1. INTRODUCTION

Lactational function of mammary gland occurs in cycles, beginning with mammary gland development, lactogenesis and concluding with mammary gland involution. During the periods between successive lactations, the mammary gland undergoes several functional transitions, including active involution after cessation of milk removal, followed by redevelopment and colostrum formation prior to parturition. However the specific cellular and biochemical changes that occur in the mammary gland during the non-lactating period have not been fully defined (Rejman and Hurley, 1988).

The early dry period is a time of rapid functional and structural changes in the mammary gland .In the initial stage of the dry period; the mammary gland undergoes a process of involution. Loss of mammary cells programmed cell death or apoptosis occurs in the mammary gland during involution. Despite this loss of cells, the animal seems to maintain partial integrity of the mammary tissue structure throughout most of the involution period. During this period of involution, the structure and function of the gland must revert back to the non-pregnant

Mammary glands secrete a class of very important proteins during this period of involution. We have isolated a novel glycoprotein from buffalo dry secretions during this period that has a molecular weight of 40 kDa. We have named this mammary gland protein as SPB-40.It shows a high sequence homology to the proteins of Chitinase activity. A very similar protein has earlier been reported as a prominent protein in the whey secretions of non lactating cows (Rejman and Hurley, 1988).A very similar protein has also been found in large amounts in the culture supernatants of MG-63 human osteosarcoma cell line (Johansen et al 1992).However little is known about the functions of these inactive Chitinase related proteins. It is also unclear whether based on their sequence homology one could assume that they may bind to chitin like polysaccharide or glycoprotein. This molecular event would then help regulate various kinds of tissue remodeling processes and or differentiation processes. For example, one of these proteins appears only during wound repair in Cartilage (Johansen et al 1993,Hakala et al, 1993) while a second works during the earliest event of pregnancy when a newly fertilized ovum is implanted in the oviduct (Sendai et al., 1995).A third member of the Chitinase related protein was identified in specific type of cancer cells from the mammary gland of mice(Morrison and Leder,1994).It was also found latter that the normal gland expressed it once the young mouse pups were

weaned. During this period of involution, the structure and function of the gland must revert back to the non-pregnant state. It appears that SPB-40 acts normally as a protective Signaling factor that determines which cells are to survive the drastic tissue remodeling that must occur during involution. Thus, many breast epithelial cells, which have been increased in number during pregnancy, must be destroyed. These cells die by a precise programmed cell death pathway called apoptosis, but most of the breast tissues remain viable and it is assumed the SPB-40 contributes to regulating which cells in the gland are to survive. It has been reported that certain types of breast cancer cells also produce an SPB-40 like protein called BRP39 (Morrison and Leder, 1994, Aroson et al., 1997).

It has been indicated that certain cancers could surreptitiously utilize the proposed normal protective signaling by proteins of this family in order to extend their own survival and thereby allow them to invade the organ and metastasize.

In order to understand the role of SPB-40 and related proteins in apoptosis and breast cancer growth, we report here the isolation, purification, amino acid sequence analysis and detailed three dimensional structure analysis of SPB-40 by X-ray diffraction method. The Structure has revealed unusual aspects of conformational variations leading to alter the scope of functions.

2. REVIEW OF LITERATURE

2.1. Mammary gland and their lactation function

Mammary gland is the most efficient bioreactor that synthesizes and produces milk. The lactation function of the mammary gland is maintained by a delicate balance of systemic blood–borne factors and local mammary –derived factors, many of which are directly affected by the process of milk removal. Systemic factors include lactopoietic hormones such as growth hormone (Bauman, 1992; Tucker, 1994) and suckling – induced prolactin (Tucker, 1994) secretion, which generally stimulate milk secretion. Both prolactin and growth hormone maintain mammary gland function and milk secretion, in part by inhibiting programmed cell death or apoptosis (Flint et al., 2000). The inhibitions of lactation function include factors arising from competing physiological states such as pregnancy. Additionally, local control of milk section is directly linked to physical removal of milk. The factor that plays a major role in regulating milk secretion in many species is a feedback inhibitor of lactation (FIL) found in milk. FIL is thought to be produced by the mammary cell as they synthesize and secrete milk synthesis and secretion. Frequent removal of milk from the gland minimizes local inhibitory effects of FIL and increases milk secretion (Wide et al., 1987, Wilde and Knight, 1989, Wide and Peaker, 1990).

Lactational function of mammary gland occurs in cycles, beginning with mammary gland development, lactogenesis and concluding with mammary gland involution. During the period between successive lactations, the mammary gland undergoes several functional transitions, including active involution after cessation of milk removal, followed by redevelopment and colostrum formation prior to parturition. However, the specific cellular and biochemical changes that occur in the mammary gland during the non –lactating period have not been fully defined (Rejman and Hurley, 1988).

The early dry period is a time of rapid function and structural changes in the mammary gland. In the initial stage of the dry period, the mammary gland undergoes a process of involution. Loss of mammary cell by programmed cell death or apoptosis occurs in the mammary gland during involution. Despite this loss of cells, the animal seems to maintain partial integrity of the mammary tissue structure throughout most of the involution period. During this

period of involution, the structure and function of the gland must revert back to the non-pregnant state.

2.2. Maintenance of Lactation Function

The role of galactopoietic hormones such as prolactin in maintenance of lactation is well established (reviewed by Tucker 1994), although specific cellular mechanisms of action continue to be investigated. Prolactin is considered the major galactopoietic hormone in nonruminants. Prolactin is released at the time of milk removal in ruminants and non-ruminants, and it remains a key systemic modulator of milk secretion during lactation. Conversely, growth hormone is generally considered to be the predominant galactopoietic hormone in ruminants (Bauman, 1992; Tucker, 1994). Inhibition of prolactin secretion or administration of prolactin to lactating cows has little effect on milk yields (Karg and Schams, 1974; Plaut et al., 1987). However, these apparently clear–cut roles of prolactin vs. growth hormone in maintenance of lactation in nonruminants vs. ruminants are probably an oversimplification (Wilde and Hurley, 1996). For example in lactating sheep both prolactin and growth hormone seem to be important for galactopoiesis (Hooley et al., 1978; Tucker, 1994). Even in the rat, recent studies have demonstrated an important role of growth hormone, independent of the role of prolactin (Flint et al., 1992 Flint and Gardner, 1994).

2.3 Declining lactation period

Even if the there is continues milking, milkyield declines as lactation progresses. Mammary tissue function declines after peak lactation and this is due to a decrease in mammary cell number (Wilde and Knight, 1989; Stefanon et al., 2002) and the result of an imbalance between cell proliferation, and cell removal (Stefanon et al., 2002). The cell loss during the declining phase of lactation in the goat and cow is the result of programmed cell death, also called apoptosis (Quarrie et al., 1994; Wilde et al., 1997; Prince et al.,2002). Mammary gland involution is a greatly enhanced extension of these processes leading to a complete cessation of lactation function. In spite of continued milk removal (with associated removal of FIL and

stimulation of a post milking prolactin surge), milk yield in dairy cattle declines as lactation progresses. This decline occurs even with routine administration of growth hormone (bovine somatotropin). The mechanisms that control lactation decline remain important areas of investigation. Mammary involution is a greatly enhanced extension of these processes leading to a complete cessation of lactation function. The timing of inhibition of milk yield in cattle coincides approximately with the period of increasing placentally derived plasma estrogen(Robertson and King, 1979). Estrogen may have an effect on the transition of mammary function from a lactating state to an involutingstate(Athie et al.,1996;Bachman,1982).

2.4 Mammary gland involution

Cessation of milk removal leads to rapid changes in the mammary tissue and initiation of the process of mammary involution (Hurley, 1989; Wagner et al., 2002). Changes in composition of mammary secretions during the early phases of involution indicate rapid changes in the normal mechanisms involved in milk synthesis and secretion (Hurley and Rejman, 1986; Hurley et al., 1987; Noble and Hurley, 1999). These changes in mammary gland secretion include a rapid decline in lactose concentration in the mammary secretions, indicating that lactose synthesis, and the associated water transport mechanism, decline soon after cessation of milk removal. However, total protein concentrations increase in early involution, partially because of water resorption from the secretion and partly due to increased concentrations of lactoferrin, serum albumin and immunoglobulin and other unidentified proteins. Lactoferrin is a major protein found in mammary secretions during involution(Rejman et al., 1989). Its synthesis is increased during involution in contrast to milk-specific proteins such as casein whose synthesis is decreased (Hurley and Rejman, 1993 Hurley et al., 1994).

Involution–associated ultrastructural changes in bovine mammary cells begin within 48 hours after cessation of milk removal (Holst et al., 1987; Hurley., 1989; Li et al., 1997). The most apparent change is the formation of large stasis vacuoles in the epithelial cells (Host et al., 1987), formed largely as a result of intracellular accumulation of milk fat droplets and secretory vesicles (Hurley, 1989). A substantial reduction in fluid volume in the gland occurs between day 3 and 7 of involution (Hurley, 1989), accounting for the reduction in luminal volume. By day 28, the

collapsed alveolar structures that remain are considerably smaller than during lactation, with a very small lumen. General alveolar structure is maintained throughout involution specifically in sheep, cow and goat.

Involution can be divided into at lease two phases (Lund et al., 1996; Alexander et al., 2001; Li et al., 1997), broadly comprising the following: an initial reversible phase whereby the gland maintains its gross morphology but undergoes a substantial increase in the rate of epithelial cell apoptosis (Strange et al., 1992); and a secondary irreversible phase, which involves the destruction of basement membrane by matrix metalloproteinase's, phagocytic clearance of milk, and apoptotic bodies and alveolar collapse(Streuli et al., 1999). Immune cells are present at all stages of mammary development, including involution (Paape et al., 2000), but the precise role of the immune system during post lactation regression has yet to be fully established.

2.5 Histological mount of mammary gland

Histological and ultrastructural work on the bovine mammary gland during involution (Holst et al., 1987; Hurley, 1989) provides no evidence for the extensive tissue degeneration observed in other species, such as rodents and other (Helminen and Ericsson, 1968a, 1968b; Helminen and Ericsson, 1971). Limited authophagocytic processes occur only transiently during the initial two days after cessation of milking. Formation of autophagocytic structures in rodent mammary tissue is characteristic of involution (Helminen and Ericsson, 1968a, 1968b; Helminen and Ericsson, 1971). A detachment of epithelial cells from the basement membrane and their loss from the tissue has been reported in rodents and other species (Helminen and Ericsson, 1968b; Richards and Benson, 1971; Pulan et al., 1996; Accorsi et al., 2002). This leaves characteristic bare spaces on the basement membrane and myoeplthelial cells. No such situations are observed in the involuting bovine gland (Holst et al., 1987). Recently, the involution process in the mouse has been characterized by examining the role of apoptosis.

2.6 Genes expressed during involution

Many genes have been shown to be differentially regulated during involution (Clarkson et al., 2004; Strange et al. 1992; Wilde et al., 1999). Around 145 genes that were specifically

upregulated during the first 4 days of involution**:** of these, 49 encoded immunoglobulin genes. A further 12 genes, including those encoding the signal transducer and activator of transcription 3 (STAT3), the lipopolysaccharide receptor (CD14) and lipopolysaccharide-binding protein (LBP), were involved in the acute-phase response, demonstrating that the expression of acute-phase response genes can occur in the mammary gland itself. Other genes identified suggested neutrophil activation early in involution, followed by macrophage activation late in the process. Immunohistochemistry and histological staining confirmed the infiltration of the involuting mammary tissue with neutrophils, plasma cells, macrophages and eosinophil. Oligonucleotide microarrays are useful tool for identifying genes that are involved in the complex developmental process of mammary gland involution. Chitinase 3-like 1 gene is also expressed during the involution period, the SPX-40, HCGP-39, BRP39, and GP38K belongs to this group (**Table 2.1**)**.** The genes identified are consistent with an immune cascade, with an early acute-phase response that occurs in the mammary gland itself and resembles a wound healing process.

Table 2.1: List of involution-specific genes clustered with the use of Pearson correlation clustering (Stein *et al.,* 2004)

DMT Probe Set	Log ratio	Gene description
101996_at	1.53	Protein tyrosine phosphate, non-receptor type 2
99582_at	1.27	Tumor-associated calcium signal transducer 1
104725_at	2	Small ras-like GTPase Tc10
101410_at	2.96	Claudin 4/mCPE Receptor
93347_at	0.88	RAB24, member of Ras oncogene family
93728_at	2.73	Transforming growth factor β_1 induced transcript 4, TSC-22
102255_at	3.41	Oncostatin M receptor
97948_at	3.39	Retinoblastoma 1
93294_at	4.31	Fibroblast inducible secreted protein, connective tissue growth factor
92858_at	4.89	Secretory leukoprotease inhibitor
93536_at	1.61	Bcl2-associated X protein
161650_at	1.94	Secretory leukoprotease inhibitor
95680_at	1.22	Protein phosphate 1, regulatory (inhibitor) subunit 2
99952_at	7.05	Chitinase 3-like 1
95681_at	2.09	Protein phosphate 1, regulatory (imhibitor) subnit 2
99099_at	0.93	Accute –phase response factor APRF/STAT3
102751_at	1.9	Transforming growth factor,β_3
102313_at	1.15	GTP cyclohydrolase 1
96766_s_at	2.2	Receptor tyrosine kinase, Tyro3

A rapid loss of tissue function and degeneration of the alveolar structure and massive loss of epithelial cells characterize mammary involution in the mouse. This cell loss is due to programmed cell death or apoptosis (Strange et al., 1992, Walker et al., 1989). Apoptosis is both a natural and systematic method of cell suicide, which takes place during normal morphogenesis, tissue remodeling and in response to infection or irreparable cell damage (Wyllie et al., 1980, Schwartzman and Cidlowski, 1993). There are two distinct types of cell death, apoptosis and necrosis, which may be distinguished by morphological, biochemical and molecular changes in dying cell. The process of apoptosis was originally distinguished from necrosis on the basis of its ultrastructure (Kerr, 1971, Kerr et al., 1972). Apoptosis may be identified by a characteristic pattern of morphological changes: nuclear and cytoplasmic condensation, nuclear fragmentation and formation of apoptotic bodies (Walker et al., 1989; Strange et al., 1992). These changes are associated with cleavage of chromatin into discrete sized oligonucleosome fragments by a calcium dependent endonuclease (Arends et al., 1990), resulting in the appearance of oligonucleosomal DNA laddering in ethidium bromide stained gels (Wyllie et al., 1980).

DNA laddering is a phenomenon seen in laboratory tests; it is a sensitive indicator of programmed cell death, specifically of apoptosis.Endonuclease activation is a characteristic feature of apoptosis. This degrades genomic DNA at internucleosomal linker regions and produces 180- to 185- base-pair DNA fragments. On *agaros*e gel electrophoresis, these give a characteristic "laddered" appearance. The dying cell's morphological changes are short-lived and difficult to detect. DNA laddering has therefore become a sensitive method to distinguish apoptosis from ischemic or toxic cell death.Morphology consistent with apoptotic cell death can be observed in the murine mammary gland within two days of milk stasis. The nucleus and cytoplasm condense, the chromatin become fragmented and marginated, and apoptotic bodies are formed (Walker et al., 1989; Strange et al., 1992). This cell loss results in extensive disintegration of alveolar structure during the early period of involution in the mouse. DNA laddering characteristic of apoptosis also has been detected in goat mammary tissue during early (Stefanon et al., 2002) and during late lactation in the cow (Wilde et al., 1997). This suggests that removal of secretory epithelial cells by apoptosis is a normal physiological event in the ruminant mammary gland, even during lactation. In addition, milk stasis has been demonstrated to stimulate DNA laddering in both goat and cow mammary tissue (Quarrie et al., 1994; Wilde et

al., 1997). These observations suggest that mammary epithelial cells are lost during involution in the goat and bovine mammary gland. However, this process of cell loss doesn't seem to be as dramatic as that observed in the mouse. In spite of the loss of cells, bovine mammary alveoli retain general structural integrity throughout involution (Holst et al., 1987).

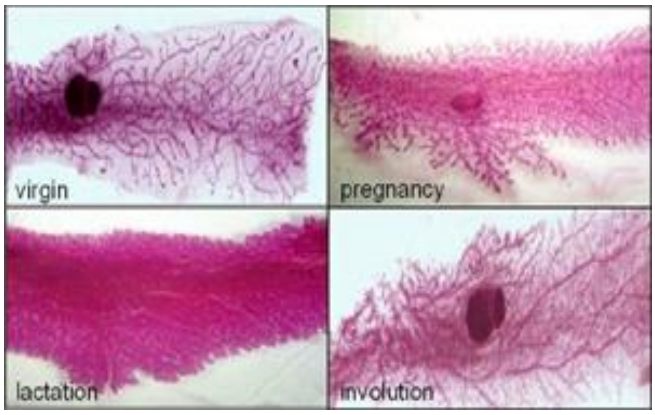


Figure.2.1 Mammary Gland Morphogenesis and Breast Cancer

While the role of cell loss in the mouse mammary gland during involution is dramatic, the impact of mammary gland apoptosis in the bovine does not lead to complete degeneration of the tissue structure. While extensive apoptosis of epithelial cells occur rapidly when lactation ceases, the mechanism of apoptosis induction is not yet known. These changes are associated with cleavage of chromatin into discrete sized oligonucleosome fragments by a calcium dependent endonuclease (Arends et al., 1990), resulting in the appearance of oligonucleosomal DNA laddering in ethidium bromide stained gels (Wyllie et al., 1980).

Apoptosis is a highly regulated process. One important regulator of apoptosis is Bcl-2, a 26-kDa protein that protects cells against apoptosis in a variety of experimental systems (Reed, 1994). Bcl-2 protein is primarily localized to the nuclear envelope, the endoplasmic reticulum, and the outer mitochondrial membranes (Krajewski et al., 1993, Lithgow et al., 1994). Apoptosis is also regulated by a number of genes, including Bcl-2 and related family members (such as Bcl-x, Bax, and Bad), which have significant structural homology with the Bcl-2 gene (Hockenbery et al., 1991; Korsmeyer, 1992; Boise et al., 1993; Oltvai et al., 1993). Apoptosis plays an important role during the involution of the mammary gland. The 6-fold increase in the ratio of Bcl-xS to

Bcl-xL during the first 2 days of involution suggests that apoptotic proteins play a role in predisposing mammary alveolar cells to cell death immediately after lactation.

2.8. Growth Factor involved in Involution

The growth and development of the mammary gland is regulated by a complex set of factors including hormones, cell substratum interactions, and growth factors and their associated receptors. Activation of growth factor receptors leads to the recruitment of a number of cytoplasmic signaling molecules. Mammary gland involution is characterized by a decrease in milk protein synthesis, extensive apoptosis of the secretory alveolar cells, and structural remodeling of the gland, which requires the activation of matrix metalloproteinases (MMPs) and the inactivation of their inhibitors, the tissue inhibitor of metalloproteinases (TlMPs) (1-6). It has been reported that 50 - 80 % of the mammary epithelial cells in mice that are present during lactation undergo apoptosis during involution. Signal transducer and activator of transcription-3 (STAT3) in the mammary gland results in delayed involution and decreased epithelial cell apoptosis, suggesting that STAT3 is critical for the normal process of apoptosis during involution. Growth factors or docking molecules then results in the activation of a number of molecules. STAT5a (Lavnilovitch et al., 2002) is involved in milk production during lactation, and its phosphorylation decreases at the beginning of involution. STAT3 (Copeland et al., 1995; Miyoshi et al., 2001) is required for normal involution and its phosphorylation increases at the beginning of involution decrease Akt in these cells at the onset of involution may be necessary for apoptosis to proceed.

2.9. Anti-apoptotic proteins during involution

Akt may play important roles in tumorigenesis by inhibiting apoptosis and or promoting metastasis. Activation of Akt alone can interfere with the apoptotic process of mammary gland involution and promote tumor progression by providing an important cell survival signal. Activation of Akt provides complementary cell survival signals that are required for mammary tumorigenesis. Activation of Akt can interfere with normal mammary gland involution by

attenuating apoptotic death in the involuting mammary gland. Expression of activated Akt is involved in promoting tumor progression by providing a critical cell survival pathway. Mammary epithelial expression of Akt can result in profound delays in mammary gland involution, a process involving extensive apoptotic cell death. Recent studies suggest that activated Akt can cooperate with these signaling pathways to efficiently induce mammary tumorigenesis.

Activation of the antiapoptotic protein kinase Akt is induced by a number of growth factors that regulate mammary gland development. Akt is expressed during mammary gland development, and expression decreases at the onset of involution(Kathryn et al.,2001). Analysis of mammary glands from the mice reveals a delay in both involution and the onset of apoptosis under the control of the mouse mammary tumor virus (MMTV) promoter (Strange et al., 1992; Walker et al., 1989;Rudolph-Owen et al; 1998). Expression of tissue inhibitor of metalioproteinase-1 (TIMP-1), an inhibitor of matrix metalloproteinases (MMPs), is prolonged and increased in the transgenic mice, suggesting that disruption of the MMP: TIMP ratio may contribute to the delayed mammary gland involution observed in the transgenic mice.

2.10. Bcl-2 gene Family and related proteins in mammary gland Involution and breast cancer

The Bcl-2 gene family regulates tissue development and tissue homeostasis through the inter play of survival and death factors (Strange et al., 1992). Family members are characterized as either pro-apoptotic or anti-apoptotic, depending on cellular context. In addition to its anti-apoptotic effect, Bcl-2 also inhibits progression through the cell cycle. Functional interactions between family members as well as binding to other cellular proteins modulate their activities. Mammary gland tissue, similar to many other tissues, expresses a number of different Bcl-2 relatives including bclx, bax, bak, bad, bcl-w, bfl-1, bcl-2 (Jager et al; 1997; Li et al., 1997) as well as the bcl-2 binding protein Bag-1. Bcl-2 is expressed in the non-pregnant mammary gland and early pregnancy. In contrast, expression of bcl-x and bax continues through late pregnancy, is down regulated during lactation, and up regulated with the start of involution. Bak, bad, bcl-w, and bfl-1 are also up regulated during involution. The specific roles of individual gene products are investigated using dominant gain of function and loss of function mice. Finally different Bcl-2 family members are commonly over- or under-expressed in human breast cancers. Bcl-2 expression in human breast cancers has been associated with a good prognosis, while decreased Bax expression has been linked to poor clinical outcome (Jager et al., 1997; Li et al., 1997).

2.11. Proteins secreted during involution

As chitinase 3-like 1 proteins gene are also unregulated with the start of involution as referred in the (**Table 2.1**). SPB-40 or YKL-40 belongs to this group of proteins, which is also expressed during involution period. Understanding the role of these family members play in regulating mammary epithelial cell survival is salient to both normal mammary gland physiology and the development of new therapeutic approaches to breast cancer.

SPB-40 is a novel 40 kDa glycoprotein secreted during the process of mammary gland involution in Goat (Mohanty et al., 2003), Sheep, Porcine and Buffalo (Srivastava et al., 2006a, c and Ethayathulla et al., 2006), Bovine (Kumar et al., 2006a) the specific in vivo function of this protein is not yet known. The expression pattern of SPB-40/YKL-40 is associated with many

cancers and inflammatory disorderness as shown in **(Figure 2.2)**. However, its possible mechanism of action has been proposed. A very similar 39-kDa glycoprotein has earlier been reported in the whey secretions of non-lactating cow (Rejman and Hurley, 1988). Similar type of protein BRP39 (Morrison et al., 1994) is also expressed during termination of involution in mice.

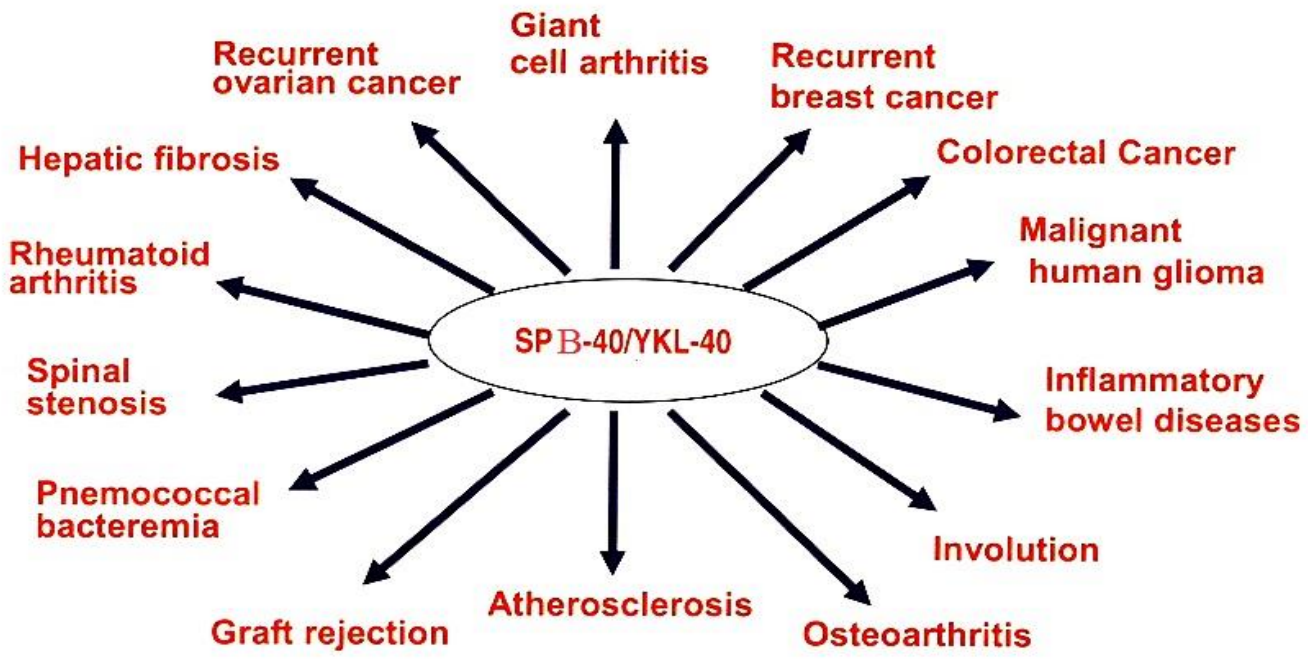


Figure 2.2: **Circumstantial expression of SPB-40 / YKL-40 occurs in a wide variety of cell types and in particular from the cells located in tissues with increased remodeling / degradation or inflammation of the extra cellular matrix.**

Lactoferrin is an iron-binding protein whose concentrations increase markedly during active involution

(Rejman et al., 1989). It is thought to compete with bacteria for iron. Lactoferrin impedes iron utilization by bacteria and causes bacteriostasis as a result of its iron sequestering properties(Sanchez et al., 19920. LF may play a role in the mammary gland by affecting phagocyte function(Campbell, 1982). LF also may limit the oxidative degeneration of cellular components that occur during periods of tissue disruption such as during inflammation and involution.

Insulin like growth factor binding proteins-5 (IGFBP-5) has been reported to be present in high concentration after 48 hours of mammary involution induced by removal of suckling young (Tonner et al., 1997). IGFBP-5 has a role in controlling extensive cell death when the mammary gland involutes at the end of lactation (Flint et al., 2000). Various apoptosis-related proteins like TGF- β1 (auto/paracrine inducer) and its receptor (TGF-βRIII), Bax (promoter), Bcl2 (inhibitor) and CPP-32 (executor) of apoptosis are expressed during mammary gland involution of goat and sow (Wareski et al., 2001; Motyl et al., 2001). The increase in Bcl-2 level in remnant lobuloalveolar tissue is probably the molecular mechanism that limits the secretory tissue involution. The induction of CPP-32 (caspase 3) in the dry period is accompanied by a progressive loss of mammary epithelial cells and increase in apoptotic cell number. The increase in the expression of the above proteins in the late lactation and dry period indicates their involvement in the induction (TGF-βI and TGF-βIII), regulation (Bax and Bcl-2) and execution (CPP-32) of apoptosis in the course of mammary gland involution in goat (Wareski et al., 2001) and in sow (Motyl et al., 2001). Transgenic mice overexpressing insulin-like growth factor-l (IGF-I) and insulin-like growth factor binding protein 3 (IGFBP-3) in the mammary gland also exhibit delayed involution and decreased apoptosis. IGF-I has been shown to be involved in both proliferation and suppression of apoptosis, and IGFBP-3 may be involved in facilitating the actions of IGF-I in the mammary gland.

Concentrations of milk-specific components (caseins, α-lactalbumin, β-lactoglobulin, and milk fat) decline slowly during the first 2 - 3 weeks of the dry period, but do not disappear completely. But the Lactose concentrations decline rapidly (Aslam et al., 1994; Hurley and Rejman, 1986 Hurley, 1987). Proteins of serum origin (immunoglobulins, serum albumin) increase in concentration during the first week of involution. All classes of immunoglobulins (Ig) increase, including IgG1, IgG2, IgA, and IgM. There is a transitory increase in the selective transport of IgG1 from about days 2 to 4 of involution. Concentrations of plasmin, plasminogen, and plasminogen activators also increase in mammary secretions during involution (Aslam and Hurley, 1997b; Politis et al., 1989), which have been implicated in the tissue remodeling. The plasmin system may also be responsible for the hydrolysis of milk proteins in the mammary tissue during involution(Aslam et al., 1994; Aslam and Hurley, 1997b). However, other protease sources, such as leukocytes and nonplasmin proteases (Grieve and Kitchen, 1985), might also contribute to the fate of milk proteins during involution. Matrix metalloproteinases (MMPs), a

family of zinc dependent endopeptidases are also expressed which play a key role in the tissue remodeling events occurring during involution (Uria and Werb, 1998).

NAGase (N-acetyl-β-D-glucosaminidase) activity increases substantially during mammary gland involution. NAGase is aIysosomal enzyme that is secreted in large quantities in the mammary gland during involution and inflammation. The specific function of NAGase in the gland is not known, but sometimes its activity in mammary secretions is used as an indicator of tissue changes that accompany involution and inflammation.

2.12. Cells in secretions of the involuting mammary gland

Very few cells (less than 2 %) in mammary secretions during involution are epithelial cells. Most cells in the secretion are leukocytes. Total leukocyte concentrations in mammary secretions increase rapidly in early involution.

PMN (polymorphonuclear neutrophils): These are phagocytic leukocytes. They predominate for the first 3 - 7 days. **Macrophages** are the predominant cell type after 7 days. These are also phagocytes. Many are filled with ingested fat droplets and other debris. They play a major role in removing large quantities of fat and cellular debris, including dead PMN. Macrophages are also the predominant cell type in the colostrum at parturition**. Lymphocytes** are always present. They increase in proportion roughly in parallel with the macrophages, but may only become the predominant cell type during the mid-dry period. Specific function of Iymphocytes in the involuting mammary gland is not known.

Tissue remodeling and differentiation processes also has been found In Culture supernatants of **MG 63 human osteosarcoma cell line** (Johansen et al., 1992), cultures of human synovial cells (Nyirkos and Golds, 1990) and human cartilage cells (Johansen et al., 1993; Hakala et al., 1993) and cancer cells from the mammary gland of mice (Morrison and Leder, 1994) where specific type of proteins were observed.

2.13 SPB-40 and other homologous proteins

SPX-family

In order to understand the role of SPX-40 proteins in tissue remodeling and apoptosis, these proteins were isolated, purified, amino acid sequence analysis and detailed three-dimensional structure analysis of SPX-40 from various sources such as Sheep (Srivastava et al., 2006a), Porcine (Srivastava et al., 2006c), Bovine (Kumar et al., 2006a), Goat (Mohanty et al., 2003) and Buffalo (Ethayathulla et al, 2006) has been determined by X-ray diffraction method. SPS-40 from dry secretion of sheep (Srivastava et al. 2006a), MGP-40/SPG-40 from dry secretion of goat (Mohanty et al., 2003; Kumar et al., 2006b), SPB-40 from bovine (Kumar et al., 2006a), SPP-40 from Porcine (Srivastava et al., 2006c), SPB-40 from buffalo (Ethayathulla et al., 2006) and SPU-40 from camel have very high sequence identity are expressed in the mammary gland during non-lactating period once the milking in animal is stopped.The structure has revealed features indicating unusual aspects of conformational variations leading to alter the scope of functions.

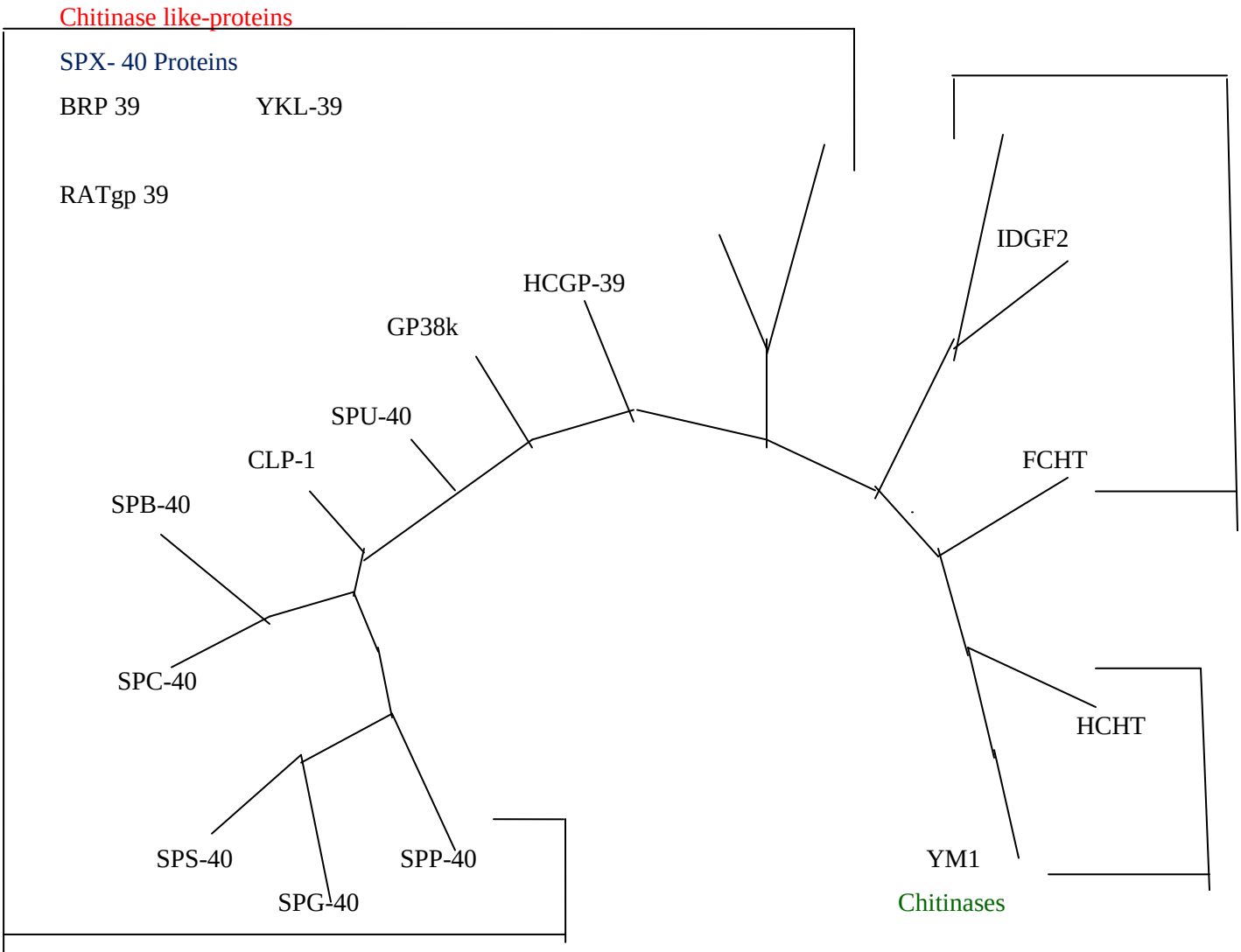
Various groups from time to time have also reported similar kind of proteins expected to be the species-specific homologue of SPB-40. The N-termini sequence of SPB-40 and their homolog are very similar as shown in **(Table 2.2).**

S.No.	Source	Protein	N-terminal sequence
1.	Buffalo	SPB-40	YKLICYYTSW
2.	Goat	MGP-40	YKLICYYTSW
3.	Bovine	SPC-40	YKLICYYTSW
4.	Sheep	SPS-40	YKLICYYTSW
5.	Porcine	SPP-40	YKLICYYTSW
6.	Chitinase like protein-1	CLP-1	YKLVCYYTSW
7.	Porcine smooth muscle cells	GP38k	YKLVCYYTSW
8.	Human cartilage glycoprotein	HCGP-39	YKLVCYYTSW
9.	Rat activated macrophage glycoprotein	Ratgp39	YKLVCYYTNW
10.	Mouse breast regressing protein	BRp39	YKLVCYFTSW
11.	Mouse macrophage novel lectin	YM1	YQLMCYYTSW

Table 2.2 N-terminal amino-acid sequences of first ten amino acids of SPB-40 and its sequence similarity with other proteins.

The sequence identities of more than 80 % to some other mammalian glycoproteins, such as bovine chondrocyte chitinase - like protein (CLP-1; GenBank accession number: AF011373), human chondrocyte glycoprotein (YKL-40/HCGP-39; GenBank accession number: M80927) (Johansen et al., 1993; Hakala et al., 1993), porcine heparin-binding glycoprotein (GP38k; GenBank accession number: U19900) (Shackelton et al., 1995) and rat cartilage glycoprotein (RATgp39; GenBank accession number: AF062038). All these proteins have identical chain lengths and similar glycosylation sites. They all contain 5 cysteine residues with two disulfide bridges while identical Cys20 is unpaired. Yet another similar protein, BRP39 (breast regression protein with M.w. 39 kDa) was identified from a specific type of cancer of mammary gland of mice (Morrison &Leder, 1994) that also showed a sequence identity of 69 % to SPB-40. BRP-39, a "breast regression protein'', of 361 amino acids (matured protein) is induced in mammary epithelial cells of mouse a few days after weaning, a stage where the mammary gland undergoes involution. The same protein is also expressed by murine mammary tumors initiated by neu/ras oncogenes (Morrison and Leder, 1994; Aronson et al., 1997). Due to very similar amino acid

sequences and chemical properties, these mammalian proteins form a subclass of closely related mammalian glycoproteins (Group I). This group I proteins are also homologous to chitinases with an approximate sequence identity of 52 % with human chitinase (Renkema et al., 1995). The chitinases have a well-defined carbohydrate-binding groove in which oligomers of N-acetylglucosamine (chitin polymers) bind preferentially. The active site of chitinases involves three acidic amino acids, Asp, Glu and Asp. The sequence comparison of SPS-40 with chitinases shows that the residue corresponding to Glu is changed to Leu. This has abolished the chitin-hydrolyzing capability of SPS-40. Inspite of several similarities, chitinases are distinct functionally with different properties from group I proteins. Therefore, they are classified here as group II proteins. There is yet another class of closely related proteins designated as chitinase - like proteins. These proteins are also catalytically inactive due to mutations of one of the catalytic residues but are similar to chitinases in the folding of their polypeptide chain. They lack glycosylation sites and show mutations of several carbohydrate-binding residues. The prominent protein among them is a novel mammalian lectin, YM1 (Sun et al., 2001). Although initially YM1 was reported to be binding to carbohydrates (Sun et al., 2001) but later on it was described that it does not possess an ideal carbohydrate-binding site (Tsai et al., 2004). Overall, YM1 is homologous to both groups I and II proteins but has several unique features that are distinct from the members of both groups. The proteins of YM1 subclass will be referred to hereafter as group III proteins. Because of sequence and structural similarities among the proteins of groups I, II and III, they essentially belong to a single superfamily to be referred hereafter as SPB-40 superfamily. Phylogeny of SPB-40 proteins, chitinase- like proteins and chinitases generated using Phylip 3.6 (PHYLogeny Inference Package is shown in (**Figure 2.3**)



Name	Description	Name	Description
FCHT	Fungal Chitinase	HCHT	Human Chitinase
YM1	Mammalian Macrophage protein	IDGF2	Imaginal disk growth factor (Drosophila)
YKL-39	Human glycoprotein	BRP39	Breast regression protein (Mouse)
RATGP39	Rat glycoprotein	HCGP-39	Human chondrocyte glycoprotein
GP38K	Porcine glycoprotein (smooth muscle cells)	CLP-1	Bovine chondrocyte glycoprotein
SPU-40	Camel signaling glycoprotein	SPP-40	Porcine signaling glycoprotein
SPB-40	Buffalo signaling glycoprotein	SPC-40	Bovine signaling glycoprotein
SPS-40	Sheep signaling glycoprotein	SPG-40	Goat signaling glycoprotein

Figure 2.3:**Phylogeny of SPB-40 proteins,chitinase-like proteins and chitinases generated using Phylip 3.6(PHYLogenyInfernce package)**

Hence the SPB-40 family proteins are CLP-1 (Bovine chondrocyte glycoprotein), GP38k (Porcine glycoprotein from smooth muscle cells) (Shackelton et al., 1995), HCGP-39 (Human chondrocyte glycoprotein) (Johansen et al., 1993; Hakala et al., 1993), RATgp39 (Rat cartilage glycoprotein) and BRP39 (Breast regression protein from mouse) (Morrison and Leder, 1994, Aronson et al., 1997) are categorized as group I proteins. The chitinases HCHT (Human chitinase) and FCHT (Fungal chitinase) as group II and the YM1 (Mammalian Macrophage protein) (Sun et al., 2001), IDGF2 (Imaginal disk growth factors from Drosophila) (Varela et al., 2002) and YKL-39 (Human glycoprotein) are chitinase-iike proteins and referred as group III proteins.

2.14. Proposed functions of SPB-40 and their homologous proteins

Group I (SPB 40 family proteins)

The exact physiological function of SPB-40 and related proteins are not yet known. However, on the basis of its expression pattern during mammary gland involution, it appears that SPB-40 acts normally as a protective signaling factor that determines which cells are to survive the drastic tissue remodeling that occurs during involution. During this period of involution, the structure and function of the gland must revert back to the non-pregnant state. Thus, many breast epithelial cells, which have been increased in number during pregnancy, must now be destroyed. These cells die by a precise programmed cell death pathway called apoptosis, but most of the breast tissues remain viable and it is assumed that SPB-40 contributes to regulating which cells in the gland are to survive (Mohanty et al., 2003; Kumar et al., 2006a; Srivastava et al., 2006a; Kumar et al., 2006b; Srivastava et al., 2006b).

The 39-kDa-whey glycoprotein/SPB-40 secreted in bovine during nonlactating period is also secreted under similar conditions. Rejman and Hurley (1988) have suggested this protein to be a marker protein of mammary gland involution. The molecular mass of this protein was approximately calculated to be 39 kDa in SDS-PAGE. Biochemical experiments on incubation of 39-kDa protein with endoglycosidase H suggest that this is a glycoprotein with high mannose

oligosaccharides. This is immunologically distinct from other milk-associated or Serum proteins. Its particular presence in mammary secretions during involution provides an important marker protein for mammary function (Rejman and Hurley, 1988).

BRP-39: Similar proposal with respect to the function of BRP-39 has also been coined because of its similar expression pattern during mammary gland involution. Mammary involution involves programmed cell death and apoptosis, and possibly BRP-39 participates in the various signal transduction pathways during this period that lead to apoptosis of the regressing cells (Aronson et al,.1997).As BRP-39 has also been observed during neu/res initiated cancer (Morrison and leder,1994), it has been proposed that certain cancers could surreptitiously utilize proposed normal protective signaling by BRP-39 in order to extend their own survival and thereby allow them to invade the organ and metastasize. This class overexpresses a growth factor receptor on their cell surfaces that is called HER-2/neu.In human breast cancer paitents, those who are positive for excess HER-2/neu protein have a negative prognosis for remission and survival. Upto a third of breast cancers contain extra copies of theHER-2/neu gene, and biotech companies have recently developed tests to measure the level of HER-2/neu in breast cancer paitents. In late 1998, anew anti-HER-2/neu drug called Herceptin that attacks the growth factor receptor was approved by the FDA.So far it has shown positive therapeutic results, especially when combined with a chemotherapeutic agent like Taxol. Breast cancer is the leading type of cancer in women and is the second leading cause (after lung cancer) of cancer death among women. In the United States in 1998, an estimated 178,700 new cases of breast cancer were expected to be diagnosed, and 43,500 women were expected to die of this disease. Approximately 2,000,000 women have been diagnosed with breast cancer (NCI statistics 1998). Breast cancer also occurs in men, although far more rarely than in women (approximately 1600 cases in men were diagnosed in treatment for breast cancer in men is guided by our understanding of the disease in women).

YKL-40/HCGP~39:

YKL-40 may play a role in events associated with increased turnover of extracellular matrix. The high serum level of YKL-40 in serum of patients with recurrent breast cancer indicates a poor prognosis for the patient's survival suggesting that it is has a role in cancer growth (Johansen et al., 1995). The spontaneous increase in the production of YKL-40 in the early phase of culture appears to represent a cellular response to changes in the extracellular matrix environment. This, coupled with the profound suppressive effects of IL-1β and TGFβ on YKL-40 production, identifies a novel regulatory pattern for this major chondrocyte-derived protein. Gene expression microarray analysis shows YKL-40 to be a potential serum marker for malignant characters in glioma. Serum YKL-40 level is high in patients with hepatic fibrosis as well as rheumatoid arthritis. Biochemically it has been proved that the expression of HER-2/neu& YKL-40 in breast cancer is co-localized. By immunohistochemical studies, it has been proved that YKL-40 is localized in breast cancer.

Human cartilage glycoprotein 39 (HCGP-39) is a glycoprotein secreted by articular chondrocytes, synoviocytes and macrophages. Increased levels of HCGP39 have been demonstrated in synovial fluids of patients with rheumatoid or osteoarthritis. The increased secretion of HCGP-39 under physiological and pathological conditions with elevated connective-tissue turnover suggests its involvement in the homoeostasis of these tissues. HCGP-39 promotes the growth of human synovial cells. A dose-dependent growth stimulation lung cell lines was exposed to HCGP-39 in a concentration range from 0.1 to 2 nM, which is similar to the effective dose of the well characterized mitogen, insulin-like growth factor-1. At suboptimal concentrations, the two growth factors work in a synergistic fashion. The use of selective inhibitors of the mitogen-activated protein kinase and the protein kinase B (AKT) signaling pathways indicates that both are involved in mediating the mitogenic response to HCGP-39. Phosphorylation of both extracellular signal regulated kinases and AKT occurred in a dose- and time-dependent fashion upon addition of HCGP-39. Activation of these signaling pathways could also be demonstrated in human chondrocytes. Thus, HCGP-39 initiates a signaling cascade in connective-tissue cells, which leads to increased cell proliferation, suggesting that this protein could play a major role in the pathological conditions leading to tissue fibrosis. Expression of the chitinase 3-like protein, HCGP-39 is associated with conditions of increased matrix turnover and tissue remodeling. High levels of this protein have been found in sera and synovial fluids of

patients with inflammatory and degenerative arthritis. In order to assess the role of HCGP-39 in matrix degradation induced by inflammatory cytokines, its effect on the responses of connective tissue cells to TNF-α and IL-1 with respect to activation of signaling pathways and production of matrix metalloproteases and chemokines. Stimulation of human skin fibroblasts or articular chondrocytes with IL-1 or TNF-α in the presence of HCGP-39 resulted in a marked reduction of p38 and SAPK/JNK phosphorylation **(Figure 2.4),** while nuclear translocation of NFκB proceeded unimpeded. HCGP-39 suppressed the cytokine-induced secretion of the matrix metalloproteases MMP1, MMP3 and MMP13, as well as secretion of the chemokine IL-8. The suppressive effects of HCGP-39 were dependent on Pl3 kinase activity, and treatment of cells with HCGP39 resulted in AKT-mediated serine/threonine phosphorylation of the apoptosis signal-regulating kinase, ASK1. This process could therefore be responsible for the down-regulation of cytokine signalling by HCGP-39. The activation of ASK1 following binding of TNF-α or IL-1 to their respective receptors involves several specific kinase-mediated steps and adapter molecules in addition to those illustrated here (TRADD: TNFR1-associated via death domain, TRAF: TNF receptor associated factor, IRAK: IL-1 receptor associated kinase). HCGP-39 binding to a putative receptor leads to PI3 kinase mediated phosphorylation of AKT, which can down- regulate ASK1 activity by catalyzing phosphorylation of serine at position 83. This in turn leads to decreased activation of JNK and p38 and finally decreased production of MMPs and chemokines. Activation of the NF-κB pathway is mediated at the level of TRAF-2 and TRAF-6, upstream of ASK1, and thus is not affected by HCGP-39.

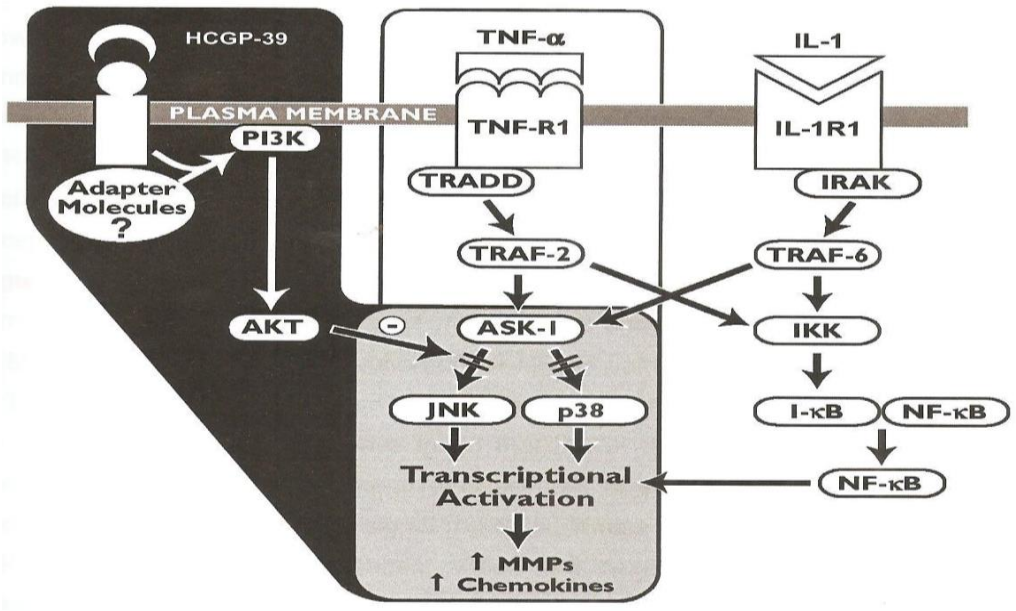


Figure 2.4: The regulatory pathway utilized by HCGP-39 is shown by Schematic representation of signaling cascades mediating responses to TNF-α and IL-1 in connective tissue cells. TNF-α Signaling is represented by its interaction with R1 receptor The activation of ASK1 following binding of TNF-α or IL-1 to their respective receptors. (Ling and Recklies, 2004).

Recklies et al., (2002) have found that HCGP-39 stimulated DNA synthesis and proliferation of human connective-tissue cells at subnanomolar concentrations and activates both extracellular signal-regulated kinase (ERK) and protein kinase B (AKT)-mediated signaling cascades, which are associated with the control of mitogenesis suggesting that this protein could play a role in the pathological conditions leading to tissue fibrosis. The activation of cytoplasmic signal-transduction pathways suggests that HCGP-39 interacts with one or several signaling components on the plasma membrane. Whether this occurs through specific receptors on the cell surface of connective tissue cells remains to be determined, and the nature of potential HCGP-39 ligands at the moment is speculative (Recklies et al., 2002).

Gp38k is a glycoprotein expressed by cultured vascular smooth muscle cells during the time of transition from a proliferating monolayer culture to a non-proliferating multilayered (differentiated) culture. Expression continues as the cell culture forms multicellular nodules. Its appearance during nodule formation suggests a role in smooth muscle cell differentiation (Millis et al., 1985; Shackelton, 1995). Because this transition period involves active cell migration

(Malinda et al., 1999) have evaluated the effects of exogenously added gp33k on vascular endothelial cell (HUVEC) migration and chemotaxis. It was demonstrated that gp38k acts as a chemoattaractant for HUVECs and stimulates cell migration in Boyden chambers at a level comparable to that achieved with the known endothelial cell chemoattractantbFGF. These results have provided evidence that gp38k may function in angiogenesis by stimulating the migration and reorganization of vascular endothelial cells**.**

Oviductin: Oviduct-specific glycoproteins are associated with zonapellucidae and/or vitelline membranes of oviductal eggs and developing embryos or are selectively sequestered in the perivitelline space of eggs (Sendai et al., 1995). Oviductin is expressed under the influence of estrogen and its proposed role is involvement in fertilization or early cell division by the fertilized egg or its implantation in the uterus (Desouza and Murray, 1994; Aronson et al., 1997).

Group II (Chitinases)

These group I proteins are also homologous to chitinases showing the sequence identity to the extent of 52 % with human chitinase (Renkema et al., 1995) and 25 % sequence identity with fungal chitinase. Chitinases have been studied well and their three-dimensional structures have shown that they consist of a $(\alpha/\beta)_8$ topoisomerase (TIM) barrel fold and a small $(\alpha + \beta)$ domain (Fusetti et al., 2002). The chitinases have a well-defined carbohydrate-binding groove in which oligomers of N-acetylglucosamine (chitin polymers) bind preferentially. The binding site contains a number of aromatic residues that are involved in carbohydrate recognition. The active site of chitinases is situated in the carbohydrate-binding groove and has three acidic amino acids, Asp, Glu and Asp. The sequence comparison of SPB-40 with chitinases shows that the residue corresponding to Glu is changed to Leu119 in the entire group I proteins. This has abolished the chitin hydrolyzing capability of group I proteins. Inspite of several similarities, chitinases are distinct functionally with different properties from group I proteins. Therefore, they are classified here as group II proteins.

Group III

IDGF2: A role for CLPs during embryogenesis has been observed in Drosophila. Four novel type CLP growth factors, called as imaginal disk growth factors (IDGFs) 1-4, which are secreted from embryonic yolk cells and the fat body of the embryo and larva, have mitogenic activity on imaginal disk cells (Kawamura et al., 1999). Recently the three dimensional structure of IDGF2 has been known (Varela et al., 2002). The structure elucidates an occluded sugar binding groove suggesting that the protein may not bind / may partially bind to saccharides / oligosaccharides assuming several conformational changes to open the partially occluded binding site. YM1 and IDGFs present 15 - 30 % amino acid sequence homology to family 18 glycosyl hydrolases.

YM1: Recently elucidated molecular structure of the murine protein YM1 (Sun et al., 2001; Tsai et al., 2004) that belongs to the vertebrate/mammalian chitinase like protein family possessing no enzymatic activity. This protein is structurally very similar to catalytic domain of the SerratiamarcescenschitinaseA (Perrakis et al., 1994). Further analysis indicates that YM1 binds preferentially to oligomers of glucosamine as well as to heparin, suggesting that this protein may act by binding to carbohydrate ligand, possibly cell-membrane associated heparin sulfate proteoglycans. Interestingly, YM1 did not bind to N-acetylglucosamine oligomers, the native substrate for chitinases. A similar lack of interactions has also been observed between chitin oligomers and HCGP-39, suggesting that binding to chitin, possibly as a defense mechanism against fungal or parasitic infection is probably is not the major function of this protein (Recklies et al., 2002). However, the recent crystal structure of HCGP-39 in complex with chitin oligomer under some special condition suggests that HCGP-39 is binding to chitin oligomer. However, it remains speculative whether, it actually can act as a physiological ligand in in-vivo conditions(Houston et al., 2003).

However, IDGFs and YM1 and other SPB-40 homologous proteins have no known catalytic activity due to mutation of critical glutamic acid residue. In this respect, several proteins with sequence homology to chitinases but no detectable enzymatic activity has been described in vertebrates, indicating that the typical chitinase-like fold may be present in proteins with a wide range of biological functions other than the chitin degradation (Varela et al., 2002). The conserved consensus sequence 'DXXDXDXE' observed in chitinases has been modified in these-

inactive chitinase-like proteins, which are 'DXXNXDXQ' in YM1 (Sun et al., 2001) and 'DXXDXAXQ' in IDGF2 (Varela et al 2002) that does not completely fulfill this consensus.

2.15 Structural properties

The SPB-40 is homologue to other proteins such as BRP39, HCGP-39, GP38k, RGP39 have a molecular mass of around 40 kDa consisting of 361, 362, 362 and 362 amino acid residues respectively (Varela et al., 2002) while mature oviductin has a molecular mass of approximately 57 kDa (Sendai et al., 1995). These are all secretory proteins secreted from various sources in special circumstances where tissue remodeling takes place. All these proteins have a 'pro' sequence of 21 amino acids residues having a role in signal transduction and extracellular secretion (Holst et al., 2001). HCGP-39 and the other proteins cited above share high sequence homology and structural relationship with family 18 chitinases. A unifying feature of these so called mammalian chitinase-related proteins with specific reference to HCGP-39 is their molecular structure, which consists of an essentially complete chitinase-3-like catalytic domain of approximately 40 kDa (Recklies of al., 2002). They do not possess glycohydrolase activity as observed in chitinases. This lack of activity can be explained by the non-conservative substitution of Leucine or Glutamine in place of the essential glutamic acid residue, which acts as a H+ donor in the active site of these proteins (Watanabe et al., 1993; Perrakis et al., 1994; Aronson, 1997). The remainder of the sequence, which forms the active site cleft in the chitinases, is very highly conserved amongst all members of glycosyl hydrolase family 18, suggesting that structure may be similar in the enzymatically active chitinases and inactive chitinase-related proteins (Recklies et al., 2002). The (α +β) domains of SPB-40 has high sequence identity & structural folding with FK binding proteins (FKBP), Human cytosolic FKBP12 also has peptidyl-propylisomerase activity and has been shown to speed up the rate of protein folding FKBP12 is the soluble receptor for the immunosuppressant drug FK506. The tightly bound (Kd 3 nM) heterodimeric complex binds and inhibits the phosphatase calcineurin thus blocking the immunostimulatory signal transduction pathway in T cells. FK506, in complex with its 12 kDa, binding protein FKBP12, exerts its immunosuppressive effects through the inhibition of calcineurin (CN), an intracellular Ca++ -calmodulin-dependent phosphatase (Klee

and Cohen, 1988). CN inhibition, in turn, interrupts the induction of IL-2 and other T-cell activation events. The functional identification of distinct sets of antitumor activities mediated by the FKBP gene family.

2.16 Family 18 glycosyl hydrolases

Family 18 glycosyl hydrolases, which includes the chitinases, are the enzymes that catalyze the hydrolysis of β-(1, 4)—N-acetyl-D-glucosamine linkages in chitin polymers of the arthropod cuticle (Henrisaat and Davies, 1997). Glycosyl hydrolase protein family 18 shows the conservation of residues in the, β4-loop # 4, which is 'DXXDXDXE' (X represents any amino acid). The signature includes Glu315 of chitinase A that acts as proton donors in the process of hydrolysis of chitin polymers (Perrakis et al., 1994). Several proteins from different species have been analyzed **(Table 2.3)** with respect to this consensus sequence, which is highly conserved.

2.17. TIM barrel domain

The first protein that was discovered to have an eight-stranded α/β domain was Triose phosphate isomerase (TIM), which is one of the most abundant domain structures in nature, also called (β/α)8-barrel (Alber et al., 1981). It was estimated that about 10 % of the proteins in nature contain a TIM barrel domain (Wierenga, 2001). A large number of these proteins are glycosyl hydrolases. The TlM-barrel is made up of eight β-strands tethered to eight α helices by loops. The β -strands of the barrel form an intrinsic network of hydrogen bonds with the neighbouring strands and are oriented in the same direction. The overall twist associated with all the strands cause the first and the eighth strand to register in parallel held in place by hydrogen bonds causing the closure of the barrel. The loops at the amino termini of the barrel β -strands are usually short, while the loops at the carboxyl termini are long and form a unique shape that composes the structure of the active site and determines substrate specificity In addition; many (α/β)8 enzymes have additional domains that are not part of this fold. It has been suggested that the TlM-barrel provides a framework that preserves the positions of these loops. It has been reported that in many TlM-barrel active enzymes, the catalytic proton donor is located at loop # 4 of the barrel. This observation suggests that TlM-barrel enzymes are evolutionarily related and are derived from a common ancestor. Recently, based on structure and sequence similarities it was suggested that two TlM-barrel proteins in the histidine biosynthesis pathway evolved by gene duplication and gene fusion from a common half-barrel (β/α) 4 ancestor. Several inactive chitinases from plant sources such as Narbonin and Concanavalin B, few mammalian chitinase-like proteins such as SPB-40 (Mohanty et al, 2003; Kumar et al., 2006a; Srivastava et al, 2006a), HCGP-39 (Fusetti et al., 2003); YM1 (Sun et.al., 2001; Tsai et al., 2004) and IDGF2 (Varela et al., 2002) have also been reported to have (β/α) 8 TIM domain and a small (α+β) folding domain.

Table 2.3: Alignment of a highly conserved region of SPB-40 and other chitinase related proteins with consensus sequence of some chitinases belonging to family 18-glycosyl hydrolase.

SPB-40	**FDGLDLAWLWP**
MGP-40	**FDGLDLAWLYP**
SPC-40	**FDGLDLAWLYP**
SPS-40	**FDGLDLAWLYP**
SPP-40	**FDGLDLAWLYP**
Bovine (CLP-1)	**FDGLDLAWLYP**
GP38k	**FDGLDLAWISP**
HCGP/YKL-40	**FDGLDLAWLYP**
Ratgp	**FDGLDLAWLYP**
BRP39	**FDGLDLAWLYP**
Human oviductin	**FDGLDLFFLYP**
Mesocricetusauratusoviductin	**FDGLGLFFLYP**
SerratiamarcescensChitinase A	**FDGVDIDWEFP**
Coccidiodesimmitischitinase (chit1)	**FDGIDIDWEYP**
Bacillus circulanschitinase A1	**FDGVDLDWEYP**
Brugiamalaychitinase	**FDGFDLDWEYP**

2.18. Physico-chemical characteristics

The physico-chemical characteristics with respect to total number of amino acids, Molecular mass and isoelectric points (theoretical ply of SPB-40 and other homologous proteins **(Table 2.4)** were determined from, their amino acid sequences by use of proteomic tools. SPB-40 gene family proteins have molecular mass of approximately 40 kDa. The isoelectric points (pI) of these proteins are in the basic range of 8.6 to 9.1. The chitinase (Chit1) and chitinase- like protein (YM1) have pI in the acidic range.

Table 2.4 Physico-chemical characteristics of SPB-40 and homologous Proteins.

Protein	No of amino acids	Molecular Weight	Th. pI
SPB-40	361	40.5	8.5
SPG-40	361	40.5	9.1
SPC-40	361	40.5	8.6
SPP-40	361	40.5	8.7
SPS-40	361	40.5	9.0
HCGP/YKL-40	362	40.3	9.1
GP38k	362	40.5	8.6
BRP-39	360	40.8	8.8
YM1	378	42.4	5.5
Chit1	392	43.6	5.8

3 MATERIALS AND METHODS

3.1 Source

Milk was obtained from National Diary Research Institute, Karnal. The milk sample was stored in -20^{o}C till further processing was done.

3.2 Isolation

3.2.1 Removal of fats from the milk

The milk sample was centrifuged at 5000 rpm for 30 min at 4^{o}C. After centrifugation supernatant was taken. This supernatant was filtered through four layered muslin cloth to remove the insoluble material and the filtrate was again centrifuged at 10,000 rpm for 30 min at 4^{o}C. Centrifugation was done to remove the fatty contents from the milk.

3.3 Purification and Characterization

To purify SPB-40 Ion exchange chromatography and Gel filtration methods were used.

3.3.1 Ion Exchange Chromatography

Skimmed milk was further diluted twice with 50 mMTris-HCL, Ph 8.0. As we know protein of our interest i.e. SPB-40 is basic in nature (pI – 9.1), we planned to go for cation exchange column chromatography

Commercially available CM sephadex C-50 matrix (Sigma, USA) was used for cation exchange chromatography. The matrix was soaked in Milli-Q water for 2 hrs and then equilibrated with 50 mMTris-HCL buffer, pH 8.0. The matrix was then added to the diluted milk and stirred slowly for one hour manually. The gel was allowed to settle and the milk was decanted to remove the unbound. Gel was washed with an excess of 50 mMTris-HCL. pH 8.0 packed in column. (25×2.5cm) and washed the same buffer containing 0.1M NaCl, which facilitates the removal of impurities. The SPB-40 was then eluted with the same buffer containing .3M NaCl. The protein solution was wasdialysed against triple distilled water. The protein was again passed through a CM-Sephadex column (10×2.5cm) pre equilibrated with 50mM Tris-HCL, pH 8.0 and eluted with a linear gradient of 0.05-0.3M NaCl in the same buffer. The elution profile was recorded by

taking absorbance at 280 nm on spectrophotometer and four peaks were obtained. The 3rd peak corresponding to SPB-40 was pooled, which has shown two bands of 40 kDa on SDS-PAGE. The third peak containing SPB-40 was concentrated using an Amicon ultrafiltration cell.

3.3.2 Electrophoretic analysis

Sodium Dodecyl Sulphate-Poly Acryl amide Gel Electrophoresis (SDS-PAGE)

Proteins eluted from Ion-exchange chromatography was subjected to Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis (SDS-PAGE) analyses as described by Laemmli (1970). Electrophoresis was performed using multiple gel casters (G.E. health care). All the reagents are prepared according to the G.E. manual using chemicals obtained from Sigma. Table 3.1 provides formulations used for preparation of stacking and separating gels. The list of stock solutions and the composition are given Table 3.2. All the stock solutions were stored at 4^0C in brown bottles for a maximum period of 30 days. 10%-4% separating-stacking gel system was caste. The protein sample was denatured by boiling for 4 min. at 100^0 C in the presence of Laemmmli's sample buffer. Equal amount of protein sample was loaded to each lane and electrophoresis was carried out at a constant current of 20 mA. The run was stopped when the dye was 0.5-1.0 cm from the bottom of the slab and the gel was fixed in fixing solution (methanol: acetic acid: distilled water; 40:10:50) and stained overnight with 0.1% Coomassie Brilliant Blue R250

prepared in fixing solution.

Table 3.1: Details of composition of working solutions used for preparation of separating and stacking gels

Solutions	Seperating Gel SDS-PAGE 10%	Stacking Gel SDS-PAGE 4%
Monomer A	3.30 ml	1.33 ml
Buffer B (1.5M Tris-HCl, pH 8.8)	2.50 ml	-
Buffer C (0.5 M Tris-HCl, pH 6.6)	-	2.50 ml
Detergent D (10% SDS)	0.10 ml	0.10 ml
Catalyst E (10% Ammonium per Sulphate)	0.05 ml	0.05 ml
N,N,N,N tetra methyl ethyl diamine (TEMED)	5.00 µl	10.00 µl
DW	4.05 ml	6.10 ml

Table 3.2: Details of composition of the stock solutions used in electrophoresis.

Solution	Composition	Quantity
Monomer A	Acryl Amide	29.2g
	Bis-Acrylamide	0.8 g
	Distilled water (DW)	100.0ml
Buffer B	1.5M Tris-HCl, pH 8.8	18.15 g Tris-base
	DW	100 ml
Buffer C	0.5 M trisHCl, pH 6.6	18.15g Tris-base
	DW	10ml
Detergent D	10% SDS	10.00 g
	DW	100ml
Catalyst E	10% APS	100mg/1ml DW
	(freshly prepared)	
Sample	DW	4.0 ml
Buffer F	0.5M Tris-HCl pH 6.8	1.0ml

	Glycerol	0.8ml
	10% (w/v) SDS	1.6 ml
	2-β-mercaptoethanol	0.4 ml
	Bromophenol blue 0.50% (w/v)	0.2ml
Electrode Buffer	Tris –base	4.5 g
	SDS	1.5 g
	Glycine	22.8 g
	DW	1500 ml

3.3.3 Gel Filtration chromatography

Commercially available sephadex G-100 matrix was soaked in Milli-Q water for 2 hrs and then equilibrated with 0.5M NaCl in 50mM Tris-HCL buffer. The equilibrated matrix was used to pack the column for gel filtration chromatography. The protein solution pooled from 3rd peak of cation exchange chromatography was loaded onto the gel filtration chromatography column. Elution was carried with 0.5M NaCl in 50mM Tris –HCL buffer, pH 8.0. The elution profile was recorded by taking absorbance at 280 nm on spectrophotometer and two peaks were obtained. The 2nd peak corresponds to SPB-40, which has shown single band of 40 kDa on SDS-PAGE. This peak was pooled, dialyzed, concentrated by using ultrafiltration and used for further analysis.

3.3.4 N-Terminal sequence determination of SPB-40

The sequence of first 20 amino acids from N –terminal end of 40 kDa protein was determined by Edman degradation method using an automated protein sequencer (PPSQ-21,Shimadzu). The separated protein on the electrophoresed gel under reducing conditions was transferred on PVDF membrane (Mini Trans-Blot Electrophoretic transfer cell, Biorad) and located on the protein sequencer. The cyclic reaction in the Edman degradation was carried out by four- step process:

1. Coupling, 2. Cleavage, 3.Conversion, 4. Chromatography and the principle involved are given below:

In the coupling reaction, phenylisothiocyanate (PITC) reacts with the amino acid residue at the amino terminus under basic conditions (provide by n-methylpiperdine / methanol / water) to form a phenylthiocarbomyl derivatives (PTC- protein).Trifluoroacetic acid then cleaves off the first amino acid as its anilinothialinone derivative (ATZ- amino acid) and leaves the new amino terminus for the next degradation cycle. The ATZ amino acid is then removed by extraction with N-butyl chloride and converted to a phenylthiohydantion derivative (PTH- amino acid) with 25 % TFA / water. The PTH- amino acid is transferred to a reverse-phase C-18 column for detection at 270 nm. A standard mixture of 19 PTH- amino acids is injected onto the column for separation, which provides standard retention times of the amino acids for comparison with each Edman degradation cycle chromatogram.

3.4 Crystallization

The purified and characterized SPB-40 sample was used for crystallization experiments by using different methods.Crystallization is the (natural or artificial) process of formation of solid crystals from a uniform solution.Crystallization is also a chemical solid-liquid separation technique, in which mass transfer of a solute from the liquid solution to a pure solid crystalline phase occurs.

Process

The crystallization process consists of two major events, nucleation and crystal growth. Nucleation is the step where the solute molecules dispersed in the solvent start to gather into clusters, on the nanometer scale (elevating solute concentration in a small region), that becomes stable under the current operating conditions. These stable clusters constitute the nuclei. However when the clusters are not stable, they redissolve. Therefore, the clusters need to reach a critical size in order to become stable nuclei. Such critical size is dictated by the operating conditions (temperature, supersaturation, etc.). It is at the stage of nucleation that the atoms arrange in a defined and periodic manner that defines the crystal structure – note that "crystal structure" is a special term that refers to the relative arrangement of the atoms, not the macroscopic properties of the crystal (size and shape), although those are a result of the internalcrystal structure.

The crystal growth is the subsequent growth of the nuclei that succeed in achieving the critical cluster size. Nucleation and growth continue to occur simultaneously while the supersaturation exists. Supersaturation is the driving force of the crystallization; hence the rate of nucleation and growth is driven by the existing supersaturation in the solution. Depending upon the conditions, either nucleation or growth may be predominant over the other, and as a result, crystals with different sizes and shapes are obtained (control of crystal size and shape constitutes one of the main challenges in industrial manufacturing, such as for pharmaceuticals). Once the supersaturation is exhausted, the solid-liquid system reaches equilibrium and the crystallization is complete, unless the operating conditions are modified from equilibrium so as to supersaturate the solution again.

Many compounds have the ability to crystallize with different crystal structures, a phenomenon called polymorphism. Each polymorph is in fact a different thermodynamic solid state and crystal polymorphs of the same compound exhibit different physical properties, such as dissolution rate, shape (angles between facets and facet growth rates), melting point, etc. For this reason, polymorphism is of major importance in industrial manufacture of crystalline products.

Protein Crystallization

Proteins, like many molecules, can be prompted to form crystals when placed in the appropriate conditions. In order to crystallize a protein, the purified protein undergoes slow precipitation from an aqueous solution. As, a result, individual protein molecules align themselves in a repeating series of "unit cells" by adopting a consistent orientation. The crystalline "lattice" that forms is held together by non-covalent interactions (Rhodes, 1993).The goal of crystallization is usually to produce a well-ordered crystal that is lacking in contaminants and large enough to provide a diffraction pattern when hit with X-ray. This diffraction pattern can then be analyzed to discern the protein's three dimensional structure. Protein crystallization is inherently difficult because of the fragile nature of protein crystals. Proteins have irregularly shaped surfaces, which results in the formation of large channels with any protein crystal. Therefore, the non-covalent bonds that hold together the lattice must often be formed through several layers of solvent molecules (Rhodes, 1993). In addition to overcoming the inherent fragility of protein crystals, the successful production of X-ray worthy crystals is dependent upon a number of environmental factors because so much variation exists among proteins, with each individual requiring unique condition for successful crystallization. Therefore, attempting to crystallize a protein without a proven protocol can be very tedious. Some factors that require consideration are protein purity, pH, and concentration of protein, temperature, and precipitants. In order for sufficient homogeneity, the protein should usually be at least 97% pure. pH conditions are also very important, as different pH's can result in different packing orientations. Buffers, such as Tris-HCL, are often necessary for the maintenance of a particular pH (Branden and Tooze, 1999). Precipitants, such as ammonium sulphate or polyethylene glycol, are compounds that cause the protein to precipitate out of solution (Rhodes, 1993).

Vapor Diffusion

Two of the most commonly used methods for protein crystallization fall under the category of vapor diffusion. These are known as the hanging drop and sitting drop methods. Both entail a droplet containing purified protein, buffer, and precipitant being allowed to equilibrate with a larger reservoir containing similar buffers and precipitants in higher concentrations. Initially, the droplet of protein solution contains an insufficient concentration of precipitant for crystallization, but as water vaporizes from the drop and transfers to the reservoir, the precipitant concentration

increases to a level optimal for crystallization. Since the system is in equilibrium, these optimum conditions are maintained until the crystallization is complete (Rhodes, 1993; McRee, 1993).Simply put, the hanging drop method differs from the sitting drop method in the vertical orientation of the protein solution drop within the system. It is important to mention that both methods require a closed system, that is, the system must be sealed off from the outside using an airtight container or high-vacuum grease between glass surfaces. Figure 3.1 and 3.2 depict the hanging drop and sitting drop systems, respectively (Rhodes, 1993; McRee. 1993). In sitting drop method, the protein drop sits on a pedestal above the reservoir solution, as opposed to hanging.

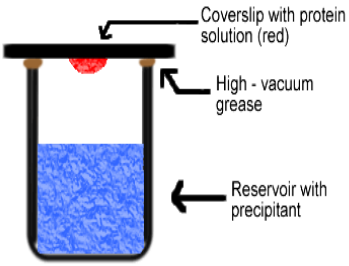


Figure 3.1: Hanging drop method

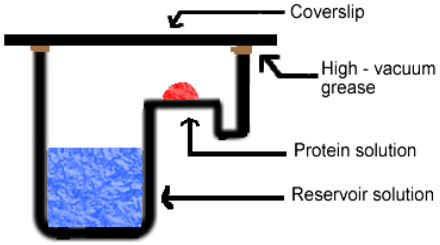


Figure 3.2: Sitting drop method

Reservoir solution (blue) usually contains buffer and precipitant. Protein solution (red) contains the same compounds, but in lower concentrations. The protein solution may also contain trace metals or ions necessary for precipitation of particular proteins. For instance, insulin is known to require trace amounts of zinc for crystallization (McRee, 1993).

High Through-Put Methods

One unavoidable aspect of crystallizing a protein from scratch is the need for a large number of experiments exploring the various conditions that are necessary for successful crystal growth. However, it is tedious to set up and screen so many different experimental conditions. Thus, high

through-put methods exist to help streamline this process. There are numerous kits available to order which apply preassembled ingredients in systems guaranteed to produce successful crystallization. Using such a kit, a scientist avoids the hassle of purifying a protein and determining the appropriate crystallization conditions. Another method that has proven successful in setting up a vast number of crystallizations involves robotics. Robotic crystallization systems use the same components described above, but carry out each step of the procedure quickly and with a large number of replicates. Each experiment utilizes tiny amounts of solution, and the advantage of the smaller size is two-fold: the smaller sample sizes not only cut-down on expenditure of purified protein, but smaller amounts of solution lead to quicker crystallizations. Each experiment is monitored by a camera which detects crystal growth (Eurekalert, 2000).

SWISS PDB Viewer

We need to install Swiss PDB Viewer from:http://spdbv.vital-it.ch/disclaim.html URL. After installation we need PDB file format of SPB-40 (4Q7N) from www.rscb.org/home/pdb/home.doSo here we were installed Swiss PDB viewer in my PC.

4 Results and Discussion

4.1. SDS Profile of crude bovine milk

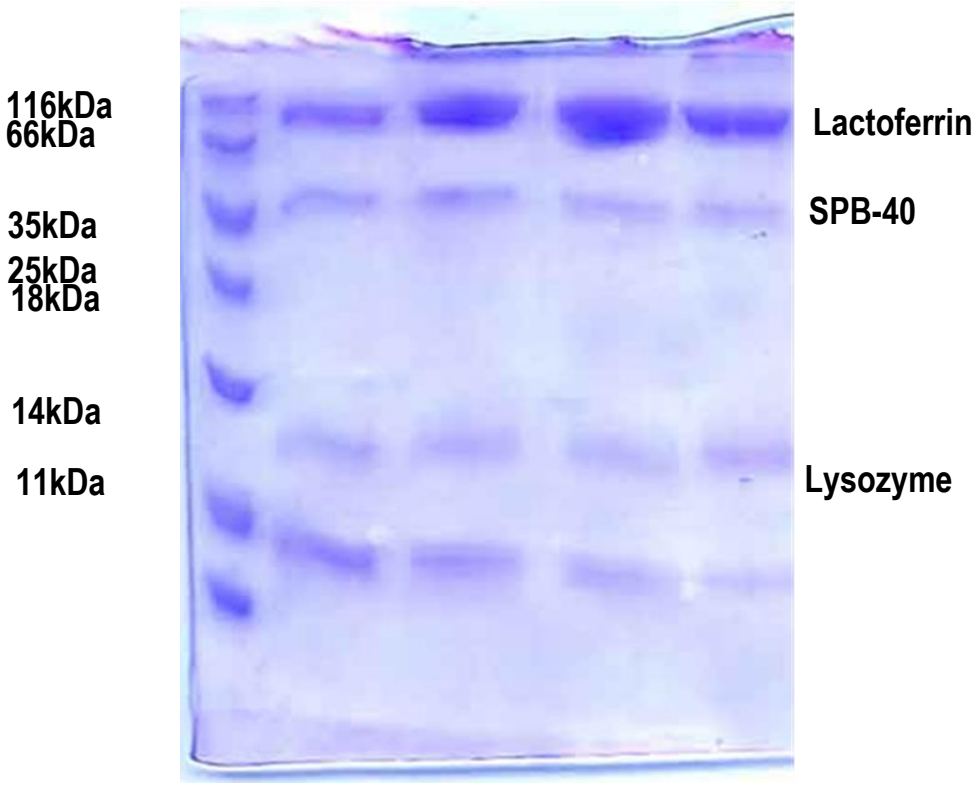


Figure 4.1: Complete Protein profile of Bovine milk

4.2. Ion Exchange Results:-

The elution profile of proteins bound to CMC-50 along the salt gradient gave following pattern: -

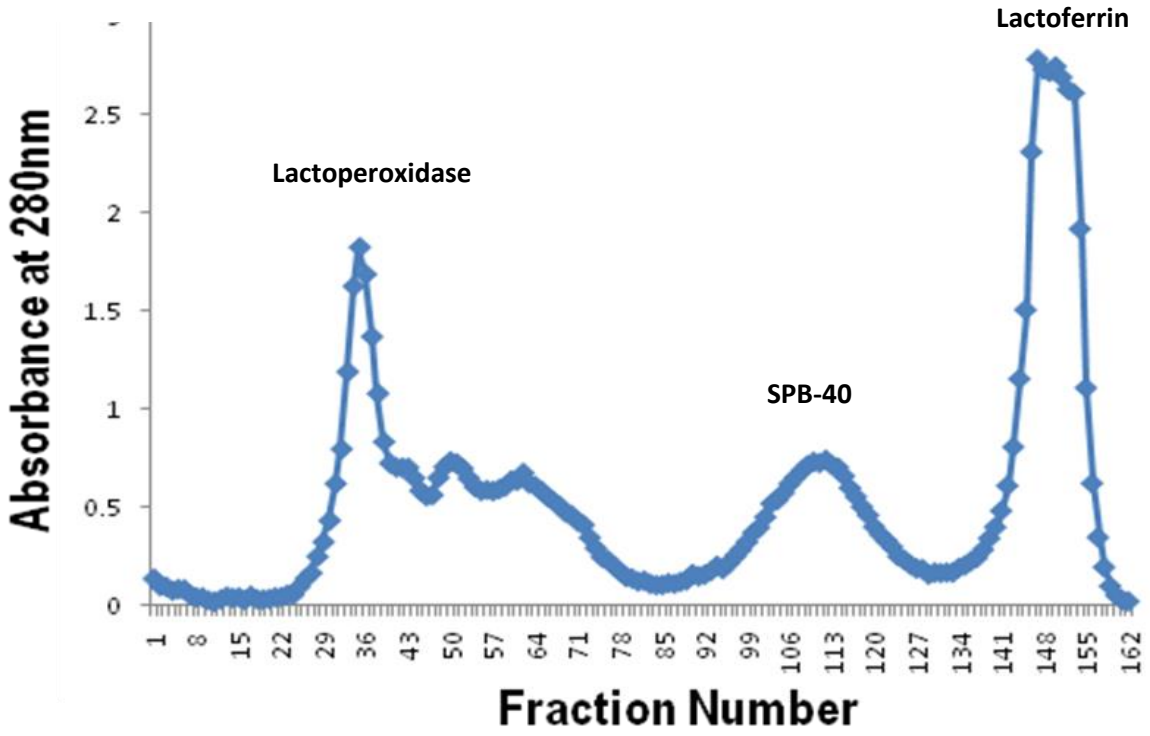


Figure 4.2: CM Sephadex C-50 Cation exchange profile

The protein got eluted nearly at 0.35 molar salt concentrations.

Peak 1 – Lactoperoxidase.

Peak 2 – SPB-40.

Peak 3– Lactoferrin.

Samples from peak 3 were run through SDS – PAGE gave following result –

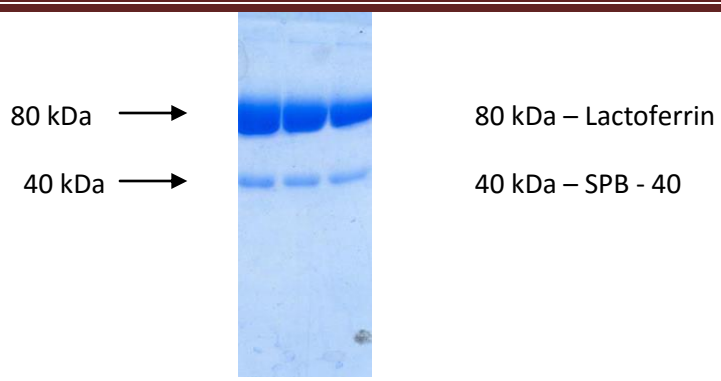


Figure 4.3: 10% SDS- PAGE profile of peak 3rd of Cation Exchange Chromatography.

This peak containing SPB – 40 was pooled out and kept for ultrafiltration and desalting by repeated washing using Amicon 50 ml unit and a 3kDa cut off membrane.

After desalting the sample was kept for Lyophilization to concentrate it for running through Sephadex G-100 (100 x 2 cm) column preequlibrated with Tris-HCl, pH 8.0 containing 0.5M NaCl. 2 peaks were obtained as shown below –

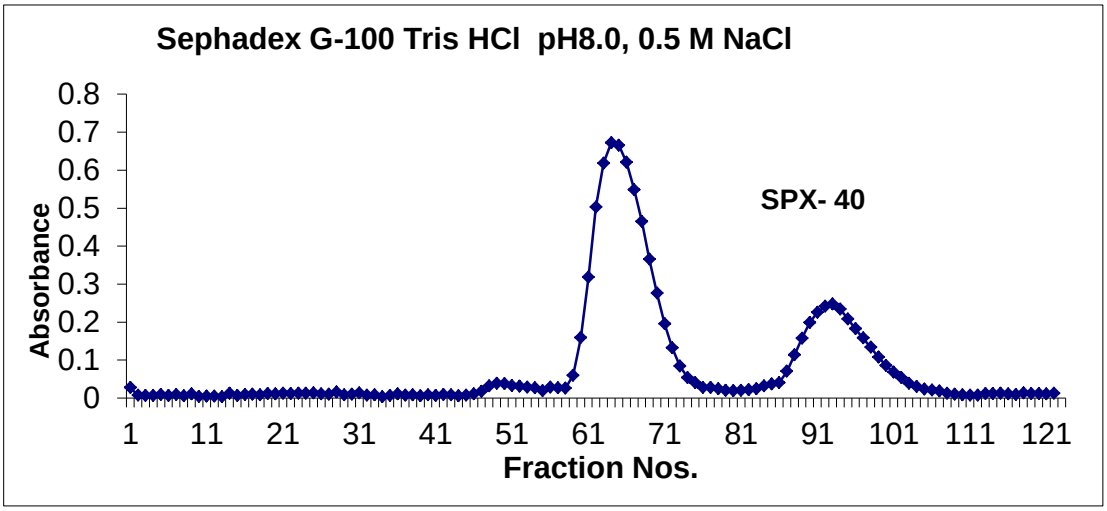


Figure 4.4: Sephadex G-100 gel filtration elution profile

The second peak in this final chromatographic step corresponds to a molecular mass of 40 kDa which was proved by running SDS-PAGE of the samples from the respective peak.

The following pattern was obtained –

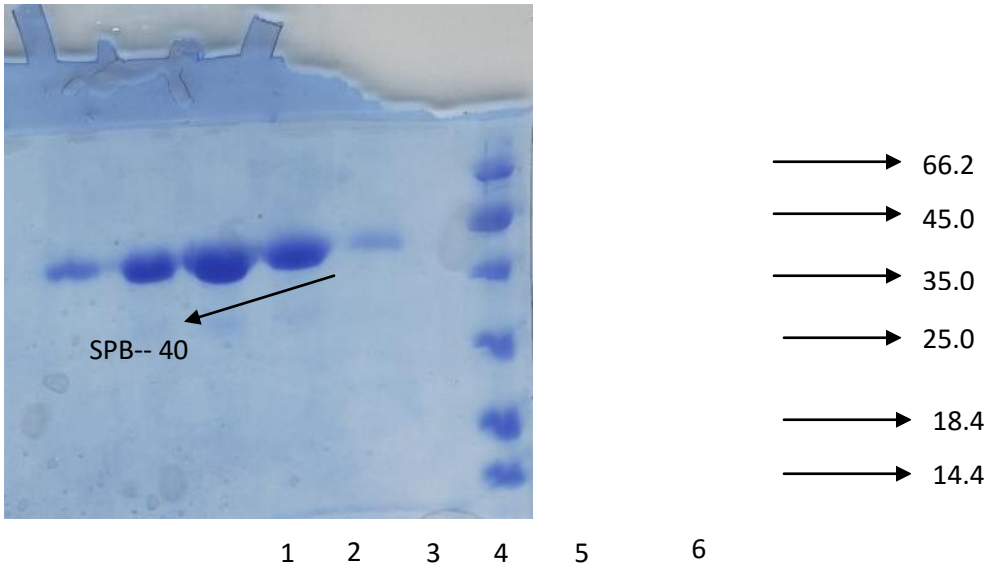


Figure 4.5: SDS-PAGE of 2nd peak

Lane 1,2,3,4 & 5 contains SPXB-40 at various concentrations.

Lane 6 – Molecular Weight Marker

This purified protein was then ultrafiltered and finally desalted by repeated washing with Mili-Q water. Now the Ultrafiltered protein was lyophilized for further use.

4.3 N–Terminal sequence determination

To confirm the identity of SPB-40, the N-terminal end of of 40kDa protein was sequenced for upto 15 amino acids. The obtained N –terminal sequence of 15 residues of SPB-40 was 100% in match with the known sequence of SPB-40.

Known Sequence **1-PAGPFRIPKCRKEFQ-15**

Obtained Sequence **1-PAGPFRIPKCRKEFQ-15**

Above results confirmed that the 40 kDa protein purified from Bovine milk was SPB-40. This purified and characterized protein sample was further used for crystallization experiments.

4.4 Crystallization of SPB-40

The purified Sample of SPB-40 were used for crystallization using Hanging drop method 30mg/ml of protein solution in 35mM TrisHCl pH 7.8 and 20 mMNacl was equlibrated with Buffer containing 20% ethanol at 298k. In crystallization experiments of the SPB-40, the protein was dissolved in the buffer to the concentration of 30 mg/ml. Thin square shaped crystals were obtained within 1 week.

The crystals obtained were as depicted below -

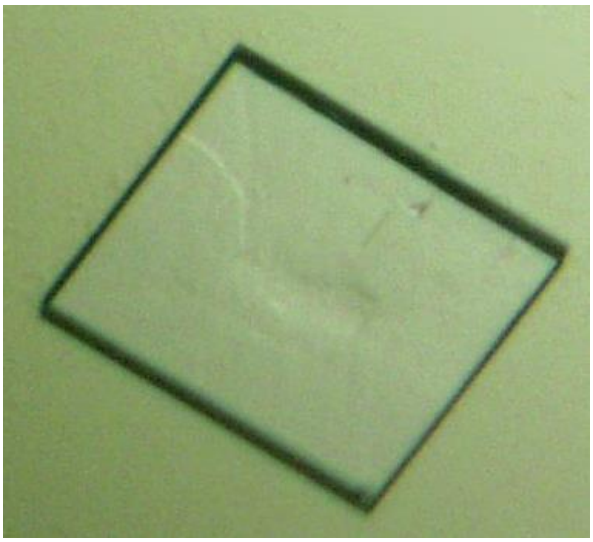


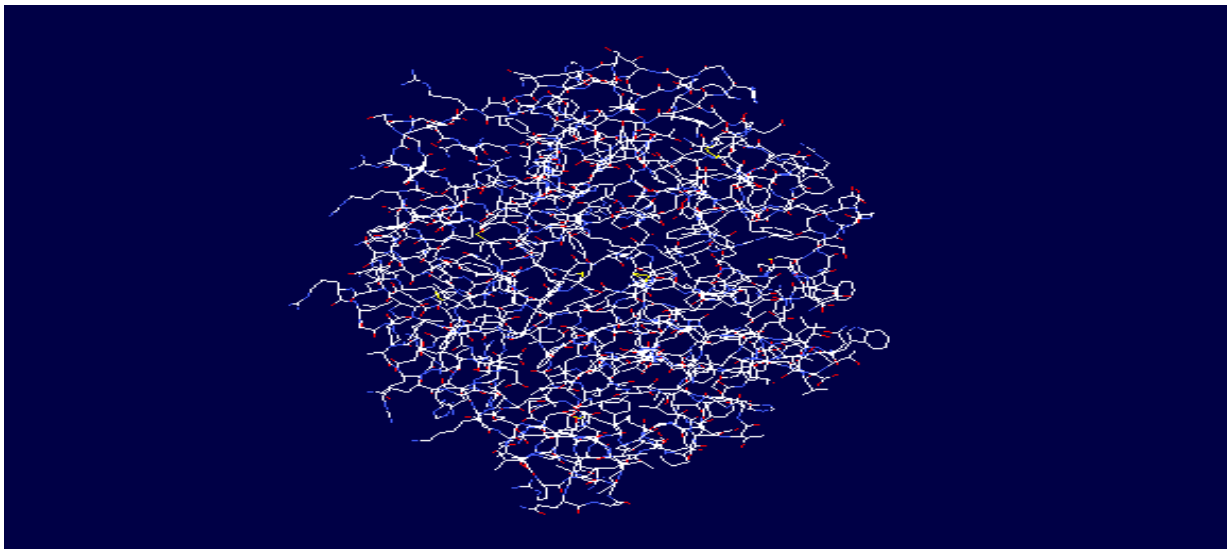
Figure 4.6: Crystal of SPB-40

4.5 SWISS PDB Viewer Result

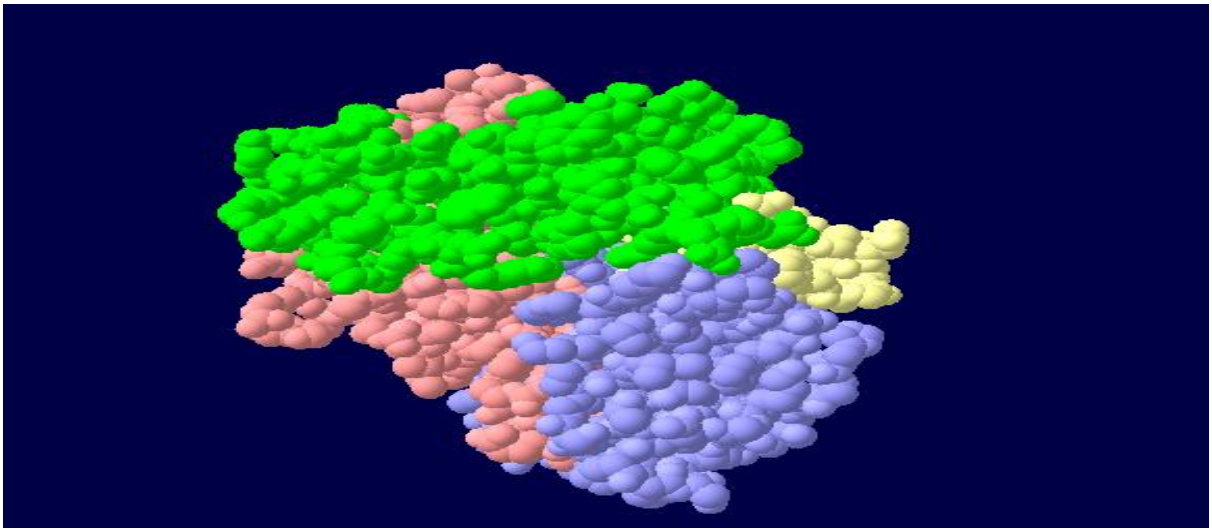
We were opened Swiss PDB viewer and select “file and open a PDB file” and then structure of SPB-40 was opened with control panel consists composition of amino acids. Now we were selected display and selected the render in 3D that showed the structure in STICKS form. When we were selected all amino acids Vander wall force of interaction in control panel and selected render structure in 3D, it showed BALL structure which shows different color of amino acid by selecting the color of amino acids in control panel.

- Then we were selected ribbons of all amino acids which show ribbon structure of SPB-40 in 3-D view.
- Then deselected all and go to control panel, select amino acid Tyrosine, select label in front of Tyrosine showed position of Tyrosine in the structure and also selected the Vander wall force of interaction on Tyrosine and hit Red color to show the Tyrosine molecules in the structure. When we were selected red color of Tyrosine in amino acids composition and selected the Vander wall force of interaction and then selected the render in 3D, then its shows Red color BALL structure of Tyrosine.
- For the torsion angle of Tyrosine, here we were removed bond and atoms of all amino acids expect Tyrosine and Valine. Now we selected torsion angle from status bar and pick and atom of Tyrosine (red in color) and valine (yellow color) where the torsion angle were found between them.
- Then we selected labels all amino acids, then opened tools on status bar and hit to compute H-Bond.
- We were also calculated force field energy, we need to open “color” and hit to calculate force field energy.
- For calculation of electrostatic potential of these two amino acids, we needed to go to “tools" and hit to “compute electrostatic potential”.
- To calculate distance between the molecules, we needed to select to distance symbol on the status bar, then pick atom to another which we want to calculate the distance and also we were calculated the angle between them.
- We were also computed Ramachandran Plot open “Wind” on status bar.

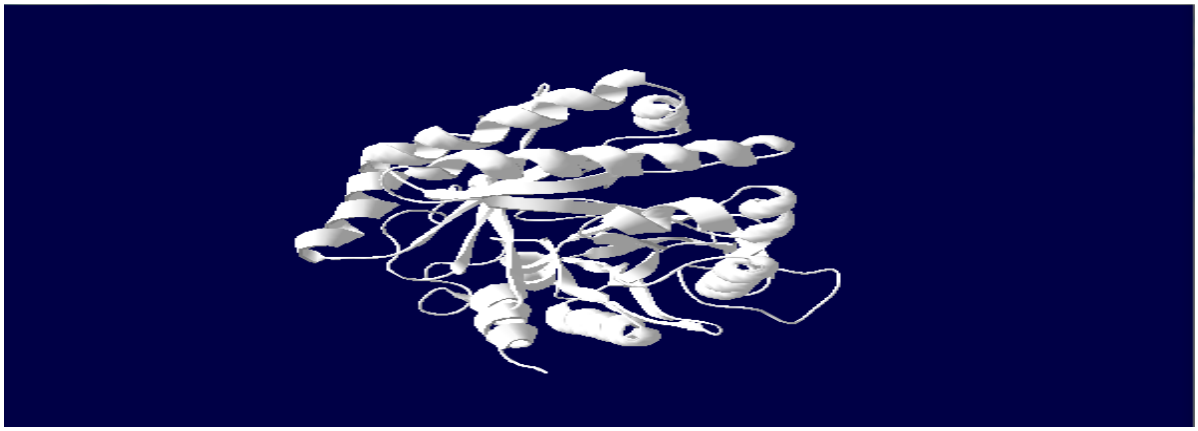
- We were downloaded PDB File of SPG-40 (2OLH), SPS-40 (2PI6), SPP-40 (1ZB5), SPB-40 (IOWQ), YM1(IVF8), Human Cartilege Gp39 (INWO) and CLP1 (2ESC) and loaded one by one by "open file" which showed the intertwined structures of protein. Now we were selected the "Wind" and then after hit "Alignment" which showed sequence similarity of these proteins.



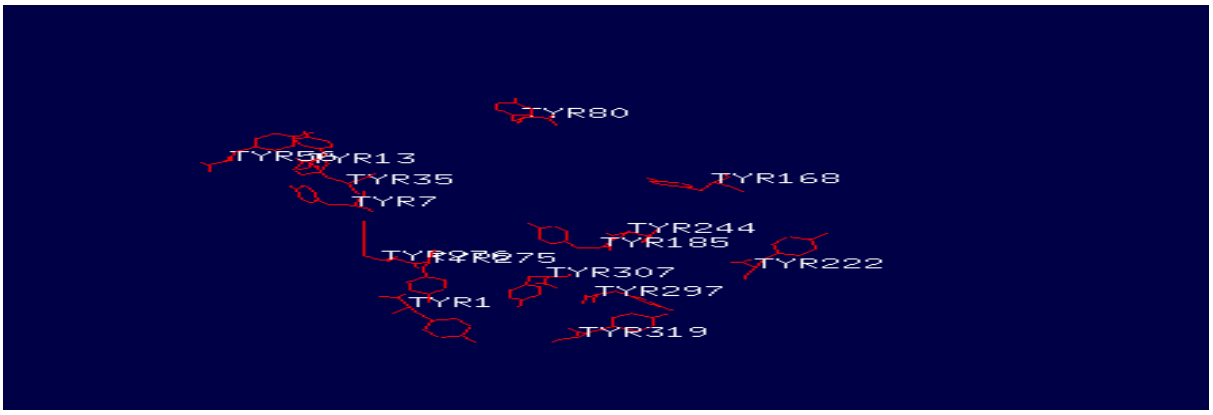
General structure of SPB-40



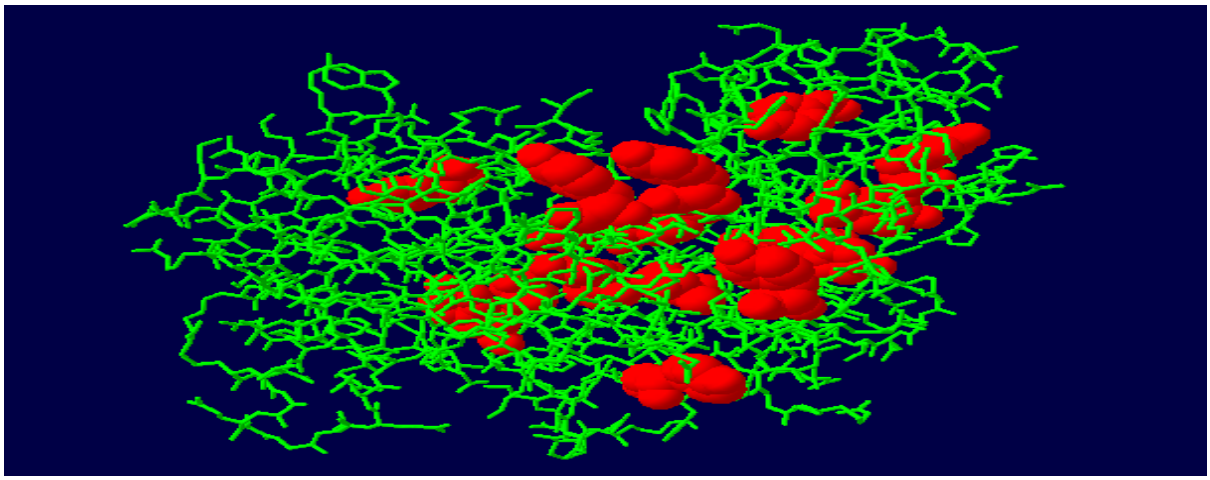
Ball structure of SPB-40



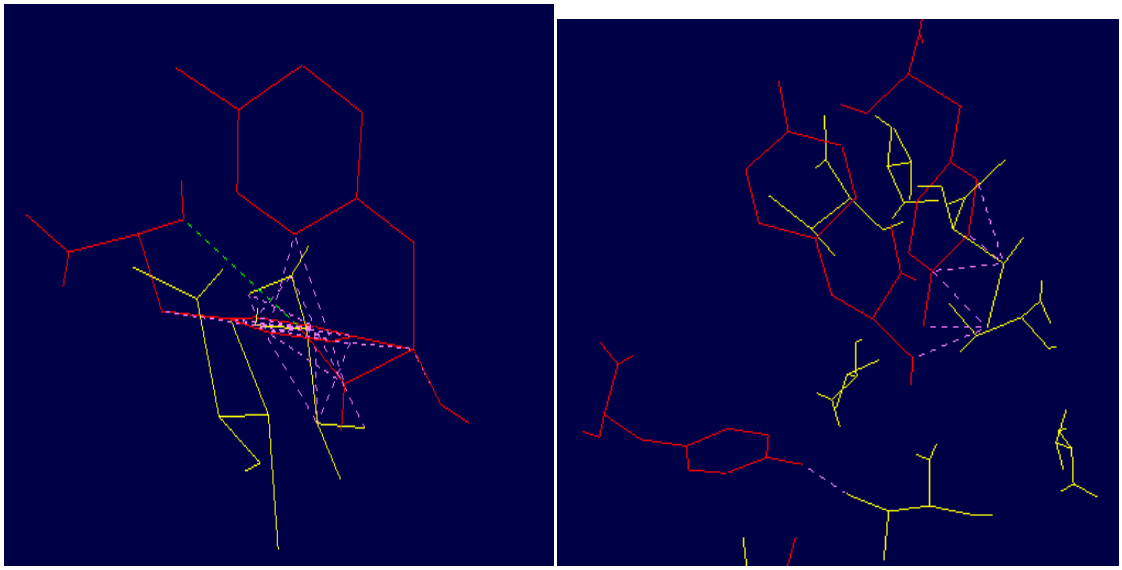
Ribbon structure of SPB-40



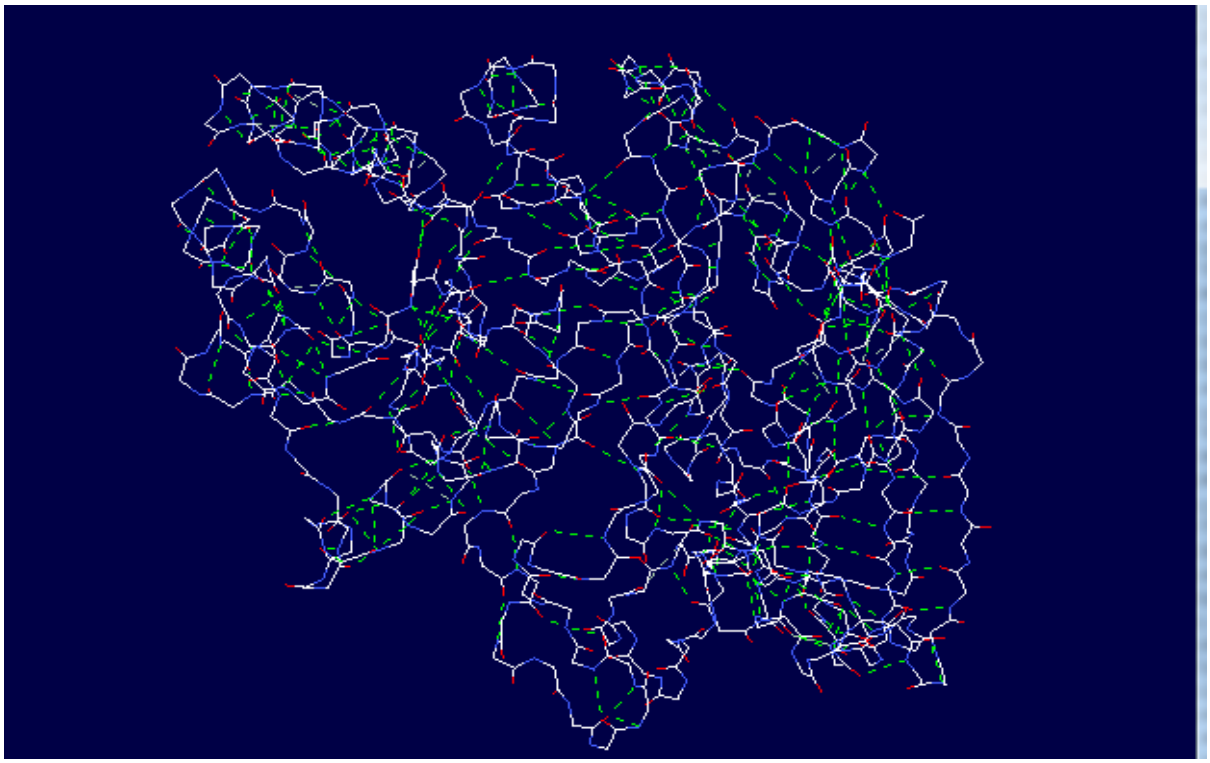
Red colour showing the structure position of Tyrosine in SPB-40



Red colour ball structure of Tyrosine in SPB-40 structure



Torsion angle between Tyrosine and Valine



Calculated H-Bond of SPB-40

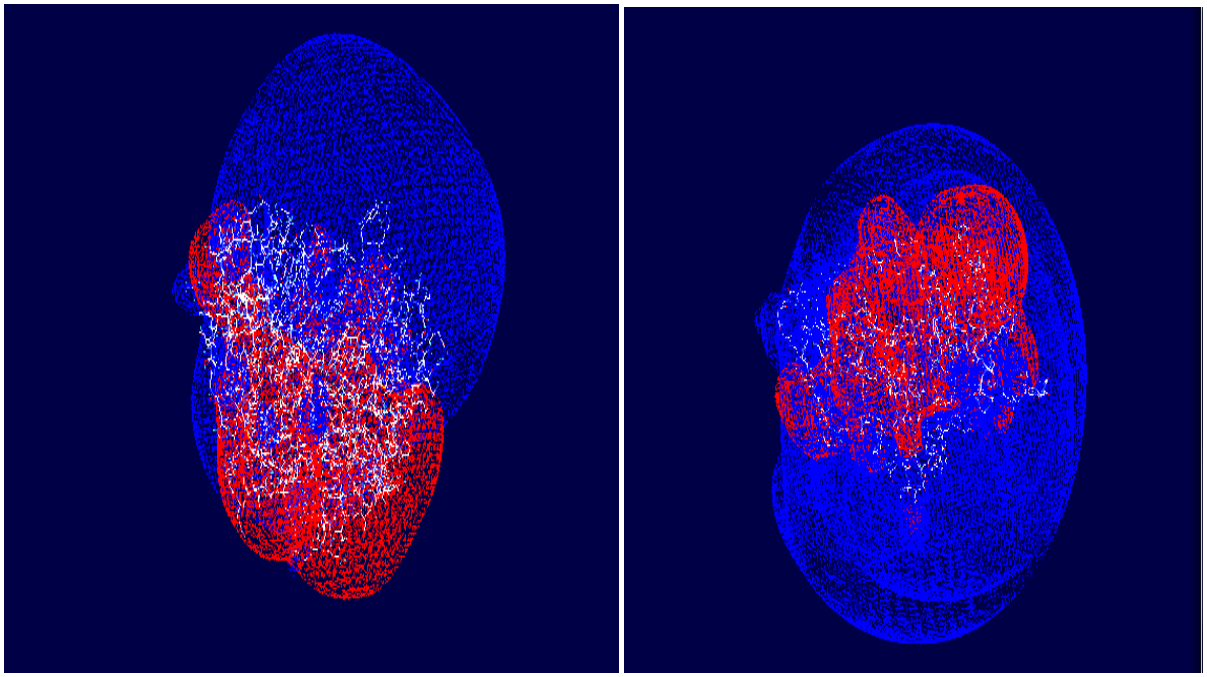
```
/ Computations were done in vacuo with the GROMOS96 43B1 parameters set, without reaction field.
/ For more information about GROMOS96, refer to: W.F. van Gunsteren et al. (1996) in Biomolecular
/ simulation: the GROMOS96 manual and user guide. Vdf Hochschulverlag ETHZ (http://igc.ethz.ch/gromos).
/ When using those results, please mention that energy computations were done with the GROMOS96
/ implementation of Swiss-PdbViewer.
/-------------------------------------------------------------------------------------------------------------------
/ residue          bonds       angles      torsion     improper   nonBonded  electrostatic constraint //      TOTAL
/-------------------------------------------------------------------------------------------------------------------
 HHT  A   1         0.000        6.183       7.540       0.000       0.00          0.00        0.0000 // E=      13.723
 TYR  A   1         3.192       13.673       0.238       1.423      80.12        251.87        0.0000 // E=     350.519
 LYSH A   2         5.577        3.811       5.628       4.461     -39.73        243.70        0.0000 // E=     223.442
 LEU  A   3         7.558       20.548       6.605      40.390     -47.29          0.00        0.0000 // E=      27.813
 ILE  A   4        19.695        7.635       6.897       2.716     -29.93          0.00        0.0000 // E=       7.007
 CYS1 A   5         6.132        3.879      11.309       2.994     -40.97          0.00        0.0000 // E=     -16.660
 TYR  A   6         9.855       14.591       5.013       6.270     -65.47          0.00        0.0000 // E=     -29.744
 TYR  A   7        20.329       16.859       4.586       5.808     -69.05          0.00        0.0000 // E=     -21.466
 THR  A   8         6.337        3.146       2.073       3.339     -40.50          0.00        0.0000 // E=     -25.602
 SER  A   9        20.905       18.338      15.637       3.371     -32.22          0.00        0.0000 // E=      26.028
 TRP  A  10        37.976       25.632       4.266      11.826     -56.28          0.00        0.0000 // E=      23.421
 SER  A  11         6.246        5.339       0.454       3.306     -33.16          0.00        0.0000 // E=     -17.819
 GLN  A  12         8.827        1.649       3.452       1.284     -39.97          0.00        0.0000 // E=     -24.754
 TYR  A  13        13.622        9.431       7.118       1.802     -37.10          0.00        0.0000 // E=      -5.128
 ARG  A  14        10.108        3.086       1.614       2.639     -45.00       -193.99        0.0000 // E=    -221.538
 GLU  A  15        10.144        5.458      11.527       0.313     -13.58       -233.91        0.0000 // E=    -220.045
 GLY  A  16         1.327        3.964       2.204       1.202     -10.71          0.00        0.0000 // E=      -2.010
 ASP  A  17         1.241        4.113       6.429       1.597     -19.47        -98.75        0.0000 // E=    -104.844
 GLY  A  18         3.667        1.614       1.652       2.967     -15.65          0.00        0.0000 // E=      -5.747
 SER  A  19         4.889        1.120       1.582       1.337     -25.30          0.00        0.0000 // E=     -16.372
 CYSH A  20         5.480        1.256       3.384       1.006     -44.32          0.00        0.0000 // E=     -33.197
 PHE  A  21         8.054       24.486       5.697       1.899     -39.94          0.00        0.0000 // E=       0.191
 PRO  A  22         5.603       29.763      16.022       1.744     -39.67          0.00        0.0000 // E=      13.458
 ASP  A  23         8.783        9.521       3.768       4.196     -10.67        -41.65        0.0000 // E=     -26.057
 ALA  A  24         2.166        2.062       5.156       1.662     -18.02          0.00        0.0000 // E=      -6.974
 ILE  A  25        12.565        1.896       4.723       1.553     -31.58          0.00        0.0000 // E=     -10.847
 ASP  A  26         2.184        6.445       5.736       0.091     -23.36       -204.68        0.0000 // E=    -213.589
 PRO  A  27        30.007       11.352      28.315       2.047     -32.77          0.00        0.0000 // E=      38.949
 PHE  A  28        16.539       20.377       4.100       2.821     -37.23          0.00        0.0000 // E=       6.606
 LEU  A  29        13.104        6.321       3.569       1.479     -40.72          0.00        0.0000 // E=     -16.250
 CYS2 A  30         5.641        3.126       2.075       1.656     -36.77          0.00        0.0000 // E=     -24.272
 THR  A  31         7.783        4.698       6.004       1.966     -26.76          0.00        0.0000 // E=      -6.312
 HISA A  32        15.800       10.678       4.513       2.637     -46.70          0.00        0.0000 // E=     -13.072
 VAL  A  33         3.424        4.132       2.564       3.281     -31.37          0.00        0.0000 // E=     -17.963
 ILE  A  34        25.834        3.792       2.795       0.820     -25.75          0.00        0.0000 // E=       7.488
 TYR  A  35        19.767       13.659       4.645       4.959     -73.11          0.00        0.0000 // E=     -30.082
 SER  A  36         7.868       11.655       5.292       1.405     -34.04          0.00        0.0000 // E=      -7.817
 PHE  A  37        21.000       18.873       3.106       3.803     -66.83          0.00        0.0000 // E=     -20.046
 ALA  A  38         2.574        2.057       3.921       0.144     -30.01          0.00        0.0000 // E=     -21.317
 ASN  A  39         7.316        3.815       8.273       2.153     -39.81          0.00        0.0000 // E=     -18.248
 ILE  A  40         9.697        8.094       7.053       0.567     -23.24          0.00        0.0000 // E=       2.174
 SER  A  41         0.519        4.319       2.848       0.059     -16.04          0.00        0.0000 // E=      -8.292
 ASN  A  42         4.082        2.959       4.008       3.241      -5.98          0.00        0.0000 // E=       8.313
 ASN  A  43         2.142        3.715       0.970       0.721     -22.93          0.00        0.0000 // E=     -15.380
 GLU  A  44         5.510        3.529       4.017       2.883     -25.13       -201.75        0.0000 // E=    -210.942
 ILE  A  45        10.947        5.695       2.436       4.956     -26.45          0.00        0.0000 // E=      -2.413
 ASP  A  46         5.120        2.178       7.400       2.171     -29.58        -88.69        0.0000 // E=    -101.401
 THR  A  47         4.012        3.759       3.648       0.145     -26.23          0.00        0.0000 // E=     -14.668
 TRP  A  48        43.673       22.039       3.585       4.077     -48.82          0.00        0.0000 // E=      24.560
 PHE  A 105        12.879       26.501       1.751       9.250     -53.93          6.35        0.0000 // E=       2.808
 LEU  A 106         9.119        7.947       1.069       4.662     -47.35        -11.53        0.0000 // E=     -36.082
 ARG  A 107        10.029      102.142      26.469       1.386     -10.22       -264.76        0.0000 // E=    -134.947
 THR  A 108         6.176        2.354       0.618       1.178      -2.63        -13.60        0.0000 // E=      -5.908
 HISA A 109        14.823       11.051       6.173       4.834     -43.04         37.79        0.0000 // E=      31.624
 GLY  A 110         1.524        2.305       3.710       0.745     -10.52         33.12        0.0000 // E=      30.880
 PHE  A 111        14.433       17.801       4.322       3.909     -74.67         16.08        0.0000 // E=     -18.119
 ASP  A 112         3.578       11.641       7.786       1.187     -30.23         43.40        0.0000 // E=      37.358
 GLY  A 113         4.077        2.036       2.120       3.150     -23.00         27.87        0.0000 // E=      16.246
 LEU  A 114         9.256        3.968       3.793       0.777     -54.68        -13.89        0.0000 // E=     -50.774
 ASP  A 115         3.880        7.586       6.584       0.579     -50.67        -15.65        0.0000 // E=     -47.688
 LEU  A 116         6.384        5.579       6.833       2.984     -47.85        -12.06        0.0000 // E=     -38.122
 ALA  A 117         7.158        1.767       1.989       0.521     -21.13        -12.11        0.0000 // E=     -21.805
 TRP  A 118        35.879       26.737       5.276       8.123     -75.45          9.56        0.0000 // E=      10.128
 LEU  A 119         6.458        6.227       1.715       0.301     -30.06         -3.51        0.0000 // E=     -18.868
 TRP  A 120        24.116       35.465       6.266       7.004     -33.08          9.55        0.0000 // E=      49.327
 PRO  A 121         6.461       12.330      23.681       0.267     -40.83         10.75        0.0000 // E=      12.659
 GLY  A 122         1.345        2.038       4.133       0.017     -12.08         33.11        0.0000 // E=      28.561
 TRP  A 123        47.469       25.793       0.097       7.571     -35.77         -6.32        0.0000 // E=      38.834
 ARG  A 124        14.140        7.152       4.243       5.248     -18.45       -242.65        0.0000 // E=    -230.319
 ASP  A 125         7.579        5.416       5.528       0.295     -31.15        -10.51        0.0000 // E=     -22.841
 LYSH A 126        25.091        4.333      10.355       0.232     -34.68         -6.32        0.0000 // E=      -0.986
 ARG  A 127        53.156       21.962      19.801       4.369     -22.82       -251.85        0.0000 // E=    -175.378
 HISA A 128        15.755        9.882       2.653       2.234     -42.33         -7.85        0.0000 // E=     -19.652
 LEU  A 129        24.116       10.076       5.673       4.710     -32.47         -7.31        0.0000 // E=       4.790
 THR  A 130        19.149        2.262       1.298       3.875     -22.07        -12.25        0.0000 // E=      -7.740
 THR  A 131         9.058        3.152       1.216       2.642     -21.01        -14.80        0.0000 // E=     -19.751
 LEU  A 132        13.587        8.065       3.086       3.257     -46.18        -13.80        0.0000 // E=     -31.979
 VAL  A 133        12.765        2.593       5.272       1.181     -30.82         -9.62        0.0000 // E=     -18.627
 LYSH A 134        18.069        4.969       9.607       2.220     -31.20         -8.83        0.0000 // E=      -5.171
 GLU  A 135         1.757        3.543       2.196       1.221     -25.40        -16.03        0.0000 // E=     -32.715
 MET  A 136         7.155        2.381       3.460       2.088     -53.48         -5.92        0.0000 // E=     -44.321
 LYSH A 137        11.599        5.744      11.645       4.082     -35.27         -9.19        0.0000 // E=     -11.390
 ALA  A 138        18.556        5.933       0.745       4.658     -23.43         -8.99        0.0000 // E=      -2.530
 GLU  A 139         3.230        3.508       1.746       0.409     -31.98        -11.39        0.0000 // E=     -34.480
 PHE  A 140        23.171       22.507       3.718       6.631     -47.76         -6.23        0.0000 // E=       2.044
 VAL  A 141         9.562        1.998       4.224       2.317     -12.22         -4.53        0.0000 // E=       1.361
 ARG  A 142        15.273       10.742      17.348       1.040     -31.66       -257.46        0.0000 // E=    -244.721
 GLU  A 143         7.057        7.011      11.624       0.391      -3.65         -7.73        0.0000 // E=      14.705
 ALA  A 144         3.999        1.417       0.509       0.060     -20.88         -6.46        0.0000 // E=     -21.356
 GLN  A 145         4.702        3.121       5.147       3.004     -14.50       -155.41        0.0000 // E=    -153.933
 ALA  A 146         7.788        2.949       4.024       0.103      -4.68         37.27        0.0000 // E=      47.462
 GLY  A 147         7.627        3.209       9.718       0.787      -1.56         34.73        0.0000 // E=      54.512
 THR  A 148         4.493        7.954       3.133       6.883      -9.19        -10.19        0.0000 // E=       3.081
 GLU  A 149         2.095        1.899       6.096       0.943      -8.11          5.66        0.0000 // E=       8.586
 GLN  A 150        10.763        3.038       4.223       0.452     -28.16       -150.13        0.0000 // E=    -159.812
 LEU  A 151        12.791        1.787       4.967       3.935     -52.87          3.77        0.0000 // E=     -25.619
 LEU  A 152        10.448        3.675       2.991       0.385     -40.83        -15.01        0.0000 // E=     -38.340
 LEU  A 153         3.894        5.731       3.745       0.347     -49.77        -13.25        0.0000 // E=     -49.308
 SER  A 154         4.763        0.887       4.877       0.341     -30.92        -19.84        0.0000 // E=     -39.898
 ALA  A 155        12.670        0.642       2.562       0.970     -29.89        -11.90        0.0000 // E=     -24.950
 ALA  A 156         4.024        2.283       4.971       0.169     -30.26         -0.74        0.0000 // E=     -19.555
 VAL  A 157        24.900        1.531       3.825       0.541     -35.41        -15.64        0.0000 // E=     -20.257
 THR  A 158         5.627        5.325       1.209       0.648      -5.90         -7.90        0.0000 // E=      -0.999
 ALA  A 159         0.794        2.135       4.590       2.260     -23.48         41.01        0.0000 // E=      27.305
 GLY  A 160         2.608        0.523       3.795       1.519     -14.47         29.37        0.0000 // E=      23.348
```

GLU	A	49	5.269	8.807	4.206	0.480	-38.64	-4.39	0.0000 // E= -24.261
TRP	A	50	40.894	23.223	1.625	5.048	-33.52	-17.22	0.0000 // E= 20.051
ASN	A	51	12.769	4.187	2.630	0.977	-21.24	-195.72	0.0000 // E= -196.398
ASP	A	52	3.622	5.849	5.802	2.949	-33.14	-16.53	0.0000 // E= -31.446
VAL	A	53	26.491	13.413	0.521	9.279	-14.38	-0.40	0.0000 // E= 34.929
THR	A	54	8.110	3.667	2.689	0.199	-11.81	-7.89	0.0000 // E= -5.034
LEU	A	55	8.096	17.743	8.162	12.860	-43.72	-8.97	0.0000 // E= -5.830
TYR	A	56	8.858	23.424	3.189	3.806	-63.13	-50.52	0.0000 // E= -74.381
ASP	A	57	4.522	2.658	3.032	0.113	-10.01	0.32	0.0000 // E= 0.638
THR	A	58	23.545	2.043	1.454	1.363	-28.02	1.26	0.0000 // E= 1.644
LEU	A	59	2.335	5.001	2.026	1.697	-42.00	-5.73	0.0000 // E= -36.674
ASN	A	60	25.869	10.175	2.881	7.403	-34.98	-173.11	0.0000 // E= -161.766
THR	A	61	8.509	12.433	13.474	13.809	7.70	-22.43	0.0000 // E= 33.493
LEU	A	62	5.075	8.152	1.449	1.716	-44.25	-13.91	0.0000 // E= -41.770
LYSH	A	63	21.812	7.761	5.671	1.999	-40.76	-1.12	0.0000 // E= -4.638
ASN	A	64	3.447	1.952	4.307	2.977	-11.49	-153.68	0.0000 // E= -152.489
ARG	A	65	7.582	9.620	26.782	2.172	-27.92	-248.21	0.0000 // E= -229.979
ASN	A	66	6.122	2.928	1.968	2.565	-32.28	-154.89	0.0000 // E= -173.594
PRO	A	67	17.730	22.333	22.860	0.167	-23.06	-21.83	0.0000 // E= 18.191
ASN	A	68	7.389	4.984	10.520	0.167	-17.57	-154.08	0.0000 // E= -148.592
LEU	A	69	7.744	3.687	0.787	2.745	-42.97	5.45	0.0000 // E= -22.558
LYSH	A	70	17.300	11.480	1.740	2.678	-42.39	-13.83	0.0000 // E= -23.022
THR	A	71	11.048	4.897	3.248	4.859	-35.99	-15.59	0.0000 // E= -27.529
LEU	A	72	9.152	8.202	7.610	4.064	-55.02	-10.38	0.0000 // E= -36.370
LEU	A	73	11.669	3.986	1.647	1.062	-51.85	-12.50	0.0000 // E= -45.984
SER	A	74	1.627	1.716	6.820	0.215	-36.91	-22.29	0.0000 // E= -48.818
VAL	A	75	4.869	32.771	8.036	42.340	-34.51	21.40	0.0000 // E= 74.910
GLY	A	76	2.699	3.521	2.208	4.909	-19.38	67.23	0.0000 // E= 61.192
GLY	A	77	2.502	2.205	4.156	0.173	-23.61	42.81	0.0000 // E= 28.241
TRP	A	78	46.473	24.682	13.014	5.229	-17.32	-2.78	0.0000 // E= 69.300
ASN	A	79	9.701	2.681	1.640	1.595	-11.06	-141.33	0.0000 // E= -136.777
TYR	A	80	27.339	18.325	5.126	4.406	-44.27	7.87	0.0000 // E= 18.799
GLY	A	81	9.740	25.303	2.212	7.826	-7.76	44.28	0.0000 // E= 81.601
SER	A	82	3.988	3.741	18.078	4.671	-19.59	12.86	0.0000 // E= 23.748
GLN	A	83	5.762	4.564	10.943	1.045	-21.88	-119.29	0.0000 // E= -118.857
ARG	A	84	6.571	8.962	6.141	3.559	-29.68	-252.35	0.0000 // E= -256.793
PHE	A	85	11.773	19.851	15.356	4.026	-63.65	3.37	0.0000 // E= -9.273
SER	A	86	1.578	3.932	0.922	2.196	-16.05	-1.87	0.0000 // E= -9.291
LYSH	A	87	9.573	1.608	20.816	1.868	-28.55	0.46	0.0000 // E= 5.780
ILE	A	88	5.204	7.612	13.985	2.019	-31.23	0.94	0.0000 // E= -1.469
ALA	A	89	13.949	3.608	1.686	0.058	-32.31	-3.26	0.0000 // E= -16.269
SER	A	90	1.239	4.371	6.325	0.641	-21.16	-4.87	0.0000 // E= -13.458
LYSH	A	91	21.722	13.598	11.702	0.604	-31.54	-5.54	0.0000 // E= 10.551
THR	A	92	4.284	12.022	11.884	12.593	-7.76	-11.56	0.0000 // E= 21.466
GLN	A	93	8.376	2.569	8.357	2.496	5.66	-151.41	0.0000 // E= -123.955
SER	A	94	4.004	2.755	5.577	0.702	-31.54	-13.12	0.0000 // E= -31.622
ARG	A	95	7.370	4.675	2.552	0.982	-46.66	-279.90	0.0000 // E= -310.988
ARG	A	96	3.589	8.878	25.020	0.604	18.24	-275.46	0.0000 // E= -219.130
THR	A	97	12.815	2.253	0.962	0.146	-22.13	-16.94	0.0000 // E= -22.890
PHE	A	98	13.609	18.716	4.546	5.013	-59.09	-7.14	0.0000 // E= -24.344
ILE	A	99	9.545	1.593	2.695	1.171	-14.75	-6.45	0.0000 // E= -6.201
LYSH	A	100	17.945	6.082	20.597	3.494	-30.19	-8.14	0.0000 // E= 9.790
SER	A	101	3.646	2.684	6.497	2.160	-23.34	-19.45	0.0000 // E= -27.802
VAL	A	102	19.973	8.130	8.208	6.181	-26.98	10.20	0.0000 // E= 25.711
PRO	A	103	13.842	29.235	17.287	1.127	-19.00	-11.90	0.0000 // E= 30.592
PRO	A	104	9.306	19.357	22.291	0.957	-13.92	-27.76	0.0000 // E= 10.238
PHE	A	105	12.879	26.501	1.751	9.250	-53.93	6.35	0.0000 // E= 2.808
LEU	A	106	9.119	7.947	1.069	4.662	-47.35	-11.53	0.0000 // E= -36.082
ARG	A	107	10.029	102.142	26.469	1.386	-10.22	-264.76	0.0000 // E= -134.947
THR	A	108	6.176	2.354	0.618	1.178	-2.63	-13.60	0.0000 // E= -5.908
HISA	A	109	14.823	11.051	6.173	4.834	-43.04	37.79	0.0000 // E= 31.624
GLY	A	110	1.524	2.305	3.710	0.745	-10.52	33.12	0.0000 // E= 30.880
PHE	A	111	14.433	17.801	4.322	3.909	-74.67	16.08	0.0000 // E= -18.119
ASP	A	112	3.578	11.641	7.786	1.187	-30.23	43.40	0.0000 // E= 37.358
GLY	A	113	4.077	2.036	2.120	3.150	-23.00	27.87	0.0000 // E= 16.246
LEU	A	114	9.256	3.968	3.793	0.777	-54.68	-13.89	0.0000 // E= -50.774
ASP	A	115	3.880	7.586	6.584	0.579	-50.67	-15.65	0.0000 // E= -47.688
LEU	A	116	6.384	5.579	6.833	2.984	-47.85	-12.06	0.0000 // E= -38.122
ALA	A	117	7.158	1.767	1.989	0.521	-21.13	-12.11	0.0000 // E= -21.805
TRP	A	118	35.879	26.737	5.276	8.123	-75.45	9.56	0.0000 // E= 10.128
LEU	A	119	6.458	6.227	1.715	0.301	-30.06	-3.51	0.0000 // E= -18.868
TRP	A	120	24.116	35.465	6.266	7.004	-33.08	9.55	0.0000 // E= 49.327
PRO	A	121	6.461	12.330	23.681	0.267	-40.83	10.75	0.0000 // E= 12.659
GLY	A	122	1.345	2.038	4.133	0.017	-12.08	33.11	0.0000 // E= 28.561
TRP	A	123	47.469	25.793	0.097	7.571	-35.77	-6.32	0.0000 // E= 38.834
ARG	A	124	14.140	7.152	4.243	5.248	-18.45	-242.65	0.0000 // E= -230.319
ASP	A	125	7.579	5.416	5.528	0.295	-31.15	-10.51	0.0000 // E= -22.841
LYSH	A	126	25.091	4.333	10.355	0.232	-34.68	-6.32	0.0000 // E= -0.986
ARG	A	127	53.156	21.962	19.801	4.369	-22.82	-251.85	0.0000 // E= -175.378
HISA	A	128	15.755	9.882	2.653	2.234	-42.33	-7.85	0.0000 // E= -19.652
LEU	A	129	24.116	10.076	5.673	4.710	-32.47	-7.31	0.0000 // E= 4.790
THR	A	130	19.149	2.262	1.298	3.875	-22.07	-12.25	0.0000 // E= -7.740
THR	A	131	9.058	3.152	1.216	2.642	-21.01	-14.80	0.0000 // E= -19.751
LEU	A	132	13.587	8.065	3.086	3.257	-46.18	-13.80	0.0000 // E= -31.979
VAL	A	133	12.765	2.593	5.272	1.181	-30.82	-9.62	0.0000 // E= -18.627
LYSH	A	134	18.069	4.969	9.607	2.220	-31.20	-8.83	0.0000 // E= -5.171
GLU	A	135	1.757	3.543	2.196	1.221	-25.40	-16.03	0.0000 // E= -32.715
MET	A	136	7.155	2.381	3.460	2.088	-53.48	-5.92	0.0000 // E= -44.321
LYSH	A	137	11.599	5.744	11.645	4.082	-35.27	-9.19	0.0000 // E= -11.390
ALA	A	138	18.556	5.933	0.745	4.658	-23.43	-8.99	0.0000 // E= -2.530
GLU	A	139	3.230	3.508	1.746	0.409	-31.98	-11.39	0.0000 // E= -34.480
PHE	A	140	23.171	22.507	3.718	6.631	-47.76	-6.23	0.0000 // E= 2.044
VAL	A	141	9.562	1.998	4.224	2.317	-12.22	-4.53	0.0000 // E= 1.361
ARG	A	142	15.273	10.742	17.348	1.040	-31.66	-257.46	0.0000 // E= -244.721
GLU	A	143	7.057	7.011	11.624	0.391	-3.65	-7.73	0.0000 // E= 14.705
ALA	A	144	3.999	1.417	0.509	0.060	-20.88	-6.46	0.0000 // E= -21.356
GLN	A	145	4.702	3.121	5.147	3.004	-14.50	-155.41	0.0000 // E= -153.933
ALA	A	146	7.788	2.949	4.024	0.103	-4.68	37.27	0.0000 // E= 47.462
GLY	A	147	7.627	3.209	9.718	0.787	-1.56	34.73	0.0000 // E= 54.512
THR	A	148	4.493	7.954	3.133	6.883	-9.19	-10.19	0.0000 // E= 3.081
GLU	A	149	2.095	1.899	6.096	0.943	-8.11	5.66	0.0000 // E= 8.586
GLN	A	150	10.763	3.038	4.223	0.452	-28.16	-150.13	0.0000 // E= -159.812
LEU	A	151	12.791	1.787	4.967	3.935	-52.87	3.77	0.0000 // E= -25.619
LEU	A	152	10.448	3.675	2.991	0.385	-40.83	-15.01	0.0000 // E= -38.340
LEU	A	153	3.894	5.731	3.745	0.347	-49.77	-13.25	0.0000 // E= -49.308
SER	A	154	4.763	0.887	4.877	0.341	-30.92	-19.84	0.0000 // E= -39.898
ALA	A	155	12.670	0.642	2.562	0.970	-29.89	-11.90	0.0000 // E= -24.950
ALA	A	156	4.024	2.283	4.971	0.169	-30.26	-0.74	0.0000 // E= -19.555
VAL	A	157	24.900	1.531	3.825	0.541	-35.41	-15.64	0.0000 // E= -20.257
THR	A	158	5.627	5.325	1.209	0.648	-5.90	-7.90	0.0000 // E= -0.999
ALA	A	159	0.794	2.135	4.590	2.260	-23.48	41.01	0.0000 // E= 27.305
GLY	A	160	2.608	0.523	3.795	1.519	-14.47	29.37	0.0000 // E= 23.348

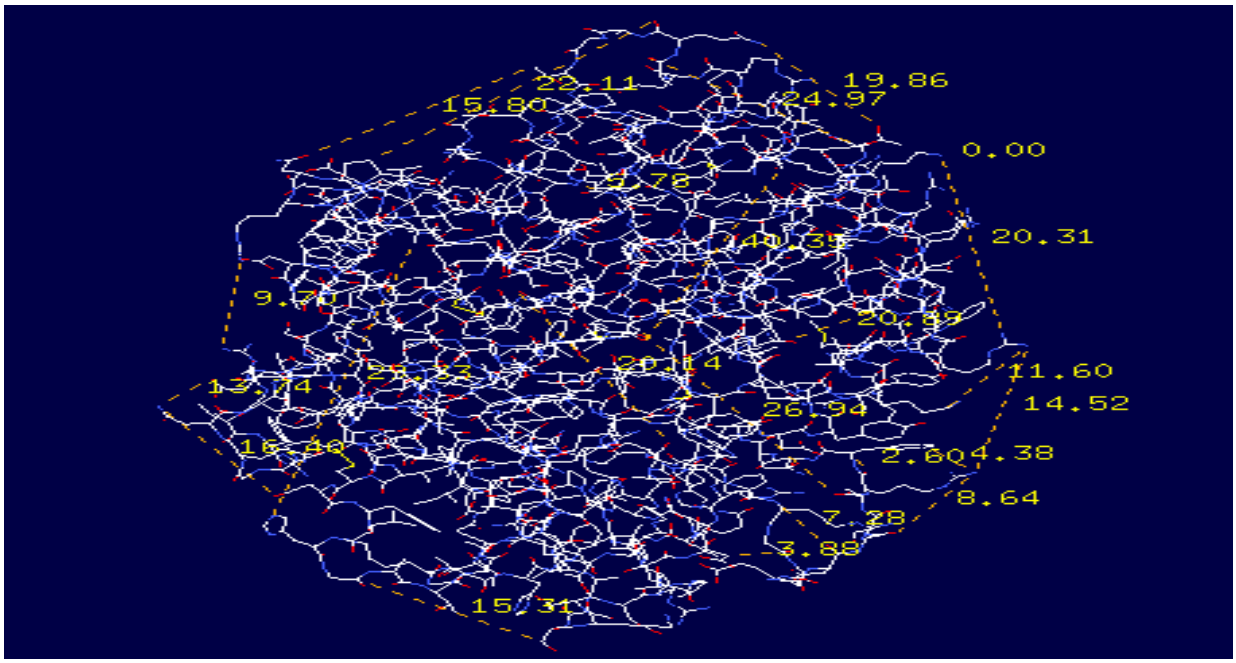
LYSH	A 161	13.478	3.699	5.393	0.117	-39.96	-9.58	0.0000 // E= -26.856
ILE	A 162	4.036	4.802	9.310	2.845	-3.92	-3.86	0.0000 // E= 13.222
ALA	A 163	2.444	3.236	1.252	0.789	-24.50	2.67	0.0000 // E= -14.106
ILE	A 164	16.240	1.578	4.164	1.740	-33.83	-0.60	0.0000 // E= -10.714
ASP	A 165	2.682	3.448	3.114	2.069	-22.62	-2.38	0.0000 // E= -13.693
ARG	A 166	12.846	10.210	8.325	0.573	-35.92	-216.81	0.0000 // E= -220.770
GLY	A 167	3.518	12.352	6.363	2.269	-26.80	35.76	0.0000 // E= 33.459
TYR	A 168	11.547	15.772	7.108	3.500	-64.83	-37.82	0.0000 // E= -64.729
ASP	A 169	6.103	4.185	4.819	2.871	-32.49	-7.32	0.0000 // E= -21.829
ILE	A 170	31.247	2.558	3.359	0.303	-13.09	2.54	0.0000 // E= 26.914
ALA	A 171	5.251	3.178	0.848	2.634	-20.79	10.21	0.0000 // E= 1.330
GLN	A 172	5.381	5.264	9.933	0.933	-9.91	-147.15	0.0000 // E= -135.548
ILE	A 173	22.215	7.752	1.105	6.325	-35.34	-9.44	0.0000 // E= -7.376
SER	A 174	4.448	5.155	0.190	2.519	-25.07	-8.38	0.0000 // E= -21.141
ARG	A 175	17.136	9.362	8.669	2.537	-2.93	-257.64	0.0000 // E= -222.873
HISA	A 176	10.392	11.168	6.192	0.615	-44.94	-3.09	0.0000 // E= -19.057
LEU	A 177	7.623	10.888	3.111	1.124	-44.87	8.50	0.0000 // E= -13.618
ASP	A 178	4.214	4.101	4.209	0.863	-28.01	5.05	0.0000 // E= -9.576
PHE	A 179	24.392	13.693	5.250	4.051	-61.77	-4.75	0.0000 // E= -19.129
ILE	A 180	7.673	6.416	8.372	0.192	-26.28	-15.85	0.0000 // E= -19.482
SER	A 181	5.573	2.062	4.538	1.668	-37.13	-20.65	0.0000 // E= -43.938
LEU	A 182	18.908	33.100	14.323	34.994	-48.11	-22.66	0.0000 // E= 30.549
LEU	A 183	12.325	4.868	8.381	1.279	-36.25	-18.91	0.0000 // E= -28.309
THR	A 184	15.421	5.280	15.411	0.296	-33.07	30.00	0.0000 // E= 33.330
TYR	A 185	17.256	18.183	4.583	5.263	-54.47	-39.39	0.0000 // E= -48.580
ASP	A 186	4.478	5.272	4.016	4.515	-14.47	1.27	0.0000 // E= 5.088
PHE	A 187	6.878	15.219	6.760	4.118	-61.88	-5.19	0.0000 // E= -34.095
HISA	A 188	18.397	10.491	5.140	0.374	-44.73	45.09	0.0000 // E= 34.770
GLY	A 189	0.577	2.716	2.757	0.094	-18.80	36.57	0.0000 // E= 23.908
ALA	A 190	7.193	3.261	2.005	4.044	-3.43	18.54	0.0000 // E= 31.614
TRP	A 191	39.291	24.809	1.620	4.047	-41.26	-6.80	0.0000 // E= 21.708
ARG	A 192	7.748	6.138	24.942	2.415	-6.73	-232.43	0.0000 // E= -197.914
GLN	A 193	7.365	9.330	11.150	5.702	-3.57	-164.83	0.0000 // E= -134.852
THR	A 194	8.047	2.468	2.080	1.818	-16.26	-10.61	0.0000 // E= -12.454
VAL	A 195	6.840	24.334	15.989	18.738	4.03	38.68	0.0000 // E= 108.608
GLY	A 196	0.367	2.489	1.503	2.398	-14.49	34.10	0.0000 // E= 26.365
HISB	A 197	18.665	8.160	10.944	2.657	-63.74	-23.26	0.0000 // E= -46.572
HISA	A 198	17.082	10.510	5.243	3.365	-39.58	-18.30	0.0000 // E= -21.683
SER	A 199	2.787	5.866	1.285	3.715	-12.25	0.89	0.0000 // E= 2.302
PRO	A 200	11.425	15.132	26.994	3.660	-36.25	-37.50	0.0000 // E= -16.540
LEU	A 201	13.706	3.722	2.161	3.751	-50.78	-13.79	0.0000 // E= -41.238
PHE	A 202	19.430	18.915	2.861	2.167	-49.98	8.89	0.0000 // E= 2.281
ARG	A 203	13.307	9.613	16.486	0.055	-26.91	-209.77	0.0000 // E= -197.212
GLY	A 204	5.997	5.217	7.919	1.012	19.80	51.00	0.0000 // E= 90.937
ASN	A 205	1.988	5.980	12.520	0.064	-4.99	-143.30	0.0000 // E= -127.732
*GLU	A 206	2.106	8.949	14.821	0.439	1036.99	7.38	0.0000 // E= 1070.684
*ASP	A 207	8.471	5.639	13.588	0.209	1154.12	0.76	0.0000 // E= 1182.791
ALA	A 208	0.641	3.598	1.973	3.897	-4.61	6.44	0.0000 // E= 11.931
SER	A 209	2.394	4.293	6.193	0.410	-6.42	6.34	0.0000 // E= 13.212
SER	A 210	7.869	5.506	4.964	0.304	9.39	3.60	0.0000 // E= 31.637
ARG	A 212	13.701	24.456	23.219	1.319	120.10	-245.44	0.0000 // E= -62.646
PHE	A 213	30.867	28.578	5.811	4.049	69.72	3.74	0.0000 // E= 142.764
SER	A 214	5.442	2.546	5.181	0.141	-22.14	-7.15	0.0000 // E= -15.979
ASN	A 215	2.880	2.915	2.180	0.713	-61.72	-179.11	0.0000 // E= -222.148
ALA	A 216	1.771	1.685	2.010	0.002	-28.59	-16.53	0.0000 // E= -39.655
ASP	A 217	4.251	2.038	3.254	3.020	-30.65	-14.37	0.0000 // E= -32.457
ASP	A 217	4.251	2.038	3.254	3.020	-30.65	-14.37	0.0000 // E= -32.457
TYR	A 218	8.756	20.939	1.719	3.861	-42.91	-53.69	0.0000 // E= -61.322
ALA	A 219	24.677	2.614	1.018	1.101	-31.12	-5.72	0.0000 // E= -7.428
VAL	A 220	14.305	2.577	3.641	0.339	-31.89	-9.15	0.0000 // E= -20.174
SER	A 221	2.514	1.125	1.242	1.271	29.31	-29.90	0.0000 // E= 5.564
TYR	A 222	9.886	14.051	5.787	2.923	-50.59	-57.74	0.0000 // E= -75.682
MET	A 223	3.473	2.736	4.758	1.252	-51.07	-12.25	0.0000 // E= -51.104
LEU	A 224	12.103	6.073	3.704	3.445	-41.19	-7.58	0.0000 // E= -23.448
ARG	A 225	1.846	20.248	10.955	0.143	115.66	-245.90	0.0000 // E= -97.054
LEU	A 226	14.310	30.915	22.554	13.554	-34.03	31.21	0.0000 // E= 78.516
GLY	A 227	0.836	7.368	2.983	0.533	-10.34	32.23	0.0000 // E= 33.609
ALA	A 228	5.798	3.617	5.481	1.927	-20.03	22.01	0.0000 // E= 18.795
PRO	A 229	3.674	12.952	22.890	0.673	-22.97	-33.07	0.0000 // E= -15.857
ALA	A 230	7.164	3.949	2.591	4.044	-19.98	-1.35	0.0000 // E= -3.585
ASN	A 231	14.972	3.801	5.770	2.123	-15.63	-154.11	0.0000 // E= -143.081
LYSH	A 232	21.430	5.786	10.246	2.842	-36.04	-15.38	0.0000 // E= -11.122
LEU	A 233	5.987	21.664	16.105	17.461	-54.95	-7.49	0.0000 // E= -1.219
VAL	A 234	12.702	3.868	2.337	0.081	-29.71	-15.78	0.0000 // E= -26.506
MET	A 235	12.721	5.149	7.577	3.598	-50.52	23.75	0.0000 // E= 2.272
GLY	A 236	7.250	1.559	1.115	0.443	-25.56	38.69	0.0000 // E= 23.493
ILE	A 237	6.166	5.275	3.515	1.004	-24.20	13.65	0.0000 // E= 5.414
PRO	A 238	17.869	11.869	20.973	1.427	-37.93	-23.84	0.0000 // E= -9.633
THR	A 239	55.965	8.296	6.429	3.410	-8.47	-19.43	0.0000 // E= 46.197
PHE	A 240	18.059	16.565	3.549	2.427	-53.89	39.82	0.0000 // E= 26.523
GLY	A 241	5.092	0.859	5.365	0.449	-20.08	20.78	0.0000 // E= 12.467
ARG	A 242	12.198	6.655	7.145	1.491	-38.56	-267.10	0.0000 // E= -278.172
SER	A 243	2.792	1.432	0.522	1.431	-30.81	-25.73	0.0000 // E= -50.363
TYR	A 244	11.695	9.124	1.467	2.141	-64.04	-49.96	0.0000 // E= -89.575
THR	A 245	1.854	1.244	6.934	1.183	-35.46	-23.21	0.0000 // E= -47.457
LEU	A 246	16.780	6.319	4.521	0.276	-48.94	-5.59	0.0000 // E= -26.633
ALA	A 247	1.633	1.167	2.505	0.396	-18.02	-2.32	0.0000 // E= -14.642
SER	A 248	5.567	4.491	9.145	0.000	-8.39	-12.61	0.0000 // E= -1.801
SER	A 249	0.905	1.560	8.274	2.797	1.46	-10.50	0.0000 // E= 4.497
LYSH	A 250	3.372	2.250	2.290	1.931	-22.01	8.52	0.0000 // E= -3.641
THR	A 251	4.417	7.825	16.792	2.153	-17.23	1.68	0.0000 // E= 15.630
ASP	A 252	5.622	11.857	18.620	0.476	-13.57	9.62	0.0000 // E= 32.624
VAL	A 253	3.498	1.669	10.798	0.436	-3.98	37.62	0.0000 // E= 50.044
GLY	A 254	0.988	6.535	3.849	0.013	-13.32	38.16	0.0000 // E= 36.230
ALA	A 255	2.561	1.843	1.384	0.324	-24.88	13.89	0.0000 // E= -4.878
PRO	A 256	12.620	14.762	16.574	0.902	-27.79	-37.54	0.0000 // E= -20.466
ILE	A 257	37.314	11.756	13.182	4.822	10.01	-3.77	0.0000 // E= 73.321
SER	A 258	4.571	4.728	4.774	1.618	-13.92	16.01	0.0000 // E= 17.783
GLY	A 259	2.054	2.276	3.005	0.010	-7.13	52.54	0.0000 // E= 52.763
PRO	A 260	14.469	19.961	14.156	0.716	-19.81	10.80	0.0000 // E= 40.293
GLY	A 261	5.401	0.309	2.249	0.588	-18.62	32.35	0.0000 // E= 22.270
ILE	A 262	15.632	4.842	3.257	2.328	0.67	23.62	0.0000 // E= 50.340
PRO	A 263	22.699	22.729	14.319	0.306	-24.13	12.10	0.0000 // E= 48.023
GLY	A 264	1.718	0.889	3.046	1.172	-16.08	41.90	0.0000 // E= 32.639
ARG	A 265	14.460	27.155	10.003	0.129	104.36	-242.70	0.0000 // E= -86.592
PHE	A 266	16.929	22.726	1.740	5.708	-52.02	17.58	0.0000 // E= 12.654
THR	A 267	7.883	3.596	3.046	1.081	-25.46	-9.29	0.0000 // E= -19.144
LYSH	A 268	9.790	9.614	16.260	1.054	-24.57	6.16	0.0000 // E= 18.311
TRP	A 269	50.426	22.767	9.416	3.978	-32.05	-11.83	0.0000 // E= 42.707
LYSH	A 270	17.739	5.919	7.857	2.101	-25.61	39.25	0.0000 // E= 47.248
GLY	A 271	2.422	2.237	4.258	0.294	-19.17	32.26	0.0000 // E= 22.293
ILE	A 272	7.202	5.539	8.006	1.407	23.23	-5.90	0.0000 // E= 39.485

LEU	A 273	9.281	9.622	3.608	3.071	-40.72	-10.87	0.0000 // E=	-26.009
ALA	A 274	1.478	3.414	3.250	0.293	-28.82	-10.17	0.0000 // E=	-30.559
TYR	A 275	16.528	8.935	3.042	5.171	-61.74	-53.88	0.0000 // E=	-81.948
TYR	A 276	3.086	17.467	0.467	4.350	-70.64	-40.99	0.0000 // E=	-86.259
GLU	A 277	10.759	6.758	3.395	2.412	-42.93	-19.61	0.0000 // E=	-39.212
ILE	A 278	21.322	3.639	1.692	1.245	-30.93	-3.43	0.0000 // E=	-6.469
CYS1	A 279	5.934	4.868	4.534	1.665	-34.62	8.92	0.0000 // E=	-8.704
ASP	A 280	4.320	5.180	5.717	5.172	-28.19	16.96	0.0000 // E=	9.161
PHE	A 281	9.675	20.877	0.612	3.220	-54.25	13.07	0.0000 // E=	-6.790
LEU	A 282	9.667	19.236	20.256	26.919	-31.39	-12.78	0.0000 // E=	31.908
HISA	A 283	18.775	7.501	3.310	1.747	-22.96	43.51	0.0000 // E=	51.879
GLY	A 284	1.448	3.790	4.230	0.111	-9.95	40.45	0.0000 // E=	40.069
ALA	A 285	7.152	0.772	2.747	0.800	-22.90	-4.00	0.0000 // E=	-15.427
THR	A 286	2.945	2.063	8.797	0.509	-19.22	-21.10	0.0000 // E=	-26.005
THR	A 287	2.375	1.304	3.320	0.877	-19.44	-15.40	0.0000 // E=	-26.962
HISA	A 288	15.215	10.773	3.368	0.239	-33.75	-15.62	0.0000 // E=	-19.772
ARG	A 289	25.848	63.070	31.739	4.173	-5.26	-258.83	0.0000 // E=	-139.261
PHE	A 290	18.897	13.594	13.835	2.816	-56.12	-10.14	0.0000 // E=	-17.117
ARG	A 291	2.309	2.213	21.634	2.392	-7.90	-238.20	0.0000 // E=	-217.553
ASP	A 292	5.411	6.541	7.271	0.024	10.95	18.63	0.0000 // E=	48.823
GLN	A 293	5.187	3.722	6.766	1.679	-46.35	-155.11	0.0000 // E=	-184.107
GLN	A 294	5.028	9.457	4.685	0.579	-20.98	-163.53	0.0000 // E=	-164.761
VAL	A 295	24.047	1.799	4.764	1.192	-30.23	15.96	0.0000 // E=	17.527
PRO	A 296	5.349	23.201	14.212	1.870	-46.33	-39.25	0.0000 // E=	-40.949
TYR	A 297	22.858	21.868	6.680	5.189	-63.72	-57.80	0.0000 // E=	-64.927
ALA	A 298	5.573	3.381	2.249	3.277	-23.25	-13.19	0.0000 // E=	-21.958
THR	A 299	1.769	3.176	2.932	4.358	-19.69	-23.37	0.0000 // E=	-30.823
LYSH	A 300	17.907	10.949	9.008	0.236	-35.75	22.72	0.0000 // E=	25.064
GLY	A 301	2.091	1.317	2.242	0.894	-7.73	39.13	0.0000 // E=	37.940
ASN	A 302	6.061	2.737	4.845	1.811	-22.01	-169.29	0.0000 // E=	-175.844
GLN	A 303	15.750	2.581	3.354	2.590	-45.53	-142.02	0.0000 // E=	-163.276
TRP	A 304	48.002	24.910	3.592	4.907	-82.98	-30.55	0.0000 // E=	-32.122
VAL	A 305	3.632	3.618	4.634	0.009	-38.55	-15.32	0.0000 // E=	-41.976
ALA	A 306	2.454	3.669	5.811	0.164	-30.82	-7.26	0.0000 // E=	-25.979
TYR	A 307	16.775	21.799	5.970	4.789	-67.23	-42.33	0.0000 // E=	-60.227
ASP	A 308	5.924	7.018	5.644	2.691	-43.86	-0.47	0.0000 // E=	-23.054
ASP	A 309	6.630	9.702	4.209	1.304	0.56	-15.86	0.0000 // E=	6.546
GLN	A 310	2.303	6.897	5.105	2.465	-30.86	-143.83	0.0000 // E=	-157.916
GLU	A 311	15.484	7.144	4.694	3.501	-22.25	11.89	0.0000 // E=	20.460
SER	A 312	3.083	4.364	4.415	5.013	-33.64	7.49	0.0000 // E=	-9.273
VAL	A 313	6.002	8.966	1.581	8.875	-38.28	3.17	0.0000 // E=	-9.689
LYSH	A 314	10.441	2.158	12.485	1.916	-41.67	-4.29	0.0000 // E=	-18.963
ASN	A 315	11.697	1.734	2.845	0.981	-35.77	-163.74	0.0000 // E=	-182.254
LYSH	A 316	14.070	4.564	3.953	0.331	-53.33	-8.28	0.0000 // E=	-38.693
ALA	A 317	1.819	1.195	0.805	1.483	-27.10	-3.38	0.0000 // E=	-25.180
ARG	A 318	19.628	5.009	9.971	3.021	-37.98	-256.62	0.0000 // E=	-256.966
TYR	A 319	8.003	22.051	2.038	2.878	-52.91	-55.72	0.0000 // E=	-73.658
LEU	A 320	13.370	9.802	6.128	6.166	-39.67	-4.44	0.0000 // E=	-8.643
LYSH	A 321	17.370	6.725	8.696	1.299	-25.75	-1.10	0.0000 // E=	7.243
ASN	A 322	5.909	3.677	1.838	2.492	-20.75	-145.37	0.0000 // E=	-152.205
ARG	A 323	23.895	18.297	21.723	0.376	21.88	-261.00	0.0000 // E=	-174.833
GLN	A 324	2.176	3.378	3.047	0.163	-12.77	-162.50	0.0000 // E=	-166.499
LEU	A 325	28.954	4.940	3.609	3.876	-45.64	4.01	0.0000 // E=	-0.257
ALA	A 326	6.840	0.959	1.555	1.214	-19.94	21.36	0.0000 // E=	11.986
GLY	A 327	7.156	1.318	2.159	2.328	-21.33	26.92	0.0000 // E=	18.553
ALA	A 328	7.032	2.198	0.587	0.401	-29.28	-13.41	0.0000 // E=	-32.472
SER	A 312	3.083	4.364	4.415	5.013	-33.64	7.49	0.0000 // E=	-9.273
VAL	A 313	6.002	8.966	1.581	8.875	-38.28	3.17	0.0000 // E=	-9.689
LYSH	A 314	10.441	2.158	12.485	1.916	-41.67	-4.29	0.0000 // E=	-18.963
ASN	A 315	11.697	1.734	2.845	0.981	-35.77	-163.74	0.0000 // E=	-182.254
LYSH	A 316	14.070	4.564	3.953	0.331	-53.33	-8.28	0.0000 // E=	-38.693
ALA	A 317	1.819	1.195	0.805	1.483	-27.10	-3.38	0.0000 // E=	-25.180
ARG	A 318	19.628	5.009	9.971	3.021	-37.98	-256.62	0.0000 // E=	-256.966
TYR	A 319	8.003	22.051	2.038	2.878	-52.91	-55.72	0.0000 // E=	-73.658
LEU	A 320	13.370	9.802	6.128	6.166	-39.67	-4.44	0.0000 // E=	-8.643
LYSH	A 321	17.370	6.725	8.696	1.299	-25.75	-1.10	0.0000 // E=	7.243
ASN	A 322	5.909	3.677	1.838	2.492	-20.75	-145.37	0.0000 // E=	-152.205
ARG	A 323	23.895	18.297	21.723	0.376	21.88	-261.00	0.0000 // E=	-174.833
GLN	A 324	2.176	3.378	3.047	0.163	-12.77	-162.50	0.0000 // E=	-166.499
LEU	A 325	28.954	4.940	3.609	3.876	-45.64	4.01	0.0000 // E=	-0.257
ALA	A 326	6.840	0.959	1.555	1.214	-19.94	21.36	0.0000 // E=	11.986
GLY	A 327	7.156	1.318	2.159	2.328	-21.33	26.92	0.0000 // E=	18.553
ALA	A 328	7.032	2.198	0.587	0.401	-29.28	-13.41	0.0000 // E=	-32.472
MET	A 329	30.450	3.509	9.198	1.663	-48.86	-11.30	0.0000 // E=	-15.334
VAL	A 330	18.912	3.276	2.210	0.016	-38.18	-13.44	0.0000 // E=	-27.211
TRP	A 331	63.694	26.491	3.506	6.612	-69.31	-22.92	0.0000 // E=	8.076
ALA	A 332	6.323	3.694	2.806	0.467	-25.38	-13.94	0.0000 // E=	-26.025
LEU	A 333	12.370	5.175	2.216	1.434	-44.06	-9.89	0.0000 // E=	-32.753
ASP	A 334	5.696	4.005	3.861	3.196	-37.17	2.40	0.0000 // E=	-18.008
LEU	A 335	11.962	6.158	3.737	0.971	-40.17	-7.80	0.0000 // E=	-25.143
ASP	A 336	10.263	4.055	12.530	2.329	-39.24	-7.47	0.0000 // E=	-17.541
ASP	A 337	21.655	5.361	3.440	3.531	-46.95	-3.21	0.0000 // E=	-16.175
PHE	A 338	21.979	14.228	1.796	4.600	3.68	0.64	0.0000 // E=	46.914
ARG	A 339	14.810	8.513	13.555	0.575	-37.35	-213.16	0.0000 // E=	-213.061
GLY	A 340	3.906	1.335	4.284	0.325	-18.64	43.50	0.0000 // E=	34.703
THR	A 341	3.577	5.975	4.630	0.049	-15.55	-10.14	0.0000 // E=	-11.464
PHE	A 342	21.389	23.064	4.300	1.907	-49.22	26.42	0.0000 // E=	27.863
CYS2	A 343	4.309	2.812	6.537	0.707	-26.61	47.52	0.0000 // E=	35.273
GLY	A 344	1.649	1.349	3.055	0.169	2.80	46.91	0.0000 // E=	55.937
GLN	A 345	14.678	15.022	13.125	8.934	-0.29	-144.05	0.0000 // E=	-92.580
ASN	A 346	5.281	5.492	13.528	0.112	-1.62	-159.27	0.0000 // E=	-136.478
LEU	A 347	6.568	25.473	16.201	4.502	-28.58	-0.34	0.0000 // E=	23.830
THR	A 348	15.797	4.915	1.451	4.055	46.74	-20.93	0.0000 // E=	52.029
PHE	A 349	15.398	13.236	3.755	4.570	-42.09	19.65	0.0000 // E=	14.514
PRO	A 350	8.044	19.656	19.633	1.351	-38.11	-22.90	0.0000 // E=	-12.320
LEU	A 351	20.245	1.970	4.483	2.296	-48.51	7.40	0.0000 // E=	-12.124
THR	A 352	7.170	4.235	1.476	4.479	-28.57	-16.36	0.0000 // E=	-27.571
SER	A 353	84.468	11.080	11.969	6.520	-28.61	-12.33	0.0000 // E=	73.102
ALA	A 354	6.650	1.061	1.269	1.805	-27.58	-2.88	0.0000 // E=	-19.674
ILE	A 355	15.259	3.237	3.853	1.948	-32.90	-1.56	0.0000 // E=	-10.157
LYSH	A 356	21.764	20.120	8.031	0.535	-21.25	-6.28	0.0000 // E=	22.924
ASP	A 357	6.503	10.918	2.826	4.320	-28.11	8.02	0.0000 // E=	4.471
VAL	A 358	4.816	3.211	3.290	0.731	-15.11	0.84	0.0000 // E=	-2.221
LEU	A 359	5.906	16.487	17.736	19.974	-42.41	5.34	0.0000 // E=	23.037
ALA	A 360	2.999	3.310	0.662	0.100	-15.61	-1.37	0.0000 // E=	-9.910
ARG	A 361	5.335	13.707	20.272	6.271	-24.64	-244.49	0.0000 // E=	-223.540
VAL	A 362	8.805	10.081	3.471	0.592	-1.83	61.85	0.0000 // E=	82.973
OXT	A 362	0.000	0.000	0.000	0.000	3.81	18.13	0.0000 // E=	21.939
NAG	A 401	0.000	0.000	0.000	0.000	0.00	0.00	0.0000 // E=	0.000
2ZO	A 402	0.000	0.000	0.000	0.000	0.00	0.00	0.0000 // E=	0.000
KJ/mol		4069.375	3170.211	2365.578	1043.025	-7736.34	-10608.18	0.0000 // E=	-7696.330

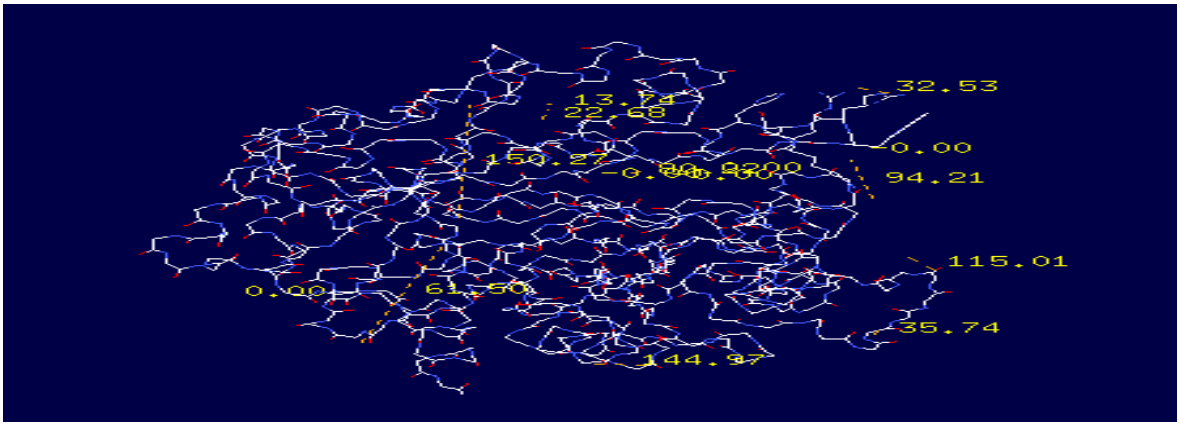
Calculated force field energy of SPB-40



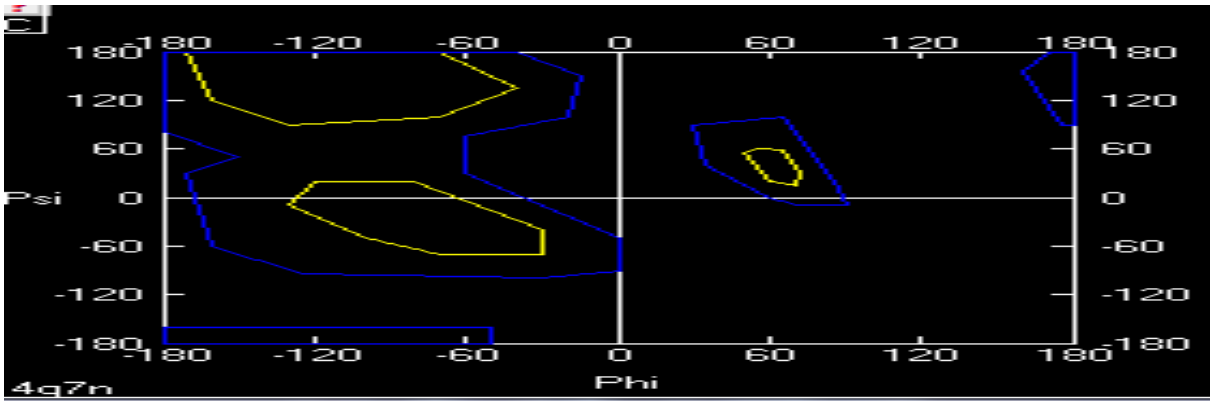
Calculated Electrostatic potential of SPB-40



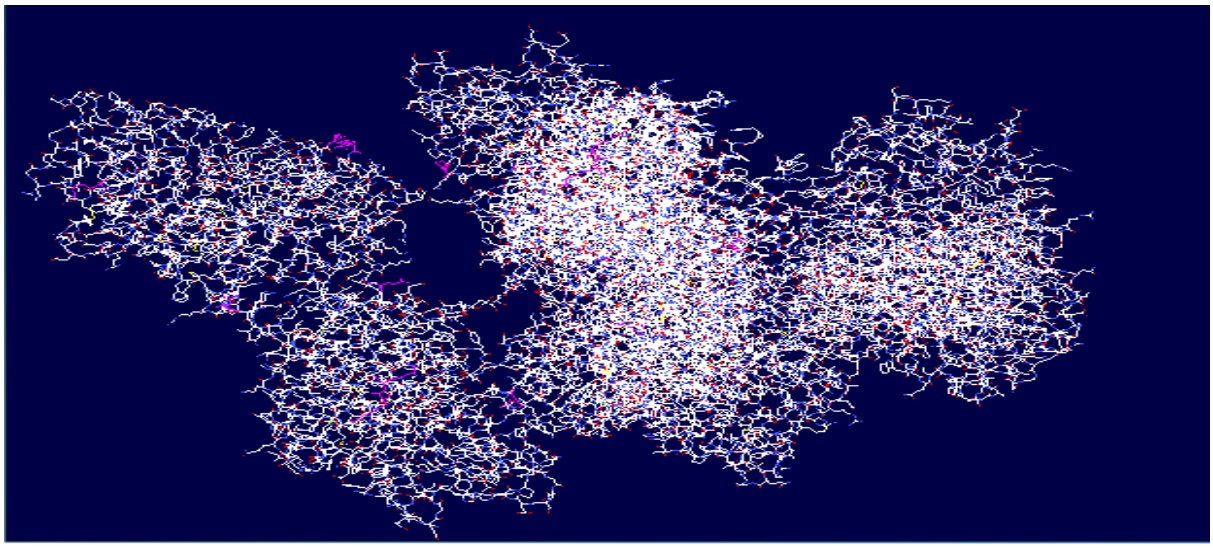
Calculated Atomic distance (in Å) of SPB-40



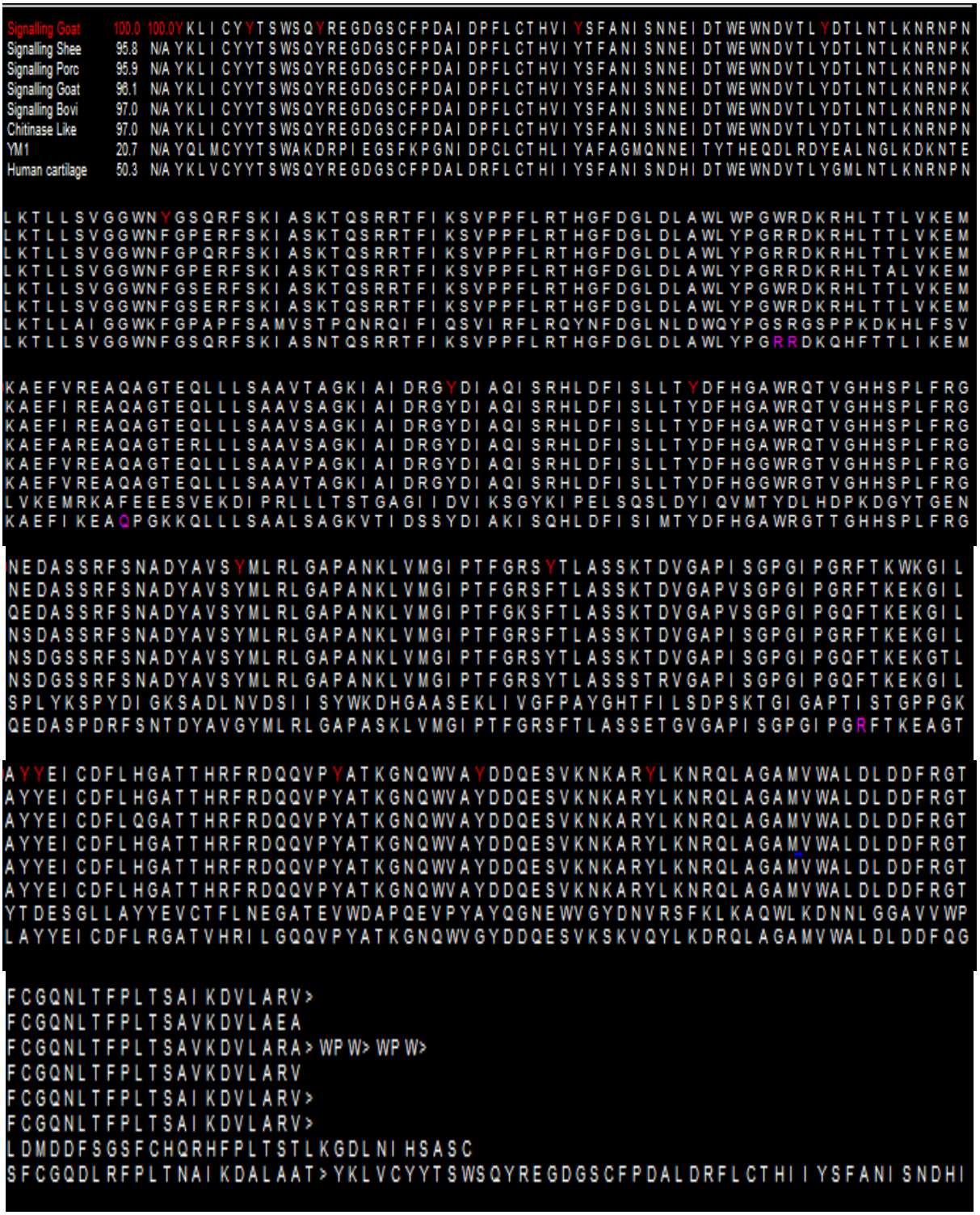
Calculated Bond angle of SPB-40



Ramachandran plot of SPB-40



Intertwines structure of SPB-40 and its homologues



Alignment of SPB-40 and its homologues protein

4.6 Discussion

The early dry period is the time of rapid functional and structural changes in the mammary gland. In the initial stages of the dry period, the mammary gland undergoes a process of involution. Loss of mammary cells by programmed cell death or apoptosis occurs in the mammary gland during involution of the mammary tissue structure throughout most of the involution period. SPB-40 is a glycoprotein, which is secreted during this period of mammary gland involution. The length of protein chain in SPB-40 is 361 residues. SPB-40 shows sequence identities of 46% and 25% with YM1 and Chit1 respectively. SPB-40 contains a beta barrel in its core with overall similar scaffolding to those observed in chitinasesand chitinases like proteins. The proteins of SPB-40 family are glycosylated and the glycosylation sites in all of them are conserved whereas the proteins of chitinases andchitinases like are not glycosylated. There is a possibility that SPB-40 may have evolved from chitinases to acquire new properties of neither binding to nor hydrolyzing chitin. The possibility exists that SPB-40 may instead be a primarily protein binding molecule.

REFERENCES

Alber, T., Banner, D.W., Bloomer, A.C., Petsko, G.A.,Phillips D., Rivers, P.S. and Wilson, I.A. (1981). On the three dimensional structure and catalytic mechanism of triose-phosphate isomerase. Philos. Trans. Lond.B.Biol. Sci.293 (1063), 159-171.

Alexander, C.M.,Selvarajan, S., Mudgett, J. and Werb, . (2001). Stromelysin-1 regulates apidogenesis during mammary gland involution. J. Cell Biol. 152, 693-703.

Arends, M.J., Morris, R.j. and Wyllie, A.H.(1990). Apoptosis.The role of endonuclease.Amer.J. Pathol. 36, 593-608.

Aslam and Hurley Aronson, N.N Blanchard, C.J. and Madura J.D. (1997).Homology modeling of glycosyl hydrolase family 18 enzymes and proteins.J.Chem.Inf.Comput.Sci. 37, 999-1005.

Aslam and Hurley, W.L. (1997a). Peptides generated from milk proteins in the bovine mammary gland during involution. J. Diary Sci. 81, 748-755.

Aslam and Hurley, W.L. (1997b). Proteolysis of milk proteins during involution of the bovine mammary gland. . J. Diary Sci. 80, 2004-2010.

Bauman, D.E. (1992). Bovine somatotropin:review of an emerging animal technology.J.Diary Sci. 75, 3432-3452.

Ethayathulla, A.S., Srivastava, D.B., Kumar, J., Singh, N., Sharma, S., Srinivasan, A. and Singh, T.P. (2006). Crystal structure of a secretory signaling Glycoprotein from buffalo at 2.8 A resolutions. (Under preparation).

Flint, D.J., Tonner, E., Beattie, J. and Panton, D. (1992).Investigation of mechanism of action of the growth hormone in stimulating lactation in rat. J. Endocr. 134, 377-383.

Flint, D.J., Tonner, E. and Allan, G.J. (2000). Insulin-like growth factor binding proteins:IGF-dependent and-independent effects in the mammary gland. J.Mammary Gland Biol.And Neoplasia 5(1), 65-73.

Hakala, B.E., White, C. and Recklies, A.D. (1993). Human Cartilage gp-39, a major secretory product of articular chondrocytes and synovial cells, ia a mammalian member of chitinase protein family.J.Biol.Chem. 268, 25803-25810.

.

Helminen, H.J and Ericsson, J.L.E. (1971).Effect of enforced milk stasis on mammary epithelium, with special reference to changes in lysosomes and lysosomal enzymes.Exp.Cell Res. 68,411-427.

Helminen, H.J. and Ericsson, J.L.E (1968a). Studies on mammary gland involution. II Ultrastructural evidence for auto-and heterophagocytisis. J. Ultrastruct. Res. 25, 214-227.

Helminen, H.J. and Ecricson, J.L.E. (1968b). Studies on mammary gland involution.III. Alterations outside auto-and heterophagocytic pathways for cytoplasmic degradation. J. Ultrastruct.Res. 25,, 228-239.

Holst, B.D., Hurley, W.L. and Nelson, D.R. (1987). Involution of the bovine mammary gland:histological an ultrastructuralchanges.J.Diary Sci. 70,935-941.

Hurley, W.L. (1987). Mammary function during the non lactating period, enzyme, lactose, protein concentrations, and pH of mammary secretions.J.Diary Sci. 69, 1642-1647.

Hurley, W.L. (1989), Warner, G.J. and Grummer, R.R. (1987). Changes in the triglyceride fatty acid composition of mammary secretions during involution.J.Diary Sci. 70, 2406-2410.

Johansen, J.S., Williamson, M.K., Rice, J.S. and Price, P.A. (1992).Identification of proteins secreted by human osteoblastic cells in culture.J.Bone. Miner. Res. 7, 501-512.

Johansen, J. S., Jensen, H. S. & Price, P.A. (1993).A new biochemical marker for joint injury.Analysis of YKL-40 in serum and synovial fluid. Br. J. Rheumatol.32, 949-955.

Kumar, J., Ethayathulla, A.S., Srivastava, D.B., Sharma, S., Singh, S.B., Srinivasan,A., Yadav, M.P., Singh, T.P.,(2006a). Crystal structure of a novel secretory signalling glycoprotein from bovine (SPC-40) at 2.1 Å resolution.ActaCryst. D62.

Li, M., Liu, X., Robinson, G., Bar-Peled, U., Wagner, K.U., Young, W.S., Henniighausen, L. and Furth, P.A. (1997). Mammary gland signals activate programmed cell death during the first stage of mammary gland involution. Proc. Natl. Acad. Sci. U.S.A. 94, 3425-3430.

Lund, L.R., Romer, J., Thomasset, N., Solberg, H., Pyke, C., Bissell, M.J., Dano, K. and Werb, Z. (1996). Two distinct phases of apoptosis in mammary gland involution:proteinase-independent and-dependent pathways. Development 122, 181-193.

Malinda, K.M., Ponce, L., Kleinman, H.K., Shackelton, L.M. and Millis, A.J. (1999). Gp38k, a protein synthesized by vascular smooth muscle cells, stimulates directional migration of human umbilical vein endothelial cells. Exp. Cell Res. 250, 168-73.

Mohanty, A.K., Singh, G., Paramasivam, M., Saravanan, K., Jabeen, T., Sharma, S., Yadav, S., Kaur. P., Kumar, P ., Srinivasan, A. & Singh, T.P. (2003).Crystal structure of a novel regulatory 40-kDa mammary gland protein (MGP-40) secreted during involution. J. Biol. Chem.278, 14451-14460.

Morrison, B. W. &Leder, P. (1994).neu and ras initiate murine mammary tumors that share genetic markers generally absent in c-myc and int-2-initiated tumors.Oncogene, 9, 3417 –3426.

Motyl, T., Gajkowska, B., wojewodzka, U,;Wareski, P., Rekiel, A. and Ploszaj, T. (2001).Expression of apoptosis-related proteins in involuting mammary gland of snow. Comp. Biochem. Physiol. Biochem. Mol. Bio. 128 (4), 635-646.

Perrakis, A., Tews, I., Dauter, Z., Oppenheim, A.B., Chet, I., Wilson, K.S. and Vorgias, C.E. (1994). Crystal structure of a bacterial chitinase at 2.3 A resolutions. Structure 2, 1169-1180.

Recklies, A.D., Whitw, C. and Andreesen, R. (1997).Molecular characterization of the gene for human cartilage gp-39 (CHI3I1), a member of the chitinase protein family and marker for late stages of macrophage differentiation. Genomics 43, 221-225.

Rejman, J.J. & Hurley, W.L. (1988).Isolation and characterization of a novel 39 kDa whey protein from bovine mammary secretions collected during the non-lactating periodBiochem. Bioph. Res. Commun., 150, 329–334

.

Renkema, G. H., Boot, R.G., Muijsers, A. O., Donker-Koopman, W. E. &Aerts, J.M.F.G. (1995).Purification and characterization of human chitotriosidase, a novel member of the chitinase family of proteins. J. Biol. Chem.270. 2198-2202.

Sendai, Y., Komiya, H., Suzuki, K., Onuma, T., Kikuchi.M., Hoshi, H. and Araki.Y. (1995).Molecular cloning and characterization of a mouse oviduct specific glycoprotein.Biol. Reprod.53(20,285-94).

Shackelton, L.M., Mann, D.M. & Millis, A.J. (1995).Identification of a 38-kDa heparin-binding glycoprotein (gp38k) in differentiating vascular smooth muscle cells as a member of a group of proteins associated with tissue remodelling.J. Biol. Chem., 270 , 13076–13083.

Srivastava, D.B., Ethayathulla, A.S., Kumar, J., Singh, N., Das, U., Sharma, S., Srinivasan, A. & Singh.T.P. (2006). Crystal structure of a signaling glycoprotein from sheep at 2.0 A resolutions. J.Struct.Biol.

Srivastava, D.B., Ethayathulla, A.S., Kumar, J.S., Singh, S.B. Singh, T.P. (2007). Crystal structure of a secretory Signaling Glycoprotein from porcine at 2.1 A resolutions. In the Press.

Steafanon, B., Colitti, M., Gabai, G., Knight, C.H. Wilde, C.J. (2002). Mammary apoptosis and lactation persistency in diary animals. J. Diary Res. 69 (1), 37-52.

Strange, R., Li, F., Saurer, S., Burkhardt, A. and Friis, R.R. (1992).Apoptotic cell death and tissue remodeling during mouse mammary gland involution. Development 115, 49-58.

Sun, Y.J., Chang, N. C., Hung, S.I., Chang, A.C., Chou, C.C. & Hsiao, C.D. (2001). The crystal structure of a novel mammalian lectin YM1, suggests a saccharide binding site. J.Biol. Chem.276, 17507-17514.

http://spdbv.vital-it.ch/disclaim.html

Tonner, E., Barber, M.C., Logan A. and Flint, D.J. (1997).Hormonal control of the insulin-like growth factor-binding protein-5 production in involuting mammary gland of rat.

Endocrinol. 138, 5101-5107.

Tsai, M.L., Liaw, S.H. & Chang, N.C. (2004). The crystal structure of Ym at 1.31 A resolution.J.Struct.Biol.148, 290-296.

Tucker, H.A. (1994). Lactation and its hormonal control. In: The physiology of Reproduction, EdsKnobil E, Neill JD, 2nd ed, Raven Press, NY, pp 1065-1098.

Varela, P.F., Llera, A.S., Mariuzza, R.A. and Tormo, J. (2002).Crystal structure of imaginal disc factor-2.A member of a new family of growth promoting glycoproteins from Drosophila melanogaster.J.Biol.Chem. 277, 13229-13236.

Walker, N.I., Bennet, R.E. and Kerr, J.F.R. (1989).Cell death by apoptosis during involution of the lactating breast in mice and rats.Am. J. Anat. 185, 19-32.

Wareski, P., Ryniewicz, Z., Orzechowski, A., Gajkowska, B., Wojewodzka, U. and Ploszaj, T. (2001).Expression of apoptosis-related proteins in mammary gland of goat. Small Rumin.Res.40 (3),279-289.

Wirenga, R.K. (2001). The TIM barrel fold:a versatile framework for efficient enzymes. FEBS Lett.492, 193-198.

Wilde, C.J., Henderson, A.J., Knight, C.H., Blatchford, D.R., Faulkner, A. and Vernon, R.G (1987).Effects of long term thrice dialy milking on mammary enzyme activity, cell population and milk yield in the goat. J. Anim. Sci. 64, 533-539.

Wilde, C.J and Knight, C.H (1989), Metabolic adaptations in mammary gland during the declining phase of lactation. J.Diary Sci. 72, 1679-1692.

Wilde, C.J and Peaker, M. (1990).Autocrine control in milk secretion. J. Agric.Sci. 114, 235 238.

Wilde, C.J., Addey, C.V.P., Li, P. and Fernig, D.G. (1997).Programmed cell death in bovine mammary tissue during lactation and involution.Exper. Physiol.82, 943-953.

Wyllie, A.H., Kerr, J.F.R. and Currie, A.R. (1980). Cell death: the significance of apoptosis. Int. Rev. Cytol.68, 251-306.

APPENDIX

SDS-PAGE Gel Electrophoresis

SDS-PAGE is a technique used in Biochemistry, Genetics and Molecular biology to separate proteins according to their electrophoretic mobility (a function of length of polypeptide chain or molecular weight as well as higher order protein folding, posttranslational modifications and other factors).

Procedure

The solution of proteins to be analyzed is first mixed with SDS, an anionic detergent which denatures secondary and non–disulfide–linked tertiary structures, and applies a negative charge to each protein in proportion to its mass. Without SDS, different proteins with similar molecular weights would migrate differently due to differences in folding, as differences in folding patterns would cause some proteins to better fit through the gel matrix than others. Adding SDS solves this problem, as it linearizes the proteins so that they may be separated strictly by molecular weight (primary structure, or number (and size) of amino acids). The SDS binds to the protein in a ratio of approximately 1.4 g SDS per 1.0 g protein (although binding ratios can vary from 1.1-2.2 g SDS/g protein), giving an approximately uniform mass: charge ratio for most proteins, so that the distance of migration through the gel can be assumed to be directly related to only the size of the protein. A tracking dye may be added to the protein solution to allow the experimenter to track the progress of the protein solution through the gel during the electrophoretic run.

Chemical ingredients and its roles

Polyacrylamide gel (PAG) had been known as a potential embedding medium for sectioning tissues as early as 1954. Two independent groups Davis and Raymond employed PAG in electrophoresis in 1959. It possesses several electrophoretically desirable features that made it a versatile medium. Polyacrylamide gel separates protein molecules according to both size and charge. It is a synthetic gel, thermo-stable,

transparent, strong, relatively chemically inert, can be prepared with a wide range of average pore sizes, can withstand high voltage gradients, feasible to various staining and destaining procedures and can be digested to extract separated fractions or dried for autoradiography and permanent recording. DISC electrophoresis utilizes gels of different pore sizes. The name DISC was derived from the discontinuities in the electrophoretic matrix and coincidentally from the discoid shape of the separated zones of ions (Anbalagan, 1999). There are two layers of gel, namely stacking or spacer gel, and resolving or separating gel.

- **Stacking gel**

The stacking gel is a large pore polyacrylamide gel (4%). This gel is prepared with Tris buffer pH 6.8 of about 2 pH units lower than that of electrophoresis buffer. These conditions provide an environment for Kohlrausch reactions, as a result, proteins are concentrated to several fold and a thin starting zone of the order of 19 μm is achieved in a few minutes. This gel is cast over the resolving gel. The height of the stacking gel region was always maintained more than double the height and the volume of the sample to be applied.

- **Resolving gel**

The resolving gel is a small pore polyacrylamide gel (3 - 30%). The Tris buffer used is of pH 8.8. In this gel, macro molecules separate according to their size. In the present experiment, 8%, 10% and 12% Resolving gel were used for separating different range of
proteins. 8% gel for 24 – 205 kD proteins, 10% gel for 14-205 kD proteins and 12% gel for 14-66 kD protein.

Chemical ingredients

- **Tris** (**tris (hydroxy methyl) aminomethane**) (**$C_4H_{11}NO_3$; mW: 121.14).**). It has been used as a buffer because it is an innocuous substance to most proteins. Its pKa is 8.3 at 20 °C and reasonably a very satisfactory buffer in the pH range 7.0 – 9.0.

- **Glycine(Amino Acetic Acid)** (**$C_2H_5NO_2$; mW: 75.07**). Glycine has been used as the source of trailing ion or slow ion because its pKa is 9.69 and mobility of glycinate are such that the effective mobility can be set at a value below that of the slowest known proteins of net negative charge in the pH range. The minimum pH of this range is somewhere around 8.0.
- **Acrylamide(C_3H_5NO; mW: 71.08)**. It is a white crystalline powder. While dissolving in water, autopolymerisation of acrylamide takes place. It is a slow spontaneous process by which acrylamide molecules join together by head on tail fashion. But in presence of free radicals generating system, acrylamide monomers are activated into a free-radical state. These activated monomers polymerise quickly and form long chain polymers. This kind of reaction is known as Vinyladdition polymerisation. A solution of these polymer chains becomes viscous but does not form a gel, because the chains simply slide over one another. Gel formation requires hooking various chains together. Acrylamide is a neurotoxin. It is also essential to store acrylamide in a cool dark and dry place to reduce autopolymerisation and hydrolysis.
- **Bisacrylamide(N,N'-Methylenebisacrylamide);($C_7H_{10}N_2O_2$;mW:154.17).** Bisacrylamide is the most frequently used cross linking agent for poly acryl-

amide gels. Chemically it is thought of having two-acrylamide molecules coupled . head to head at their non-reactive ends . Bisacrylamide was preserved at 4 °C.

- **Sodium Dodecyl Sulfate(SDS) ($C_{12}H_{25}NaO_4S$; mW: 288.38)**. SDS is the most common dissociating agent used to denature native proteins to individual polypeptides. When a protein mixture is heated to 100 °C in presence of SDS, the detergent wraps around the polypeptide backbone. It binds to polypeptides in a constant weight ratio of 1.4 g/g of polypeptide. In this process, the intrinsic charges of polypeptides become negligible when compared to the negative charges contributed by SDS. Thus polypeptides after treatment become a rod like structure possessing a uniform charge density that is same net negative charge per unit length. Mobilities of these proteins will be a linear function of the logarithms of their molecular weights.
- **Ammonium persulfate (APS)$N_2H_8S_2O_8$; mW: 228.2).**

APS is an initiator for gel formation.

- **TEMED(N,N,N',N'tetramethylethylenediamine)($C_6H_{16}N_2$; mW: 116.21**);
  Chemical for polymerisation of acrylamide gel is used for SDS-PAGE. It can be initiated by ammonium persulfate and the quaternary amine, N,N,N',N'-tetramethylethylenediamine (TEMED). The rate of polymerisation and the properties of the resulting gel depends on the concentration of APS and TEMED. Increasing the amount of APS and TEMED results in a decrease in the average polymer chain length, an increase in gel turbidity and a decrease in gel elasticity. Decreasing the amount of initiators shows the reverse effect. The lowest catalysts concentrations that will allow polymerisation in the optimal period of time should be used. APS and TEMED are used, approximately in equimolar concentrations in the range of 1 to 10 mM.

Chemicals for processing and visualization

The following chemicals are used for processing of the gel and the protein samples visualized in it:

- **Bromophenolblue (BPB**)(**3',3",5',5"tetrabromophenolsulfonphthalein**), **($C_{19}H_{10}Br_4O_5S$; mW: 669.99)**. BPB is the universal marker dye. Proteins and nucleic acids are mostly colourless.When they are subjected to electrophoresis, it is important to stop the run before they run off the gel. BPB is the most commonly employed tracking dye, because it is viable in alkali and neutral pH, it is a small molecule, it is ionisable and it is negatively charged above pH 4.6 and hence moves towards the anode. Being a small molecule it moves ahead of most proteins and nucleic acids. As it reaches the anodic end of the electrophoresis medium electrophoresis is stopped. It can bind with proteins weakly and give blue colour.

- **Glycerol ($C_3H_8O_3$; mW: 92.09).** It is a preservative and a weighing agent. Addition of glycerol (20-30 or 50%) is often recommended for the storage of enzymes. Glycerol maintains the protein solution at very low temperature, without freezing. It also helps to weigh down the sample into the wells without being spread while loading.
- **Coomassie Brilliant Blue R-250 (CBB) ($C_{45}H_{44}N_3NaO_7S_2$; mW: 825.97)**. CBB is the most popular protein stain. It is an anionic dye, which binds with proteins non-specifically. The structure of CBB is predominantly non-polar. So is usually used

(0.025%) in methanolic solution (40%) and acetic acid (7%). Proteins in the gel are fixed by acetic acid and simultaneously stained. The excess dye incorporated in the gel can be removed by destaining with the same solution but without the dye. The proteins are detected as blue bands on a clear background. As SDS is also anionic, it may interfere with staining process. Therefore, large volume of staining solution is recommended, approximately ten times the volume of the gel.

- **n-Butanol ($C_4H_{10}O$; mW: 74.12).** Water saturated butanol is used as an overlay solution on the resolving gel.
- **Dithiothreitol (DTT; $C_4H_{10}O_2S_2$; mW: 154.25).** DTT is a reducing agent used to disrupt disulfide bonds to ensure the protein is fully denatured before loading on the gel; ensuring the protein runs uniformly. Traditionally the toxic and less potent 2-mercaptoethanol was used.

SDS gradient gel electrophoresis of proteins

As voltage is applied, the anions (and negatively charged sample molecules) migrate toward the positive electrode (anode) in the lower chamber, the leading ion is $\underline{Cl}^-$ (high mobility and high concentration); glycinate is the trailing ion (low mobility and low concentration). SDS-protein particles do not migrate freely at the border between the Cl^- of the gel buffer and the $\underline{Gly}^-$ of the cathode buffer. Friedrich Kohlrausch found that Ohm's law also applies to dissolved electrolytes. Because of the voltage drop between the Cl^- and Glycine-buffers, proteins are compressed (stacked) into micrometer thin layers. The boundary moves through a pore gradient and the protein stack gradually disperses due to an frictional resistance increase of the gel matrix. Stacking and unstacking occurs continuously in the gradient gel, for every protein at a different position. For a complete protein unstacking the polyacrylamide-gel concentration must exceed 16% T. The two-gel system of "Laemmli" is a simple gradient gel. The pH discontinuity of the buffers is of no significance for the separation quality, and a "stacking-gel" with a different pH is not needed.

Ion Exchange Chromatography

Ion Exchange Chromatography relies on charge-charge interactions between the proteins in your sample and the charges immobilized on the resin of your choice. Ion exchange chromatography

can be subdivided into cation exchange chromatography, in which positively charged ions bind to a negatively charged resin; and anion exchange chromatography, in which the binding ions are negative, and the immobilized functional

group is positive. Once the solutes are bound, the column is washed to equilibrate it in your starting buffer, which should be of low ionic strength, then the bound molecules are eluted off using a gradient of a second buffer which steadily increases the ionic strength of the eluent solution. Alternatively, the pH of the eluent buffer can be modified as to give your protein or the matrix a charge at which they will not interact and your molecule of interest elutes from the resin. If you know the pH you want to run at and need to decide what type of ion exchange to use paste your protein sequence into the titration curve generator. If it is negatively charged at the pH you wish, use an anion exchanger; if it is positive, use a cation exchanger. Of course this means that your protein will be binding under the conditions you choose. In many cases it may be more advantageous to actually select conditions at which your protein will flow through while the contaminants will bind. This mode of binding is often referred to as "flow through mode". This is a particularly good mode to use in the case of anion exchange. Here one could use this type of mode to bind up endotoxins or other highly negatively charged substances well at the same time relatively simply flowing your protein through the matrix.

Charged groups:

The presence of charged groups is a fundamental property of an ion exchanger. The type of group determines the type and strength of the ion exchanger, their total number and availablity determines the capacity.There is a variety of groups which has been choosen for use in ion exchangers; some of these are:

Anion exchangers Functional group:-

Diethylaminoethyl (DEAE) $-O-CH_2-CH_2-N+H\ (CH_2CH_3)_2$

Quaternary aminoethyl (QAE) $-O-CH_2-CH_2-N+(C_2H_5)_2-CH_2-CHOH-CH_3$

Quaternary ammonium (Q) $-O-CH_2-CHOH-CH_2-O-CH_2-CHOH-CH_2-N+(CH_3)_3$

Cationexchangers Functional group:-

Carboxymethyl (CM) $-O-CH_2-COO$

Sulphopropyl (SP) $–O-CH_2-CHOH-CH_2-O-CH_2-CH_2-CH_2SO_3$

Methyl sulphonate (S) $-O-CH_2-CHOH-CH_2-O-CH_2-CHOH-CH_2SO_3$

Sulphonic and quaternary amono groups are used to form strong ion exchangers,the other groups form weak ion exchangers.The terms strong and weak refer to the extent of variation of ionization with pH not the strength of binding. Strong ion exchangers are completely ionized over a wide pH range, whereas with weak ion exchangers, the degree of dissociation and thus exchange capacity varies much more markedly with pH.

Some properties of strong ion exchangers are:

- Sample loading capacity does not decrease at high or low pH values due to loss of charge from the ion exchanger.
- A very sample mechanism of interaction exists between the ion exchanger and the solute.
- Ion exchange experiments are more controllable since the charge characteristics of the media do not change with changes in pH.
- **Anion Exchange Chromatography (AEC)**

- The surface charge of the solutes (proteins, nucleic acids, endotoxin) which bind will be net negative, thus to get binding of a specific protein one should be above the pI of that protein. Commonly used anion exchange resins are Q-resin, a Quaternary amine; and DEAE resin, DiEthylAminoEthane AEC is often used as a primary chromatography step due to its high capacity, (Matrices can bind from 10 to 100 mg of protein per ml) and ability to bind up and separate fragmented nucleic acids and lipopolysaccharides from the initial slurry.
-
- Typically, AEC is performed using buffers at pH's between 7 and 10 and running a gradient from a solution containing just this buffer to a solution containing this buffer with 1M NaCl. The salt in the solution compet **Akta Prime (Amersham Biosciences)** and releases the protein from its bound state at a given concentration. Proteins separate because the amount of salt needed to compete varies with the external charge of the protein. Uses of AEC include initial clean up of a crude slurry, separation of proteins

from each other, concentrating a protein, and the removal of negatively charged endotoxin from protein preparations.

- **Cation Exchange Chromatography (CEC)**

The surface charge of the solutes (proteins, nucleic acids, endotoxin) which bind will be net positive, thus to get binding of a specific protein one should be below the pI of that protein. Commonly used cation exchange resins are S-resin, sulfate derivatives; and CM resins, carboxylate derived ions CEC is less commonly used compared to AEC, largely due to the fact that often proteins do not stick to this resin at physiological pHs and one is reluctant to titrate a protein through its isoelectric point to get it to adhere to the resin. Nonetheless, it is as powerful as AEC for initial separations with equivalently high capacity. Typically, CEC is performed using buffers at pH's between 4 and 8 and running a gradient from a solution containing just this buffer to a solution containing this buffer with 1M NaCl. Uses of CEC include initial clean up of a crude slurry, separation of proteins from each other, concentrating a protein, and as a common first purification step for proteins expressed under acidic conditions such as in *P.pastoris.*

Gel filtration chromatography

Gel filtration chromatography is a separation based on size. It is also called molecular exclusion or gel permeation chromatography. In gel filtration chromatography, the stationary phase consists of porous beads with a well-defined range of pore sizes. The stationary phase for gel filtration is said to have a fractionation range, meaning that molecules within that molecular weight range can be separated.

Proteins that are small enough can fit inside all the pores in the beads and are said to be included. These small proteins have access to the mobile phase inside the beads as well as the mobile phase between beads and elute last in a gel filtration separation. Proteins that are too large to fit inside any of the pores are said to be excluded. They have access only to the mobile phase

between the beads and, therefore, elute first. Proteins of intermediate size are partially included - meaning they can fit inside some but not all of the pores in the beads. These proteins will then elute between the large ("excluded") and small ("totally included") proteins.

- Consider the separation of a mixture of glutamate dehydrogenase (molecular weight 290,000), lactate dehydrogenase (molecular weight 140,000), serum albumin (MW 67,000), ovalbumin (MW 43,000), and cytochrome c (MW 12,400) on a gel filtration column packed with Bio-Gel P-150 (fractionation range 15,000 - 150,000).

Gel Filtration Set up

When the protein mixture is applied to the column, glutamate dehydrogenase would elute first because it is above the upper fractionation limit. Therefore it is totally excluded from the inside of the porous stationary phase and would elute with the void volume (V_0). Cytochrome c is below the lower fractionation limit and would be completely included, eluting last. The other proteins would be partially included and elute in order of decreasing molecular weight.

These separations can be described by this equation

$$V_r = V_o + KV_i$$

where V_r is the retention volume of the protein, V_0 is the volume of mobile phase between the beads of the stationary phase inside the column (sometimes called the void volume), V_i is the volume of mobile phase inside the porous beads (also called the included volume) and K is the partition coefficient (the extent to which the protein can penetrate the pores in the stationary phase, with values ranging between 0 and 1). In the mixture of proteins listed above, the partition coefficient (K) for glutamate dehydrogenase would be 0 (totally excluded), K = 1 for cytochrome c (totally included) and K would be between 0 and 1 for the other proteins, which are within the fractionation range for the column.

In practice, gel filtration can be used to separate proteins by molecular weight at any point in a purification of a protein. It can also be used for buffer exchange - a protein dissolved in a sodium acetate buffer, pH 4.8, can be applied to a gel filtration column that has been equilibrated with tris buffer, pH 8.0. Using the tris buffer, pH 8.0, as the mobile phase, the protein moves into the tris mobile phase as it travels down the column, while

the much smaller sodium acetate buffer molecules are totally included in the porous beads and travels much more slowly than the protein.

Affinity chromatography

Affinity chromatography is a method of separating biochemical mixtures based on a highly specific biological interaction such as that between antigen and antibody, enzyme and substrate, or receptor and ligand. Affinity chromatography combines the size fractionation capability of gel permeation chromatography with the ability to design a chromatography that reversibly binds to a known subset of molecules.

Affinity chromatography can be used to:

- Purify and concentrate a substance from a mixture into a buffering solution
- Reduce the amount of a substance in a mixture
- Discern what biological compounds bind to a particular substance, such as drugs
- Purify and concentrate an enzyme solution.
- Affinity chromatography can be used in a number of applications, including nucleic acid purification, protein purification from cell free extracts, and purification from blood.

Principle

The immobile phase is typically a gel matrix, often of agarose; a linear sugar molecule derived from algae. Usually the starting point is an undefined heterogeneous group of molecules in solution, such as a cell lysate, growth medium or blood serum. The molecule of interest will have a well known and defined property which can be exploited during the affinity purification process. The process itself can be thought of as an entrapment, with the target molecule becoming trapped on a solid or stationary phase or medium. The other molecules in solution will not become trapped as they do not possess this property. The solid medium can then be removed from the mixture, washed and the target molecule released from the entrapment in a process known as elution. Possibly the most common use of affinity chromotography is for the purification of recombinant proteins.

Ultrafiltration:

Ultrafiltration or UF is a pressure driven membrane separation process that separates particulate matter from soluble components in the carrier fluid (such as water).UF membranes typically have pore sizes in the range of 0.01 - 0.10 µm and have a high removal cabability for bacteria and most viruses, colloids and silt (SDI). The smaller the nominal pore size, the higher the removal cabability. Most materials that are used in UF are polymeric and are naturally hydrophobic. Common polymeric materials used in UF include: Polysulfone (PS), Polyethersulfone (PES), Polypropylene (PP), or Polyvinylidenefluoride (PVDF). Although these materials can be blended with hydrophilic agents, they can reduce the membranes ability to be cleaned with high strength disinfectants such as hypochlorite that impacts removal of bacterial growth.

The OMEXELL™ UF module utilizes a double-walled hollow fiber (capillary) PVDF membrane which has a very small nominal pore diameter for PVDF material that allows for the removal of all particulate matter, bacteria and most viruses and colloids. Despite the small pore diameter, the membrane has a very high porosity resulting in a flux similar to that of micro-filtration (MF) and can effectively replace MF in most cases.Systems designed with OMEXELL™ UF use an outside-in flow configuration which allows for less plugging, higher solids loading, higher flow area and easy cleaning. The primary flow design is dead-end filtration but the module can be operated using a concentrate bleed. Dead-end filtration uses less energy and has a lower operating pressure than the concentrate bleed, therefore reducing operating costs.

Typically, OMEXELL™ UF is operated at a constant permeate flow. The transmembrane pressure (TMP) will naturally increase over time and the module can be cleaned periodically by back flushing and air scouring to remove the fouling layer. Disinfectants
and other cleaning agents can be used to fully remove and prevent performance loss due to biological growth as well as other foulants.

Centrifugation

Centrifugation is a process that involves the use of the centripetal force for the separation of mixtures, used in industry and in laboratory settings. In chemistry and biology, increasing the

effective gravitational force on a test tube so as to more rapidly and completely cause the precipitate (“pellet”) to gather on the bottom of the tube. The remaining solution is called the “supernate” or “supernatant liquid”. The supernatant liquid is then either quickly decanted from the tube without disturbing the precipitate, or withdrawn with a Pasteur pipette. The rate of centrifugation is specified by the acceleration applied to the sample , typically measured in revolutions per minute (RPM) or g. The particles settling velocity in centrifugation is a function of their size and shape, centrifugal acceleration, the volume fraction of solids present, the density difference between the particle and the liquid, and the viscosity.

Lyophilization:

Lyophilization is a process more commonly known as freeze-drying. The word is derived from Greek, and means "made solvent-loving". Lyophilization is a way of drying something that minimizes damage to its internal structure. Because lyophilization is a relatively complex and expensive form of drying, it is limited to those materials which are sensitive to heat and have delicate structures and substantial value. One of the only substances which cannot be preserved effectively by freeze-drying is mammalian cells, which are too fragile.The preferred method of preservation in the biotechnologyindustry,lyophilization is regularly used to preserve vaccines, pharmaceuticals, and other proteins.

Freeze-drying is also used to preserve special food products, eliminating the need for refrigeration. Freeze-dried food is eaten by mountain climbers and astronauts. Lyophilization is used by botanists to preserve flower samples indefinitely. Because the process of freeze-drying removes most of the water from the sample, freeze-dried materials become highly absorbent, and merely adding water can restore the sample to something close to its original state. The energy and equipment costs of lyophilization are around 2-3 times higher than those of other drying methods. The drying cycle is also longer, about 24 hours. First, the temperature of the sample is lowered to near freezing point. Then, the sample is inserted into a vacuum chamber. The more energetic molecules escape, lowering the temperature further, while the extremely low pressure causes water molecules to be drawn out of the sample. Attached to the vacuum chamber is a condenser, which converts the airborne moisture into liquid and siphons it away.

Great care is taken throughout the process to ensure that the structure of the sample remains constant. For instance, the sample could merely be frozen by the vacuum rather than being

frozen under atmospheric pressures, but that would cause shrinkage in the sample, damaging its structure irreversibly.

The primary mechanism that allows for freeze-drying is sublimation, whereby ice is directly converted to water vapor, without passing through the intermediary stage of a liquid. Rather than through heating, this is done by removal of pressure so that the ice boils without melting. The result is a sample whose structure is largely preserved, which can be stored at room temperatures and pressures.

Spectrophotometry:

Spectrophotometry is the quantitative study of electromagnetic spectra. It is more specific than the general term electromagnetic spectroscopy in that spectrophotometry deals with visible light, near-ultraviolet, and near-infrared. Also, the term does not cover time-resolved spectroscopictechniques.The most common spectrophotometers are used in the UV and visible regions of the spectrum, and some of these instruments also operate into the near-infrared region as well.Visible region 400-700nm spectrophotometry is used extensively in colorimetry science. Ink manufacturers, printing companies, textiles vendors, and many more, need the data provided through colorimetry. They usually take readings every 20 nanometers along the visible region, and produce a spectral reflectance curve. These curves can be used to test a new batch of colorant to check if it makes a match to specifications. Traditional visual region spectrophotometers cannot detect if a colorant has fluorescence. This can make it impossible to manage color issues if one or more of the printing inks is fluorescent. Where a colorant contains fluorescence, a bi-spectral fluorescent spectrophotometer is used. There are two major setups for visual spectrum spectrophotometers, d/8 (spherical) and 0/45. The names are due to the geometry of the light source, observer and interior of the measurement chamber. Scientists use this machine to measure the amount of compounds in a sample. If the compound is more concentrated more light will be absorbed by the sample; within small ranges, the Beer-Lambert law holds and the absorbance between samples vary with concentration linearly. Samples are usually prepared in cuvettes; depending on the region of interest, they may be constructed of glass, plastic, or quartz.

Protein sequencing:

Proteins are found in every cell and are essential to every biological process, protein structure is very complex: determining a protein's structure involves first protein sequencing - determining the amino acid sequences of its constituent peptides; and also determining what conformation it adopts and whether it is complexed with any non-peptide molecules. Discovering the structures and

functions of proteins in living organisms is an important tool for understanding cellular processes, and allows drugs that target specific metabolic pathways to be invented more easily.The two major direct methods of protein sequencing are mass spectrometry and the Edman degradation reaction. It is also possible to generate an amino acid sequence from the DNA or mRNA sequence encoding the protein, if this is known. However, there are a number of other reactions which can be used to gain more limited information about protein sequences and can be used as preliminaries to the aforementioned methods of sequencing or to overcome specific inadequacies within them.

Sample preparation:

Samples for sequencing should be free of primary amines (e.g., Tris or glycine) and should contain less than 2 µmol equivalent of non-volatile salt (e.g., 100 µl of 20 mMNaCl).Ethanol-precipitation, trichloroacetic acid precipitation, reverse phase chromatography, gel filtration, ion-exchange separation, Centri-Plus concentrator or dialysis can remove these materials. Chromatography and SDS-PAGE are two commonly used approaches for sample preparation. SDS-PAGE followed by electrobloting of protein samples onto PVDF membranes is the preferred method for N-terminal sequence analysis. PVDF membranes, pore size 0.2 micron, are known to have a superior protein binding capacity. To avoid high glycine background during Edman degradation, electroblotting must be performed in 10 mM CAPS (3-[cyclohexylamino-1-propanesulfonic acid]) buffer with 10% methanol, pH 11.0 for atleast 15- 30 minutes for 30 KDa size protein. Following the transfer, rinse the membrane 4-5 times with Milli-Q water to reduce the left over Tris and glycine contaminats from electrophoresis.

Staining

The proteins can be visualized on the blot by staining with Coomassie blue (originally developed as acid wool dyes in Kumasi, Ghana in the late 1800's), Ponceau S, or Amido Black (the latter two are less sensitive stains). Coomassie blue is sensitive enough to detect 50-200 ng of protein on PVDF. PVDF-based membranes have higher protein-binding capacity and result in better average initial and repetitive yields (please note that

nitrocellulose membranes are not compatible with the reagents and organic solvents used in automated protein sequencing). The size of an excised piece of PVDF membrane containing the protein band should be smaller than 40 mm2. As a rule of thumb, if the Coomassie blue stained protein band on the blot is visible on a photocopy, then there is enough material for sequencing. If necessary, a standard protein can be blotted on the PVDF to verify that the protein was electroblotted efficiently. Contaminating salts can be removed by washing the membrane with MilliQ water, prior to sequencing the sample on

PVDF membrane.ProSorb cartridge, a device specifically designed for samples in solution containing salts, will be provided on request.

Turn-Around Time

For a single sequence the turn-around time is 2-3 working days. Additional sequences will take 1-2 additional days per sample.

Purity

Liquid samples should contain only one protein or peptide. Buffers, SDS, salts, amino acids, primary amines, and other contaminants must be removed from your sample. These contaminants affect Edman degradation reaction on the instrument, contaminate the instrument or affect PTH amino acid detection. Samples submitted on PVDF (blotted from SDS-PAGE gels) should have well separated bands to minimize contamination.

Crystallization:

A protein will stay in solution only up to a certain concentration. Once this limiting concentration is reached, the solution will no longer remain homogeneous, but a new state or phase will appear. This phenomenon forms the basis of all protein crystallization experiments. By changing the solution conditions, the crystallographer tries to exceed the solubility limit of the protein so as to produce crystals.

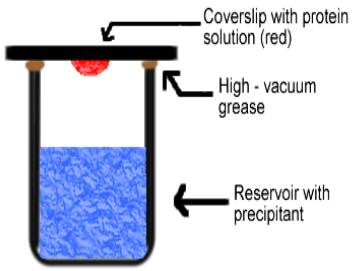
This plan rarely runs smoothly. After changing the solution conditions, one of several difficulties is usually encountered: (i) nothing happens, i.e., the protein solution remains homogeneous; (ii) a new phase appears, but it is not a crystal. Instead, it is an aggregate or a liquid; or (iii) crystals do form, but they are unsuitable for structure determination because they give a poor X-ray diffraction pattern.It is often possible to overcome these difficulties by trial and error—repeated crystallization attempts with many different conditions—but this strategy does not always work. Even when it is successful, the lessons learned cannot be easily generalized; the conditions which work with one protein are not necessarily optimal for a different protein. The goal of crystallization is usually to produce a well-ordered crystal that is lacking in contaminants and large enough to provide a diffraction pattern when hit with x-ray. This diffraction pattern can then be analyzed to discern the protein's three-dimensional structure. Protein crystallization is inherently difficult because of the fragile nature of protein crystals. Proteins have irregularly shaped surfaces, which results in the formation of large channels within any protein crystal. Therefore, the noncovalent bonds that hold together the lattice must often be formed through several layers of solvent molecules (Rhodes, 1993). In addition to overcoming the inherent fragility of protein crystals, the successful production of x-ray worthy crystals is dependent upon a number of environmental factors because so much variation exists among proteins, with each individual requiring unique condition for successful crystallization. Therefore, attempting to crystallize a protein without a proven protocol can be very tedious. Some factors that require consideration are protein purity, pH, concentration of protein, temperature, and precipitants. In order for sufficient homogeneity, the protein should usually be at least 97% pure. pH conditions are also very important, as different pH's can result in different packing orientations. Buffers, such as Tris-HCl, are often necessary for the maintenance of a particular pH (Branden and

Tooze, 1999). Precipitants, such as ammonium sulfate or polyethylene glycol, are compounds that cause the protein to precipitate out of solution (Rhodes, 1993).

Vapor Diffusion:

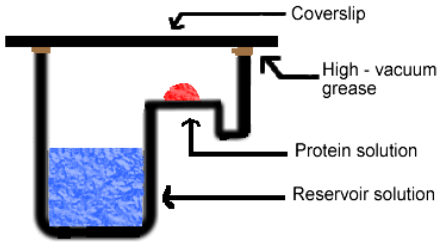
Two of the most commonly used methods for protein crystallization fall under the category of vapor diffusion. These are known as the **hanging drop** and **sitting drop** methods. Both entail a droplet containing purified protein, buffer, and precipitant being allowed to equilibrate with a larger reservoir containing similar buffers and precipitants in higher concentrations. Initially, the droplet of protein solution contains an insufficient concentration of precipitant for crystallization, but as water vaporizes from the drop and transfers to the reservoir, the precipitant concentration increases to a level optimal for crystallization. Since the system is in equilibrium, these optimum conditions are maintained until the crystallization is complete (Rhodes, 1993; McRee, 1993). Simply put, the hanging drop method differs from the sitting drop method in the vertical orientation of the protein solution drop within the system. It is important to mention that both methods require a closed system, that is, the system must be sealed off from the outside using an airtight container or high-vacuum grease between glass surfaces. Figures 3 and 4 depict the hanging drop and sitting drop systems, respectively (Rhodes, 1993; McRee, 1993).

Diagram of hanging drop method:



Reservoir solution (blue) usually contains buffer and precipitant. Protein solution (red) contains the same compounds, but in lower concentrations. The protein solution may also contain trace metals or ions necessary for precipitation of particular proteins. For instance, insulin is known to require trace amounts of zinc for crystallization (McRee, 1993).

Diagram of sitting drop method:



In this method, the protein drop sits on a pedestal above the reservoir solution, as opposed to hanging.

Why Crystallize Proteins?

The crystallization of proteins currently has three major applications: (1) structural biology and drug design, (2) bioseparations, and (3) controlled drug delivery. In the first application, the protein crystals are used with the techniques of protein crystallography to ascertain the three-dimensional structure of the molecule. This structure is indispensable for correctly determining the often complex biological functions of these macromolecules. The design of drugs is related to this, and involves designing a molecule that can exactly fit into a binding site of a macromolecule and block its function in the disease pathway. Producing better quality crystals will result in more accurate 3D protein structures, which in turn means its biological function can be known more precisely, also resulting in improved drug design.

SWISS PDB Viewer

Swiss PDB viewer is an application which is used to analyse several proteins at the same time. It is a user friendly interface. In Swiss PDB viewer, proteins are superimposed on each other to deduce structural alignments and to compare their active sites. Swiss PDB also helps to find any other relevant parts present in proteins. Mutation in Amino acid, angles, H-bonds, and distances between atoms can be easily found out because of intuitive graphic and menu interface.

Threading of protein primary sequence onto 3D template can also be easily done and an immediate feedback of threaded protein will be accepted by the reference structure. When submitting a request in order to build missing loops it also refines side chain packing.

Swiss-PDB Viewer can additionally scan electron density maps, and provides various tools to build into the density. In addition, various modelling tools square measure integrated and residues are mutated.

Why Crystallize Proteins?

The crystallization of proteins currently has three major applications: (1) structural biology and drug design, (2) bioseparations, and (3) controlled drug delivery. In the first application, the protein crystals are used with the techniques of protein crystallography to ascertain the three-dimensional structure of the molecule. This structure is indispensable for correctly determining the often complex biological functions of these macromolecules. The design of drugs is related to this, and involves designing a molecule that can exactly fit into a binding site of a macromolecule and block its function in the disease pathway. Producing better quality crystals will result in more accurate 3D protein structures, which in turn means its biological function can be known more precisely, also resulting in improved drug design.

SWISS PDB Viewer

Swiss PDB viewer is an application which is used to analyze several proteins at the same time. It is a user friendly interface. In Swiss PDB viewer, proteins are superimposed on each other to deduce structural alignments and to compare their active sites. Swiss PDB also helps to find any other relevant parts present in proteins. Mutation in Amino acid, angles, H-bonds, and distances between atoms can be easily found out because of intuitive graphic and menu interface.

Threading of protein primary sequence onto 3D template can also be easily done and an immediate feedback of threaded protein will be accepted by the reference structure. When submitting a request in order to build missing loops it also refines side chain packing.

Swiss-PDB Viewer can additionally scan electron density maps, and provides various tools to build into the density. In addition, various modeling tools square measure integrated and residues are mutated.