Introduction:

Gregor Johann Mendel’s observations on plant hybridization in the later half of nineteenth century formed the basis for the modern study of genetics. By the beginning of the twentieth century, the pathways for the breakdown of certain amino acids has been worked out and it was in the year 1990 that the concept of an ‘inborn error of metabolism’ was proposed by Garrod who perceived a relationship between genes and enzymes to explain the rare condition, alkaptonuria. Till date, hundreds of genetic diseases caused by the deficiency of a single enzyme have been described (Scriver et al, 2001). Some of these enzymes are actively engaged in catabolism and are present in the cellular structure known as the Lysosome.

Lysosome

The lysosome is just one component of a series of seemingly unconnected intracellular organelles, collectively known as the endosomal-lysosomal system. (De Duve & Wattiaux 1966). The endosome and lysosome form a chain responsible for the catabolism of naturally endogenous and exogenous macromolecules and the subsequent recycling of their constituent monomeric components.

The lysosome is characterized by the presence of a membrane, a low internal pH and vesicles containing many hydrolytic enzymes. (Vellodi 2005). Inactive digestive enzymes are formed in the rough ER and transported to the Golgi complex through vesicles, where they are activated. Activated digestive enzymes are surrounded by membranes that bud off from the Golgi complex, forming lysosomes. Lysosomes can fuse with food containing structure, to digest their contents (Figure 1).
Figure 1: The structure and function of lysosomes.

The membrane contains transport system that carries particles between lumen and cytosol and an electrogenic proton pump called the vacuolar proton pump, or V-type H+ -ATPase to acidify the internal lysosomal environment (Arai et al 1993). Substrate breakdown occurs within the lysosome. It is also capable of secreting its contents after fusion with the plasma membrane (Luzio et al 2000). Another important role for lysosomes involves elevation in intracellular Ca ++ which triggers fusion of lysosomes with the plasma membrane facilitating plasma membrane resealing. That the calcium regulated exocytosis of lysosomes is very important for membrane repair of the cell (Reddy et al 2001).
**Lysosomal Storage Disorders (LSDs):**

LSDs comprise a group of at least 50 distinct genetic diseases (Meikle et al. 2004), each one resulting from deficiency of a particular lysosomal protein/activity or, in few cases, from non-lysosomal activities that are involved in lysosomal biogenesis or protein maturation. The number of recognized LSDs is increasing as newer disorders are characterized biochemically and genetically.

Over the past decade, a deficiency of cathepsin K has been described, which results in an LSD called pycnodysostosis (Gelb et al. 1996), and several of the genes and proteins involved in the neuronal ceroid lipofuscinoses (Batten disease family) have been characterized continually adding more to the spectrum of Lysosomal storage disorders.

Infantile neuronal ceroid lipofuscinosis, also known as Santavuori disease, has been shown to result from a deficiency of palmitoyl protein thioesterase (Vesa et al. 1995), and classic late-infantile neuronal ceroid lipofuscinoses (Jansky–Bielschowsky disease) has been shown to result from a deficiency of tripeptidyl peptidase I (Sleat et al. 1997), both of which are lysosomal enzymes. At least eight genes are thought to be involved in this group of LSDs, but only six have been identified to date (Ezaki et al. 2004). In addition, the protein deficiency leading to Danon disease was recently identified as the lysosome-associated membrane protein LAMP-2 (Nishino et al. 2000). Undoubtedly, many more proteins and genes are involved in LSDs but are yet to be characterized.
Most LSDs are inherited in an autosomal recessive manner, with the exception of Hunter syndrome or mucopolysaccharidosis type II (MPS II), which shows X-linked recessive inheritance, Danon disease, which is X-linked dominant, and Fabry disease, which, with a high proportion of affected females, should not be described as X-linked recessive.

Lysosomal storage disorders (LSDs) are all characterized by the intra-lysosomal accumulation of unmetabolized substrates resulting in the formation of large intracellular vacuoles (Figure 2), which is the primary mechanism of the disease.

![Figure 2: Enlarged vacuoles in a lysosomal storage disease. The electron microscopy picture of a leukocyte cell from a patient affected with alpha-mannosidosis (Source: Dr. Dag Malm).](image)

However, the extensive range of disease symptoms indicates that many secondary biochemical and cellular pathways must also be activated. So, the primary accumulating metabolite affects a secondary biochemical or cellular pathway, which then subsequently causes tissue pathology, altered gene expression and the activation of tertiary biochemical pathways. (Figure 3)
Thus understandably, any macromolecule which accumulates giving an idea of the disease will qualify to be called a primary biomarker and any other marker which is secondarily activated or suppressed or signifies that alteration would best be named a secondary biomarker.

**Biochemical tools for the diagnosis of LSD’s: Advantages & Disadvantages**

Suspicion of LSD is based on clinical findings but it is not a confirmed diagnosis until the suspected enzyme deficiency is established. Lysosomal enzymes are present in all cells except mature RBC’s. The gold standard for diagnosis of LSD’s is the assay of the deficient enzyme in peripheral blood leucocytes, plasma and fibroblasts.
Enzyme assays performed on a combination of leucocytes and plasma predominantly include enzymes involved in digestion of glucose phenotypes and oligosaccharides i.e. Gaucher’s Disease; Niemann Pick disease types A, B and Mucopolysaccharidosis amongst others.

Measurement of enzyme activity in leucocytes and plasma enables the diagnosis of most LSD’s yet, a proportion may not be detected by this method i.e. in Sialodosis and Pompe disease, where the distinction between normal and affected could be very narrow. When leucocytes assays are not reliable, the enzyme analysis involves assaying hair roots that develop form progenitor cells which may be more representative and reliable.

Enzyme assay, the primary tool for diagnosis, and in the absence of any other modality is also utilized for diagnosis of carriers and for supporting prenatal diagnosis. This has a lot of disadvantages. There is a large overlap in enzyme levels among carrier and normal individuals; hence it is very difficult to detect carriers by enzyme assay. Mutation analysis forms the mainstay of diagnostic confirmation and is also a more reliable tool for prenatal diagnosis.

The other disadvantages of enzyme assays are its standardization and are fraught with possibility of ‘pseudo deficiency’. For example, in Gaucher’s disease the Glucocerebrosidase level of an individual is not always predictive of the phenotypic expression of Gaucher’s disease. Even among monozygotic twins with similarly abnormal glucocerebrosidase levels, a remarkable variability in disease manifestation can occur. Hence search for other diagnostic tools which are informative and conclusive are required. With more disorders becoming amenable to enzyme replacement therapy, estimation of enzyme levels for monitoring these patients is evidently futile.
The Concept of a biomarker

The identification of factors secreted by gaucher cells were called ‘Biomarkers of fundamental interest, because they may be of both diagnostic and prognostic significance. Quantitative biomarkers which correlate with clinical manifestation of disease that affect the quality of life, rest of complications or survival (surrogate biomarkers) may play a vital role in diagnosis. Surrogate biomarkers may thus have special significance in monitoring, treatment for these rare diseases.

Most mutations in ‘classic’ lysosomal storage disorders (LSDs) result in the delivery of a defective enzyme that has a reduced catalytic activity in lysosomes (label 1) (Figure 4).

Figure 4: The biochemical and cellular basis of lysosomal storage disorders
In some cases, another protein that is required for optimal hydrolase activity is defective or absent (label 2). LSD can be caused by the defective transport of a lysosomal hydrolase out of the endoplasmic reticulum (ER) due to a mutation that causes misfolding (label 3). Alternatively, LSD can be caused by the defective transport of a lysosomal hydrolase out of the ER because a multi-enzyme complex that is required for transport cannot be formed (label 4). In the Golgi, defective glycosylation could result in an enzyme with reduced catalytic activity (label 5). Alternatively, defective glycosylation in the Golgi could produce an enzyme that cannot reach lysosomes because it cannot bind to mannose-6-phosphate receptors (due to defective glycosylation with mannose-6-phosphate; label 6).

Defects in other transport steps from the Golgi could also lead to LSD (label 7). Several LSDs are caused by defects in integral lysosomal membrane proteins. These include defects in transporters (label 8), or in proteins that are involved in other vital regulatory events of lysosomal function (label 9).

The lysosomal hydrolases are shown in various shades of blue, and a relevant LSD example is shown for each defect when one is known. (source: nature Review, molecular cell biology volume 5, July 2004).
The utility of a biomarker in the setting of poor genotype-phenotype correlation

Within the lysosomal storage disorders, there is a general relationship between the inherited mutation and the disease manifestation, that is most severe phenotype occurs when no functional enzyme is produced, while mild expression or phenotype occurs with mutation that do not completely abolish enzyme function.

However, many unexplained discrepancies remain and mutation alone seems unable to completely predict the phenotypic expression. Phenotypic variation of a simple Mendelian condition (such as Gaucher disease) within kindreds emphasizes the influence of modulating factors. These factors may include specific modifier genes, the genetic background in general and the influence of the environment. Thus this limitation could be offset if a biomarker could reliably predict the clinical phenotype.

Concept of an ideal biomarker in reference to Lysosomal Storage Disease:

A biomarker is an indicator of the presence or extent of a biological process that is directly linked to the clinical manifestation and outcome of a particular disease. A biomarker should be readily quantifiable in body fluid, tissues or whole organs to provide indirect but ongoing and specific determination of disease severity. An ideal biomarker should have some specific characteristics.
Characteristics of the ideal biomarker:

- Easy to quantify in accessible material
- Abundance not subject to wide variation in the general population
- Reflects total burden of disease at all sites
- Expression specifically increased in the relevant disease
- Unaffected by unrelated condition and co morbid factors
- Expression increased in the established disease without overlap between untreated patients and healthy subjects
- Abundance varies rapidly in response to treatment in a manner closely correlated with established clinico-pathological parameters of disease
- Measurement is reliable, quick, reproducible and cheap

Categories of biomarker:

Biomarkers that are involved in the study of LSDs fall broadly into two categories. The first category is that of a direct biomarker which affects the macromolecular substrates of the enzyme system that is defective and thus reflect directly the pathological storage material. Examples include plasma, tissue and urinary globotriaosylceramide (Gb₃) in Fabry disease and glycosaminoglycans (GAG’s like dermatan sulfate and heparan sulfate) in mucopolysaccharidoses types I, II and III.
The second category is indirect biomarker (like Chitotriosidase and CCL18/PARC) which includes those molecules or changes in cell, tissue or organ function whose accumulation in biological samples is increased as consequence of ill understood pathological effects but may provides greater information about local burden of storage and disruptions of functions and that reflect the primary effects of lysosomal pathology in each disorder.

**Chemokines:**

Chemokines constitute a family of chemotactic cytokines that act through seven-transmembrane domain G protein-coupled receptors on their target cells (Cyster JG et al.1999, Zlotnik A et al. 2001, Yoshie O et al 2001). According to the organization of their NH$_2$-terminal Cys residues, chemokines are structurally divided into the CC, CXC, CX$_3$C, and C chemokines. Chemokines are key players in directing the migration and the activation of leukocytes throughout the body, under physiological and immunopathological conditions.

Furthermore, they are involved in various other processes, including angiogenesis, hematopoiesis, tumor growth, and metastasis. The first wave of chemokines was mainly identified on the basis of their chemotactic properties (Mantovani A et al 1999). These chemokines are in general highly inducible in multiple cell types and are responsible for the leukocyte infiltration at sites of inflammation.
The second generation of chemokines has been discovered since 1996 through bioinformatics, these chemo attractants are mostly implicated in the homeostatic trafficking of lymphocytes and dendritic cells (DC) to those specific tissues where they are constitutively expressed. However, some of these second-generation chemokines, such as the CC chemokine ligand 18 (CCL18), apparently belong to the inflammatory/inducible as well as the constitutive/homeostatic chemokines, depending on the circumstances (Mantovani A et al 1999).

**Discovery of CCL18 as biomarker:**

By searching the public GenBank expressed sequence tag (EST) database with the cDNA sequence of the human CC chemokine CCL3, (Hieshima et al. 1997) identified a series of partial cDNA sequences, encoding a polypeptide with significant (64%) sequence identity to CCL3. The full-length cDNA of this new CC chemokine was obtained from human fetal lung. The finding that this chemokine was constitutively expressed at high levels in human lung together with its inducible expression in some human cell lines led to the designation pulmonary and activation-regulated chemokine (PARC). Analogously, (Wells et al. 1997) discovered, by means of computer-assisted analysis of the EST database, a sequence coding for a new CCL3-like polypeptide, which they named macrophage inflammatory protein-4 (MIP-4).
Independently, PARC was also cloned from a cDNA library of monocyte-derived DC and from a cDNA library of macrophages alternatively activated by interleukin (IL)-4 and glucocorticoids (GC) and nominated DC-chemokine 1 (DC-CK1) and alternative macrophage activation-associated CC chemokine-1 (AMAC-1), respectively (Adena et al. 1997, Kodelja et al. 1998). Furthermore, (Guan et al. 1999) isolated, through exon trapping of genomic fragments close to the CCL3 gene, the gene for a novel CCL3-like chemokine, which corresponded to MIP-4. Following the new systematical chemokine nomenclature, PARC/MIP-4/DC-CK1/AMAC-1 has been renamed CCL18 (Zlotnik A et al. 2000).

It is likely that a twofold duplication of the CCL3 gene led to the generation of two CCL3-like genes. The fusion of the latter in combination with the inactivation of some exons (possibly by base changes around former exon intron borders) and the deletion of other sequences (marked by a cross) might have created a new transcription unit, i.e., the CCL18 gene (Figure 5) (Tasaki et al. 1999, Politz et al. 2000). The consensus genomic organization of CCL18 depicted here is composed on the basis of different reports (Hieshima et al. 1997, Adema et al. 1997, Kodelja et al. 1998, Guan et al. 1999, Tasaki et al. 1999, Politz et al. 2000).
Figure 5: Putative generation mechanism and structure of the human CCL18 gene.

Dotted boxes indicate the pseudoexons, whereas the open and shaded boxes represent the untranslated and translated parts from the actively used exons, respectively. The intron sequences are indicated as horizontal lines between the used exons. Lengths of sequences are in base pairs (bp). The length of the 3’- untranslated region of the CCL18 gene (marked by “?”) varies between 430 bp and 442 bp according to the clones isolated (Hieshima et al 1997, Adema et al. 1997, Kodelja et al.1998, Guan et al. 1999, Tasaki et al. 1999, Politz, et al. 2000). (Tasaki Y et al. 1999) demonstrated the transcription initiation site (arrow) presented here.
Several human diseases have been reported to be accompanied with elevated levels of CCL18 like for example Gaucher disease, atherosclerosis, sarcoidosis, active hepatitis C infection, hypersensitive pneumonitis, allergic contact hypersensitivity, septic as well as rheumatoid arthritis, ovarian carcinoma, gastric carcinoma and recently Whipple disease, Niemann-Pick type B disease and beta-thalassemia (Boot et al 2004, Reape et al., 1999; Kusano et al., 2000; Schutyser et al., 2001; Pardo et al., 2001; Schutyser et al., 2002; Mrazek et al., 2002; Goebeler et al., 2001; Leung et al., 2004; Desnues et al., 2005; Dimitriou et al., 2005, Brinkman et al., 2005).

**Heparin Cofactor II – Thrombin Complex**

Heparin cofactor II (HCFII) is a serine protease inhibitor (serpin) that inactivates thrombin rapidly in the presence of certain glycosaminoglycans (Bauer et al 1993). Aihara et al. (2004) measured plasma HCF II activity, HDL cholesterol level, and carotid artery plaque thickness in Japanese individuals and observed that HCF II activity decreased with age. Multiple regression analysis revealed that plasma HCF II activity and HDL cholesterol level were independently associated with decreased plaque thickness and that the antiatherogenic contribution of HCF II activity was stronger than that of HDL cholesterol.
Using crossed immunolectrophoresis; Andersson et al. (1987) were the first to demonstrate molecular heterogeneity of the HCF II molecule, the so-called 'Oslo variant,' in affected members of two Norwegian families with HCF II deficiency (612356). Their findings were consistent with an autosomal dominant pattern of inheritance; affected individuals had half the normal amount of normal HCF II and were presumed heterozygotes.

Using PCR, Blinder et al. (1989) amplified DNA fragments encoding the N-terminal 220 amino acids of HCF II from a patient with the Oslo variant. They identified a point mutation resulting in an arg189-to-his (R189H; 142360.0001) substitution in 1 allele. Blinder et al. (1989) created the same mutation in the cDNA of native HCF II by oligonucleotide-directed mutagenesis and expressed it in E. coli. The recombinant cofactor reacted with thrombin in the presence of heparin, but not dermatan sulfate, confirming that the R189H mutation is responsible for the functional abnormality in HCF II Oslo.

HCFII does not inhibit other proteases involved in coagulation or fibrinolysis. Thrombin stimulates platelet aggregation, promotes coagulation by cleavage of fibrinogen and activation of factors V, VIII, XI, and XIII, and inhibits fibrinolysis by activation of a plasma carboxyptidase (Prat et al.1989). Conversely, when thrombin binds to thrombomodulin on the surface of endothelial cells, it activates protein C, which inhibits further thrombin generation.
Thrombin also engages in a variety of activities unrelated to hemostasis (Maekawa et al 1996). For example, it causes proliferation of fibroblasts and other cells, induces monocyte chemotaxis, promotes adhesion of neutrophils to endothelial cells, and inhibits neurite outgrowth. HCFII could potentially regulate the activity of thrombin in one or more of these diverse biological processes.

Thrombin (IIa), heparin (HEP) and heparin cofactor II (HCFII) interact to form a tertiary complex via a random order bireactant mechanism. Heparin subsequently dissociates from the complex as thrombin forms a covalent complex with the heparin cofactor II. (Figure 6)

![Diagram](image)

**Figure 6: Inhibition of Thrombin by Heparin Cofactor II**

The activation of coagulation ultimately leads to the activation of prothrombin to the enzyme thrombin. Unless regulated, thrombin will act on its natural substrates that include fibrinogen, factor V, factor VIII, factor XIII, Protein C, TAFI as well as specific receptors on platelets and endothelial cells. The activity of thrombin in plasma is regulated in part through interaction with protease inhibitors.
Based on kinetic rates and physiological concentrations, the primary inhibitor of thrombin in plasma is antithrombin (ATIII), followed by heparin cofactor II (HCII) and α2macroglobulin. The thrombin-heparin cofactor II complex (T-HCII) results when thrombin cleaves a scissile bond near the C-terminus of HCII, forming a covalent, 1:1 acyl enzyme intermediate with HCII with an apparent mass of 102 kDa. Calcium is not required for this interaction, but the rate of thrombin inhibition by HCII can be accelerated 1000-fold by optimal concentrations of heparin. Unlike Antithrombin, thrombin inhibition by HCII is also enhanced by dermatan sulphate. HCII-T complexes are cleared from circulation by serpin-enzyme complex receptors on the surface of hepatocytes, with a half-life of 10 minutes (Bauer et al. 1993, Prat et al. 1989, Maekawa et al. 1996).

The presence of thrombin-HCII complexes in human plasma indicates that HC II inhibits thrombin in vivo (Liu et al., 1995, Andersson et al. 1996). HCII is synthesized by the liver and circulates in human plasma at a concentration of about 1 μM (Tollefsen et al. 1986). Turnover studies of labeled HCII in humans suggest that about 40% of the protein equilibrates with an extra vascular compartment (Sie et al. 1985), but the tissue distribution of HCII has not been thoroughly investigated. HCII has been detected in the intima of normal human arteries, and the ability of dermatan sulfate in the arterial wall to stimulate HCII is decreased in atherosclerotic lesions (Cooper et al. 1996, Shirk 2000).
During pregnancy, both the maternal and the fetal blood contain trace amounts of a dermatan sulfate proteoglycan that stimulates thrombin inhibition by HCFII (Andrew et al 1992). The placenta is rich in dermatan sulfate and may be the source of the circulating proteoglycan, suggesting that HCFII could be activated locally to inhibit coagulation in the placenta (Brennan et al 1984).

Chemotactic peptides are released when HCFII is partially degraded by neutrophil proteases, suggesting a possible direct role for HCFII in inflammation (Church et al 1991). HCFII could also participate in wound healing by regulating the mitogenic or chemotactic activities of thrombin.

Elevated levels of Heparin Co factor II thrombin reported with some other human disease like mucopolysaccharidosis indicate that HCFII-T is an excellent biomarker for MPS and represents a novel finding that may implicate GAG modulated serpins and their proteases in the pathophysiology of MPS diseases. Importantly, the elevation of HCFII-T complex appears to be correlated to disease severity and is responsive to treatment (Derrick et al. 2006, Derrick et al. 2008, Kia et al 2010).

The quest for an ideal biomarker in lysosomal storage disorders is still underway, and several newly discovered, but also old and overlooked markers might prove their relevance. Therefore, the objectives of the study were to establish the potential role of biomarker CC chemokine ligand 18 and heparin co factor II thrombin complex for the diagnosis and their use to monitor the disease progression of Gaucher disease, Niemann Pick disease and Mucopolysaccharidoses.