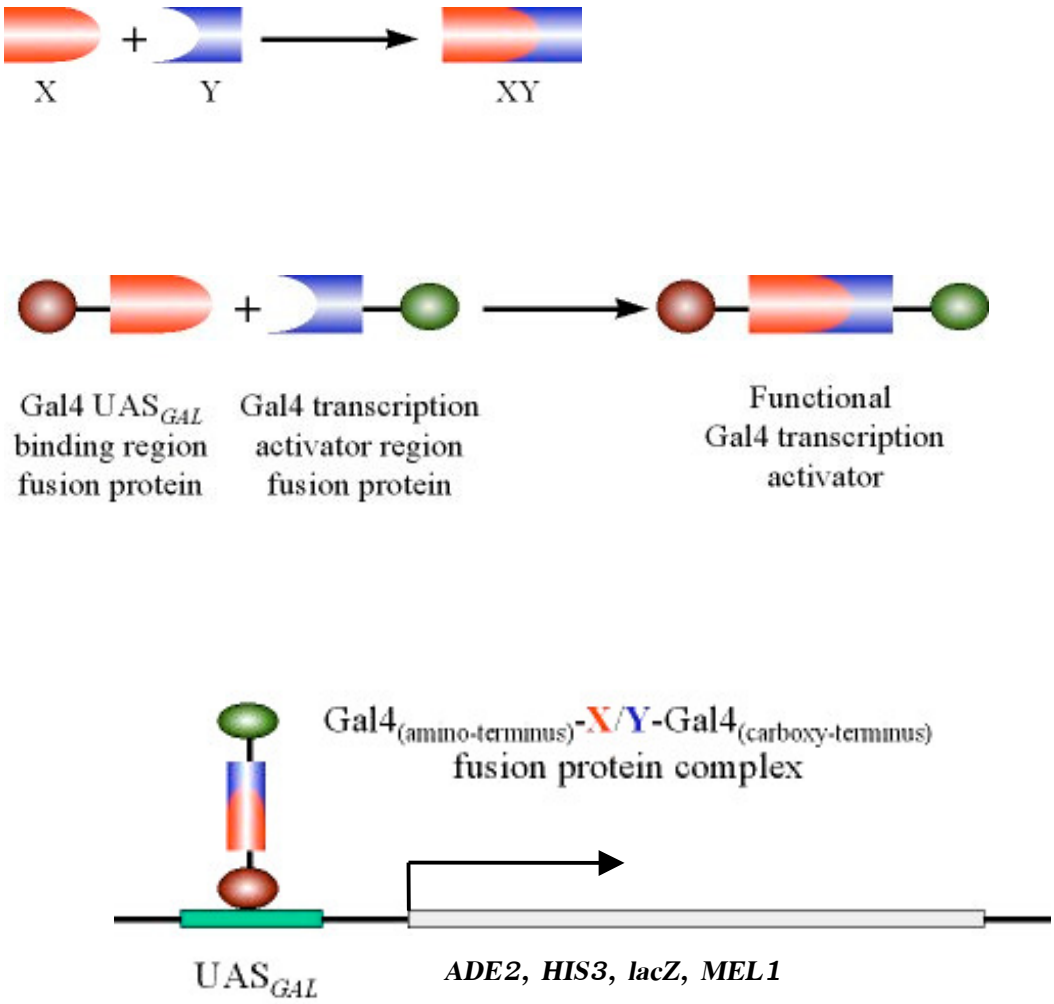

Figure 5. A schematic of the yeast two-hybrid system

Two-hybrid proteins are generated. The bait hybrid consists of protein X fused to a DNA-binding domain while the prey hybrid consists of protein Y fused to an activation domain. Neither of these alone is alone is able to activate transcription of reporter genes. However, when interaction of protein X and Y occurs, a functional transcription activator is generated and results in the transcription of reporter genes (Fields and Ok-kyu, 1989).

6.2 GST (Glutathione-s-transferase) pull-down assay

The GST fusion protein pull-down technique uses the affinity of GST for glutathione-coupled beads to purify interacting proteins from a solution of non-interaction proteins (Kaelin et al., 1991). The GST-fusion probe protein and cell lysate are mixed in the presence of glutathione agarose beads and incubated to allow protein associations to occur. The GST fusion protein and any associated molecules are collected by centrifugation and complexes are washed. The complexes can be eluted from the beads with excess free glutathione or boiled directly in SDS-PAGE gel loading buffer. The proteins are resolved by SDS-PAGE and processed for further analysis by western blotting, autoradiography, or protein staining (Figure 6). The GST pulldown technique is especially useful for probing protein interactions in solutions that might go undetected in a membrane-based assay. This method of detecting interactions is determined by the availability of antibodies to the target protein. Moreover, ^{35}S-labeled in vitro translated protein can additionally be utilized or the target protein can be tagged with an epitope (Sambrook and Russell et al., 1999).

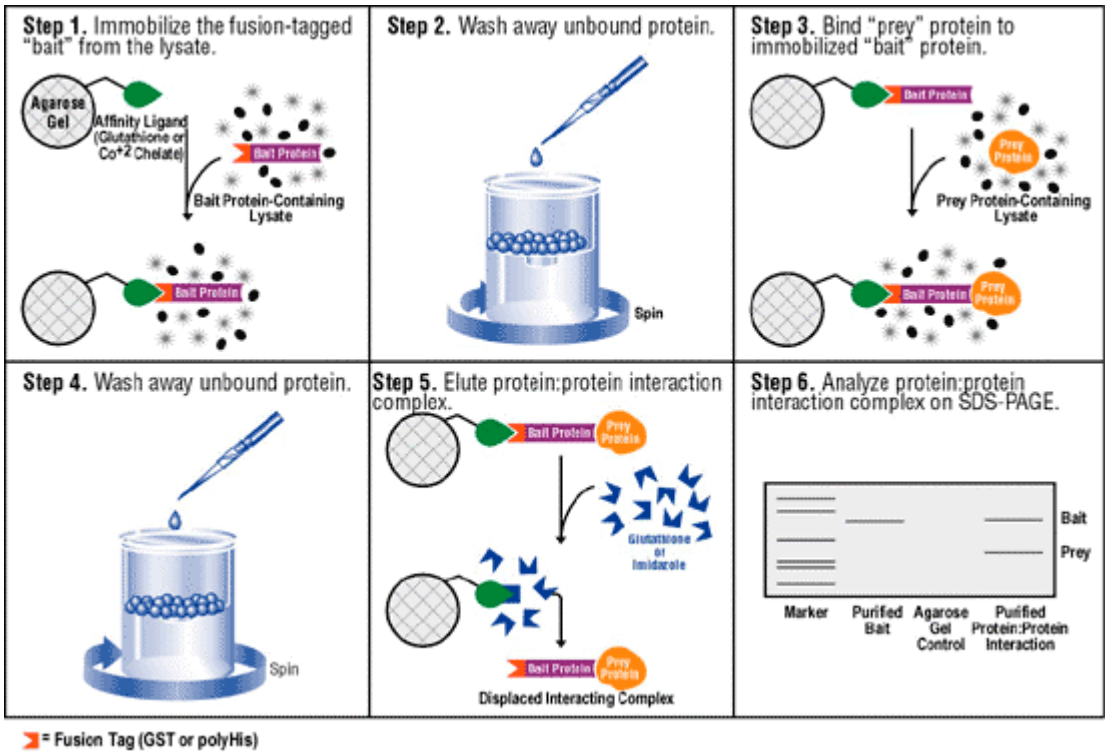

Figure 6. A schematic of GST pull-down assay (Kaelin et al., 1991)

6.3 Co-immunoprecipitation (Co-IP)

Co-immunoprecipitation is a powerful method used to study protein-protein interactions. The principle of an IP is very simple. An antibody (monoclonal or polyclonal) against a specific target antigen is allowed to form an immune complex with that target in a sample, such as a cell lysate. The immune complex is then captured on a solid support to which either protein A or protein G has been immobilized (protein A or G binds to the antibody, which is bound to its antigen). The process of capturing this complex from the solution is referred to as precipitation. Any proteins not "precipitated" by the immobilized protein A or G support are washed away. Finally, components of the bound immune complex (both antigen and antibody) are eluted from the support and analyzed by SDS-PAGE (gel electrophoresis), followed by western blot detection to verify the identity of the antigen. The use of co-immunoprecipitation to search for novel proteins that interact with a known protein is a powerful way to identify physiological interactions within the intact cells (McMahon et al., 1998, Lonergan et al., 1998).

6.4 Surface plasmon resonance (SPR)

Surface Plasmon resonance has simplified the study of macromolecule interaction by providing a format that may be used to measure molecular interactions in real time with small analytical amounts of material. The strength of SPR spectroscopy is the ability to obtain kinetic data; the microscopic rate constants for an interaction between macromolecules. In this approach one protein to be analyzed is secured to flat sensor chip in a flow chamber, after which a solution containing a prospective interacting partner across the chip and the resonance angle of reflected light is measured. This technique has been used for protein-protein, protein-small molecule, protein-nucleic acid and protein- lipid interactions (Celia et al., 1999; Raut et al., 1999)