CHAPTER 2

LITERATURE REVIEW

2.1. The biochemistry of trimethylaminuria

2.1.1. Introduction

Trimethylaminuria is a disorder that has been recorded in the literature for over a thousand years (Mitchell, 1996). It has only been recorded as a syndrome since 1970 and it is also called the fish odour syndrome or stale fish syndrome (Spellacy *et al.*, 1979; Ziegler, 1990; Anon 3, 1998; Christensen, 1999). There is no conventional cure known at present. This disorder does not seem to pose any immediate health risk because very little is known on the effects it may have on major metabolic pathways and drug metabolism. Its diagnostic characteristics include a bad smell that varies from stale fish-like to garbage-like odour.

Trimethylaminuria is known to occur in similar proportions in males and females (Mitchell, 1996). Trimethylaminuria has received much attention as a source of social ridicule in most parts of societies to the embarrassment of the sufferers. It is regarded rare and not a life threatening disorder, although it may lead to serious psychosocial problems (Todd, 1979; Anon 3, 1998). Anxiety, low self esteem, addiction to drugs and clinical depression are just a few psychosocial problems observed in patients with trimethylaminuria (Basarab *et al.*, 1999). The clinical consequences of this disorder are often grossly underestimated. The major obstacle in the management of trimethylaminuria is the paucity of knowledge among medical and dental specialists concerning the occurrence of this disorder. This is frequently the main source of frustration to the patient who seeks, at least, an explanation for his/her condition (Preti *et al.*, 1992).

There are different figures in different countries in relation to the number of individuals with this disorder as well as the carrier status (heterozygotes) of the affected persons. Studies conducted in Jordan, Ecuador and New Guinea show a prevalence of the carrier status of 1.7%, 3.8% and 11%, respectively (Mitchell *et al.*, 1997).

Countries such as England, Australia, Korea and Italy have also conducted population surveys with regard to the carrier status of their populations. The incidence of heterozygotes of the allele for impaired N-oxidation (Nitrogen-oxidation) in the world-population appears to be of the order of 1%. Therefore, about 1% of people world-wide are likely to be carrying at least one mutated FMO gene, which suggests the presence of several thousands of people with the fish odour syndrome (Ayesh *et al.*, 1993; Cashman *et al.*, 1997). The present study presents an indication of percent presence of impaired N-oxygenation associated with FMO3 enzyme malfunction. This is the first such study performed in South Africa.

2.1.2. Characteristics of Trimethylaminuria

Trimethylaminuria is an autosomal recessive human disorder affecting a small part of the population as an inherited polymorphism (Ziegler, 1993; Cashman, 1997). It is caused by mutations on the flavin monooxygenase 3 (FMO3) gene with the resultant malfunctioning protein product, which is the FMO3 enzyme. The same mutations that cause trimethylaminuria are responsible for the impaired oxygenation of xenobiotics (Anon 2, 1998). The offending smell produced by the affected individuals is essentially the trimethylamine odour. Trimethylamine smells strongly of rotting fish and the human nose can detect it at very low concentrations, that is, less than one part per million (Ayesh *et al.*, 1993). The smell is due to insufficient FMO3 enzyme available to catalyse trimethylamine conversion to its oxide. Trimethylamine oxide has 100 times less olfactory potency compared to trimethylamine (Spellacy *et al.*, 1979; Ziegler, 1993). In some cases the FMO3 produced is simply non-functional (refer to figure 2.1) such that none of the trimethylamine in the body can be metabolised and the accumulation of trimethylamine consequently yields the undesirable smell characteristic of a rotting fish.

The ability to N-oxidise trimethylaminuria is distributed polymorphically in the population, and people with the fish odour syndrome appear to be homozygous for an allele which determines an impaired ability to carry out the N-oxidation reaction (Ayesh *et al.*, 1993).

Individuals diagnosed with trimethylaminuria excrete relatively large amounts of trimethylamine in the urine, sweat and breath resulting in a fishy smell, characteristic of trimethylamine (Al-Waiz *et al.*, 1987b; Cashman *et al.*, 1995; Brunelle *et al.*, 1997). The enzyme system which, when defective, causes trimethylaminuria is known to be age, gender and species dependent and is affected by the exogenous environment (Christensen, 1999; Anon 1, 2000).

The types of trimethylaminuria can generally be classified as (Shelley and Shelley, 1984; Falls *et al.*, 1997; Zschocke *et al.*, 1999):

- Primary genetic trimethylaminuria: an inherited enzyme deficiency.
- Acquired trimethylaminuria: viral infection.
- Transient trimethylaminuria: hormonal modulation.
- Induced trimethylaminuria: intake of enzyme inhibitors and co-substrates (e.g. steroid usage) or the presence of serious liver damage.

The origin of trimethylamine in the diet is derived from eggs, fish, legumes, mayonnaise and other choline, lecithin and trimethylamine precursors (Rettie *et al.*, 1994; Mayatepek and Kohlműeller, 1998). The ingestion of these food products result in their conversion in the gut by the symbiotic bacteria to the malodour trimethylamine as shown in figure 2.1. Some leafy vegetables such as Brussels sprouts, cauliflower, Swedes and cabbage produce goitrin or its precursor (13C product) which serves as an alternative substrate for the FMO enzyme and thus inhibits the efficient conversion of the trimethylamine to its oxide (Spellacy *et al.*, 1979; Shelley and Shelley, 1984; Christensen, 1999).

The inter-relationships of metabolic processes leading to trimethylaminuria as well as the locations where these processes occur are summarised in table 2.1.

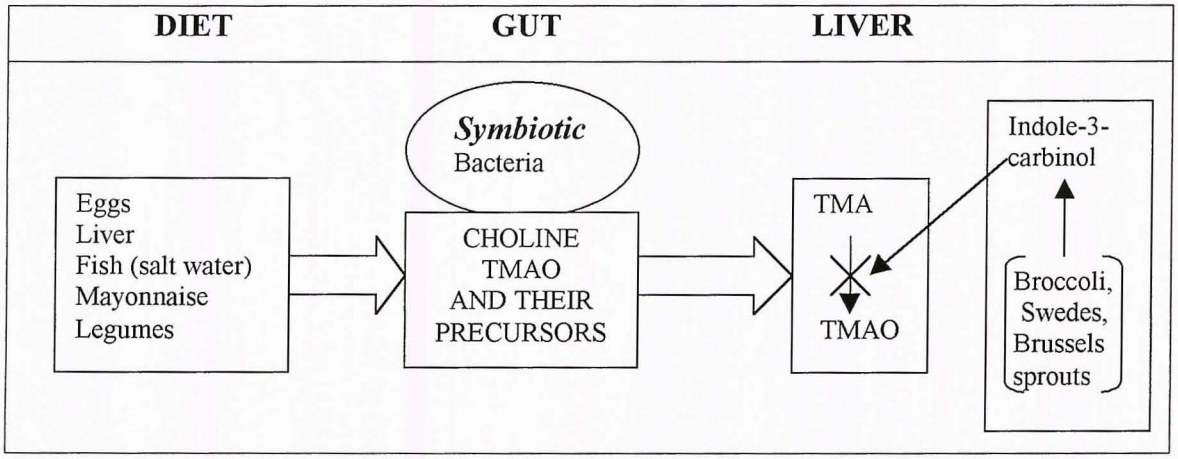

Fig. 2.1: **Metabolic inter-relationships of trimethylamine.** The cross (X) between TMA and TMAO conversion reaction in the liver indicates inhibition by Indole-3-carbinol.

Management of trimethylaminuria includes dietary limitations of foodstuffs containing trimethylamine and its precursors (Cashman *et al.*, 1999b). Usage of acid soaps and lotions also helps by converting TMA (Trimethylamine) to TMAO (Trimethylamine oxide) on the skin (Wilcken, 1994). Affected individuals are also advised to avoid foodstuffs containing inhibitors such as indole-3-carbinol as shown in figure 2.1 (Cashman *et al.*, 1999b).

2.1.3. Substrate metabolism by FMO3 enzyme

The FMO family of enzymes was originally called Ziegler's enzymes, since nobody except Dan Ziegler, who originally identified and described these enzymes, believed they existed (Anon 5, 1999). The present definition of FMO is based only on function and not on structure (Ziegler, 1993; Anon 5, 1999). The general characteristics of this enzyme's gene superfamilies are involved in chemical defence, which makes them well suited to their roles.

They possess the following properties (Anon 5, 1999):

- Wide and overlapping substance specificities.
- Conserved overall catalytic mechanism that allows versatility in reactions catalysed.
- The ability to fit in with other specific/non-specific elimination mechanisms.
- Appropriate localisation.
- Some level of biological variation and/or genetic polymorphism.
- Some physiological roles, that is, roles in dealing with endogenous substrates as well as xenobiotic substrates.

All FMOs are part of the family of the cytochrome P450 (cyt P450) enzymes. Together with other enzymes in this group, they oxidise lipophilic substrates (including trimethylamine) to more polar products, which are then eliminated with urine products through the urinary system (Itagaki *et al.*, 1996; Brunelle *et al.*, 1997). The FMO enzymes serve mainly in the catalysis of phase I detoxification oxidation of the malodour trimethylamine to its smell-free analogue, trimethylamine oxide (Ziegler, 1993). They achieve this by adding a single atom of oxygen as a hydroxyl, ketone or epoxide to their substrate. This is the main elimination route for lipophilic products and to a smaller extent, lipophilic products can be eliminated via diffusion and excretion (Anon 1, 2000).

2.1.3.1. Identity and distinction of FMO enzymes from the other oxygenases

Before 1960, xenobiotic heteroatom-containing compound (HCC) oxidation was thought to be catalysed exclusively by cyt P450 (Rettie *et al.*, 1994). The emergence of the drug metabolising capacity of FMO enzymes has changed the way scientists view human heteroatom drug metabolism. Hence, comparison between cyt P450 and FMO enzymes serves to confirm the FMO enzyme capacity to oxidise heteroatom-containing compounds or drugs.

Generally, oxidative metabolism of HCC by the cyt P450-dependent process leads to products with increased potential for toxic or carcinogenic properties, but exceptions are known to occur. On the other hand, the FMO enzymes generally convert lipophilic HCC to polar, readily excreted oxygenated metabolites that possess decreased toxic potential (Rettie *et al.*, 1994; Itagaki *et al.*, 1996). Overall HCC toxicity is dependent on further metabolic processes, both oxidation and reduction, and notable exceptions to the general statements made above do exist (Rettie *et al.*, 1994).

Large species differences in the rate of microsomal heteroatom oxygenation undoubtedly contribute to determining the way that chemicals and drugs are metabolised to detoxificated metabolites. The relative contributions of FMO and cyt P450 to the metabolism of HCC is determined predominantly by the nucleophilic nature of the heteroatom, with soft nucleophiles being preferentially oxygenated by FMO enzymes (Dolphin *et al.*, 1992). FMOs and cyt P450, both catalyse the same overall reaction (Anon 5, 1999; Anon 1, 2000):

$$\mathbf{RH + NADPH + H^{+} + O_2 \rightarrow ROH + NADP^{+} + H_2O}$$

The following are the main differentiating characteristics of the FMO mechanism of catalysis compared to cyt P450:

- Cyt P450 requires other enzymes for full monooxygenase activity, but FMO enzymes are a complete monooxygenase catalytic unit system. Liver cyt P450 involved in xenobiotic metabolism requires a common reductase enzyme as the accessory enzyme to fulfil its catalytic potential. On the other hand, FMO requires no additional redox partners or accessory enzymes (Ziegler, 1990; Ziegler, 1993; Anon 5, 1999)
- Cofactors for both enzymes are similar since they carry out the same overall reaction. The NADPH-cyt P450 reductase uses NADPH (Nicotinamide adenine dinucleotide phosphate) as a cofactor and contains FAD (Flavin adenine dinucleotide) and FMN (Flavin mononucleotide) as bound coenzymes whereas FMOs contain bound FAD and uses NADPH as a cofactor (Ziegler, 1990; Ziegler, 1993; Anon 5, 1999).

- There is a large group of cyt P450s and the exact number of forms differ between species. In mammals examined so far, there are apparently only five FMO families.
- FMOs appear to be highly conserved compared to cyt P450s (Ziegler, 1993).
- Both enzymes, cyt P450 and FMOs, require NAHPH and O_2 (Oxygen) for activity and their subcellular distribution in most tissues is virtually identical (Cashman *et al.*, 1995; Cashman *et al.*, 1999b). Therefore, one cannot use localisation and cofactor requirements to define the biochemical identity of any of the two enzymes.

The determination of the contribution of a specific enzyme requires its measurement in the presence and absence of selective inhibitors and positive effectors, respectively, along with quantitation products formed (Cashman *et al.*, 1999b).

2.1.3.2. The catalytic cycle of the FMO3 enzyme

Flavin-containing monooxygenases and cytochrome P450 are proteins that play a vital role in both endogenous metabolic functions as well as protecting organisms from the harmful effects of foreign chemicals they encounter in their environment. FMOs discriminate between essential and foreign compounds by excluding the former rather than selectively binding the latter (Ziegler, 1993). A clear understanding of the chemical steps in the catalytic cycle is essential for insight into structural features that may determine, at least in part, the substrate specificity of different forms of the flavin-containing monooxygenases (Ziegler, 1990).

The schematic representation below shows the steps in the catalytic cycle of FMO (Ziegler, 1990; Ziegler, 1993; Rettie *et al.*, 1994; Anon 5, 1999; Anon 1, 2000):

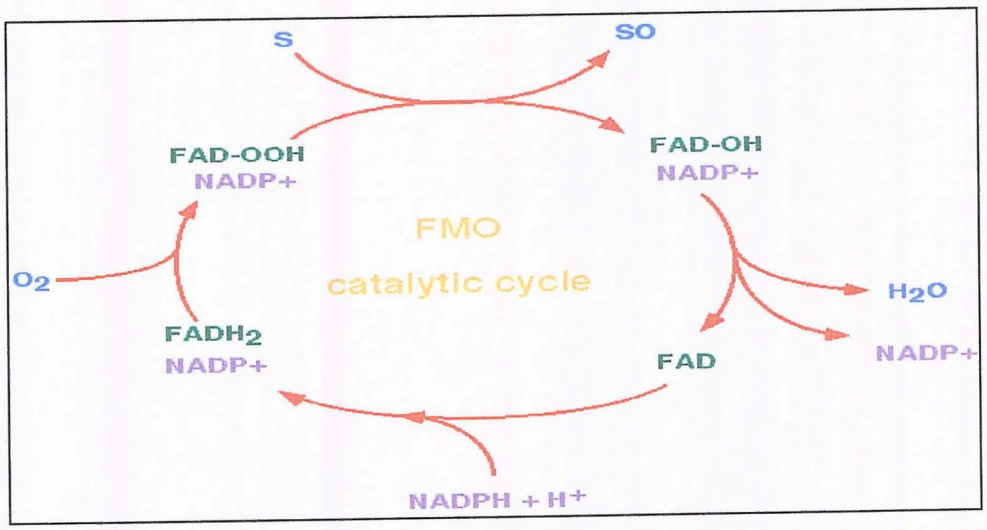

Fig. 2.2: **The catalytic cycle of FMO3 enzyme.**

The enzyme is pre-primed, that is, it is already reduced and has oxygen bound before it encounters the substrate. The FMO forms a stable NADP(H)- and oxygen-dependent 4α-hydroperoxy flavin enzyme intermediate in the absence of an oxygenatable substrate. Any substrate that can gain access to an activated site can be oxygenated. The primed, activated form of the enzyme (which has a hydroperoxyflavin group at the active site, and $NADP^+$ bound) interacts transiently with the substrate, transferring an oxygen atom and leaving a spent, deactivated hydroperoxyflavin form of the enzyme. The monooxygenated product is released immediately after it is formed (Ziegler, 1990; Ziegler, 1993).

The second oxygen is released as water and the spent $NADP^+$ is retained bound to the enzyme. Finally, molecular oxygen is bound and the activated resting form of the enzyme is regenerated. It is in this form that it is ready to hit the next molecule of substrate that wanders into the range (Anon 5, 1999).

Overall catalysis is limited by the decomposition of the flavin pseudo-base (FADOH) or by the release of $NADP^+$ as shown in figure 2.2. $NADP^+$ is a competitive inhibitor of NADPH in the FMO catalytic cycle (Ziegler, 1993).

Since both these steps occur after substrate addition and release of the oxygenated product, the mechanism predicts that at saturation all substrates are oxidised at the same velocity (i.e. $V_{max} = k$). This holds true for tertiary amines whereas for primary amines the oxidation velocity is half that of the former substrate (Cashman, 1995).

The overall size of the nucleophile appears to be a major factor that limits access to the 4α-hydroperoxyflavin in different FMO isoforms, and by simply measuring the oxidation of a functional group bearing substituents of increasing size, such differences should be readily identifiable. This pattern is known to be tissue specific (Rettie *et al.*, 1984; Ziegler, 1993).

There is no evidence that substrate binding lowers the energy of activation for oxygen transfer from the hydroperoxyflavin to substrate, or that, precise fit to substrate at more than one point on the enzyme is even necessary (Ziegler, 1990; Ziegler, 1993). FMO isoforms are thus present in the cell in their oxygen-activated form and any soft nucleophilic substrate accessible to the enzyme-bound peroxyflavin intermediate will be oxidised. Because the energy for catalysis is present in the enzyme before contact with the xenobiotic, the fit of the substrate is not nearly as stringent as with most enzymes. This feature is unique to the FMO among monooxygenase and is responsible for the extraordinary range of substrates accepted by these flavoproteins (Hines *et al.*, 1994; Anon 5, 1999).

A single point of contact between the substrate and the terminal oxygen of the hydroperoxyflavin is all that is required for product formation. It is this property, unique to FMO isoforms, which is responsible for the extraordinary broad substrate specificity of these flavoenzymes. This is because the energy required to drive the reaction is present in the enzyme before it encounters the substrate (Ziegler, 1993; Cashman, 1995). FMO3 has a unique ability to oxidise structurally dissimilar compounds with equal rates because perfect fit for the catalytic site of a substrate to an FMO enzyme is not required.

In general, it is conceivable that different structural isoforms of FMO could have the same substrate size limit and activity measurements alone could easily miss or underestimate the total number of FMO gene products in a species.

These measurements do, however, provide a convenient and rapid method for measuring the size limits of substrates accessible to these enzymes in crude tissue preparations.

This information can be used to predict access of other types of soft nucleophiles to the active sites of FMO isoforms present in that tissue preparation. Because these flavoproteins discriminate among soft nucleophiles by excluding non-substrates rather than selectively binding their substrates, it is very likely that the substrate specificity of isoforms 1 through 5 will be the same across species and tissues (Ziegler, 1993).

Size restrictions for substrates accepted by different FMO isoforms determined with thiocarbamides can be extrapolated to substrates bearing functional groups. The species distribution of isoforms determined by activity with thiocarbamides also offers an explanation for the differences in the role of FMO in N-oxidation of other substrates by microsomal enzymes from other species (Ziegler, 1993). Since the hydroperoxyflavin is common to all forms of the enzyme, it is evident that differences in access are largely responsible for differences in substrate specificity of various forms of the enzyme (Cashman, 1995).

Despite the lack of tight binding, FMO-catalysed reactions are fully enzymatic and follow saturation (Michaelis-Menten) kinetics. While a single point of contact between the substrate and the enzyme-bound oxidant will often suffice, this does not preclude more complex interactions with some substrates of a specific FMO. In addition, the observation that K_m for a homologous series of substrates bearing the same functional group often decreases with increasing lipophilicity does not invalidate the preceding analysis. This may be due to non-specific absorption of the more lipophilic analogues on the enzyme (Ziegler, 1993).

Human FMO3 activity is decreased by indole-3-carbinol and its condensation products. These chemical compounds are not substrates for FMO3 in contrast to many inhibitors; hence, they can be used as potent selective inhibitors of the FMO catalytic system (Cashman, 1995).

"Unlike other oxidases, FMO sits in the cell in the loaded, cocked position and any suitable target that wanders within range is oxidised!" These are the words of Henry Kamin concisely summing up the distinguishing facts about FMOs compared to other oxygenases (Ziegler, 1990).

2.1.3.3. Specificity and general reactivity of nitrogen (N)-, sulphur (S)- and phosphorus (P)-centres metabolised through FMO3

Only xenobiotics with an electron-rich (nucleophilic) centre can serve as substrates for FMO3 and be converted to readily excreted polar products (Cashman, 1995; Cashman *et al.*, 1997; Lang *et al.*, 1998;). The hydroperoxyflavin is not a sufficiently strong oxidant to attack the nitrogen atom in amides, carbamides, imines, nitrones, oximes etc. Amines, hydroxylamines and hydrazines are more susceptible to oxidation, and xenobiotics containing these groups are usually excellent substrates for FMO isoforms (Rettie *et al.*, 1994).

- In most molecular configurations, sulphur is generally quite nucleophilic and almost every type of functional group bearing sulphur shows substrate activity. Sulphoxidation catalysed by FMO appears to be a major route for the detoxification of drugs and pesticides bearing sulphide side chains. The oxidation of thiocarbamides and mercaptoimidazoles is catalysed exclusively by FMO (Ziegler, 1990).

- FMO isoforms appear to be ideally adapted to catalyse the detoxification of structurally diverse soft nucleophiles, e.g. alkaloids with basic side chains and organic sulphur xenobiotics. FMO isoforms share a mechanism distinct from all other oxidases bearing flavin, heme or other redox-active prosthetic groups (Christensen, 1999).

- Of the functional groups bearing nitrogen, only amines, hydroxylamines and hydrazines are sufficiently nucleophilic to serve as substrates for FMO. However, even with these restrictions, the synthetic and naturally occurring nitrogen substrates for FMO include a large number of alkaloids and medicinal amines. The medicinal amines include all antihistamines (histamine receptor antagonist, e.g. ranitidine), monoamine oxidase inhibitors (some antidepressants) and tricyclic antipsychotic drugs with basic side-chains (Zhang *et al.*, 1995).

FMO readily catalyses the oxidation of mono-cationic amines or of anionic sulphur compounds where the charge is localised on sulphur, but the addition of a second charge group anywhere on the molecule generally blocks substrate activity. A second charged group on the molecule, either anionic or cationic, apparently blocks access to the catalytic site of the FMO. Only uncharged amines or those bearing a single positive charge at physiological pH are substrates (Cashman, 1995). This effectively excludes amino acids and other possible biogenic substrates.

This suggests that the position and number of ionic groups are the principle factors that enable the enzyme to discriminate between essential and xenobiotic soft nucleophiles. Impaired oxygenation of nicotine, which is the alternative substrate for FMO, has been observed in patients suffering from trimethylaminuria (Rettie *et al.*, 1994).

FMO substrates appear to be good nucleophiles. FMO isoforms cannot monooxygenate carbon centres. This is a critical, key functional difference of these enzymes from cyt P450 (Anon 1, 2000).

Other metabolic enzymes (e.g. cyt P450) can oxygenate carbon centres but also catalyse the oxidation reaction of nitrogen, sulphur, oxygen, phosphorus and selenium atoms contained in their substrates. This means that FMO isoforms in general have less oxidising capacity than cyt P450 isoforms, since they cannot function on catalysis of compounds without a heteroatom (usually N or S) (Lomri *et al.*, 1992; Anon 1, 2000). However, FMO enzymes catalyse the oxidative metabolism of a variety of N-, S- and P-containing compounds, some of which are of toxicological importance (Overby *et al.*, 1997). Thus, FMO3 appears to be the most important enzyme form for human drug metabolism. In contrast to cyt P450s, there is a clear sequence similarity between FMO orthologs in humans and animals (Ziegler, 1993).

2.2. The molecular study of trimethylaminuria.

2.2.1. Introduction.

Trimethylaminuria is to a larger extent an inborn error of metabolism, except in few cases wherein it is a result of external factors such as liver damage (Shelley and Shelley, 1984) or induction by viral infection (Zschocke *et al.*, 1999). To gain insight and understanding towards combating trimethylaminuria as well as other associated diseases, it is important to study the molecular basis and fundamentals of this disorder.

The coding (Dolphin *et al.,* 1996) and intronic (Dolphin *et al.*, 1997a) nucleotide sequences of the FMO3 (flavin-containing monooxygenase 3) gene have already been sequenced. The FMO3 gene is transcribed into RNA (ribonucleic acid) and finally translated into a functional FMO3 enzyme. Any change in the nucleotide sequence of the FMO3 gene, whether in the intronic or exonic sequence may have unfavourable effects leading to trimethylaminuria. The severity of trimethylaminuria is dependent on the character and type of base sequence changes.

2.2.2. The FMO genes and their protein products.

FMO isoforms belong to the flavocytochrome c sulphide dehydrogenase subfamily of flavoenzymes, NAD(P)H-dependent monooxygenases and reductases (Akerman *et al.*, 1999b). The flavin-containing monooxygenase gene family contains five members that are expressed in a species- and tissue-dependent manner (Overby *et al.*, 1997; Kawaji *et al.*, 1997).

These five genes have been mapped to the long arm of chromosome 1 and the FMO3 gene is in the region 1q23-24 (Dolphin *et al.*, 1997b). FMO1 through to FMO5 are each a product of a single gene (Dolphin *et al.*, 1997a). Of the five FMO isoforms, only FMO3, 4 and 5 are expressed in adult liver, with mRNA (messenger ribonucleic acid) abundance in the order FMO3 > FMO5 » FMO4 (Dolphin *et al.*, 1997a).

In the foetus FMO1 is expressed in both the kidney and the liver whereas in adult humans, FMO1 is expressed only in the kidney (Dolphin *et al.*, 1996). The major FMO enzymes in human liver are FMO3 and FMO5. Quantitation of FMO3 and FMO5 with monospecific antibodies and recombinant isoforms as standards showed levels of FMO3 to range from 12.5 – 117 pmol/mg and 3.4 – 3.5 pmol/mg for FMO5 (Overby *et al.*, 1997). These amounts were quantified from human hepatic microsomal samples. The FMO3 concentration is consistently greater than FMO5 with ratios varying from 2:1 to 10:1 (Overby *et al.*, 1997). Further research show that there is little relationship between the expression of FMO3 and FMO5. There is also no relationship between levels of mRNA and levels of enzyme protein. Transcripts for FMO4 have been detected in samples from human liver, but no evidence for protein expression has been reported (Dolphin *et al.*, 1996)

2.2.2.1 The molecular structure of the FMO3 gene and its protein product.

The overall length of the FMO3 cDNA (complementary deoxyribonucleic acid) is 1913bp (base pairs). It contains a 5′-flanking region of 93bp, an open reading frame of 1596bp, followed by a stop codon and 221bp of part of the 3′-untranslated region. The open reading frame encodes a polypeptide of 532 amino acids, with a calculated mass of 60047Da (Ziegler, 1980) and an estimated pI (isoelectric point) of 8.3 (Phillips *et al.*, 1995)

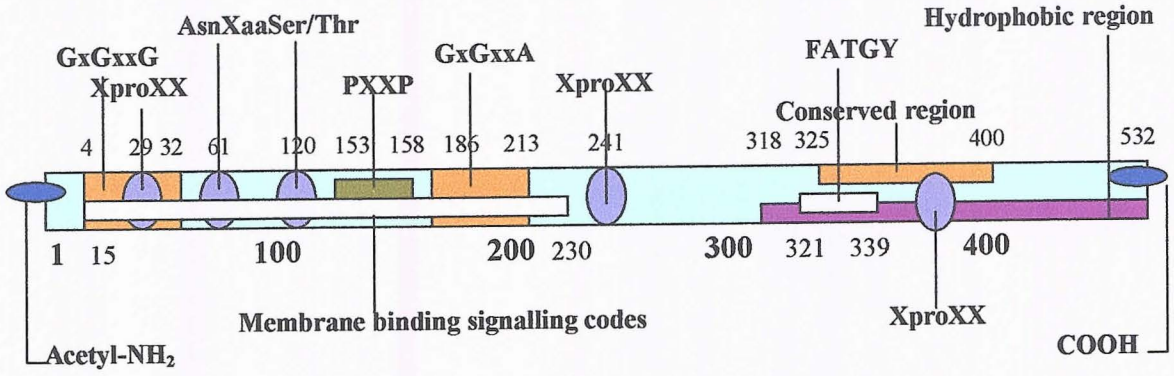

Fig. 2.3: **The FMO3 enzyme model**. The abbreviated text represents all the functional regions of the enzyme and their positions in the polypeptide sequence. The diagram is not drawn according to scale.

The FMO3 gene contains nine exons. One of these exons is noncoding whereas the remaining eight are coding exons (Overby *et al.*, 1997). Exon 1 is noncoding and exon 9 encodes the entire 3′-untranslated region of the corresponding mRNA. The translation initiation codon is located at the same relative position within exon 2 (Dolphin *et al.*, 1997a).

2.2.2.2. The FAD- and NADPH-binding domains

FMOs contain one molecule of FAD per monomer. Residues 1 – 200 as well as 450 – 532 are highly conserved regions indicating important amino acid structural and functional domains (figure 2.3). Two regions within the primary sequence of human FMO3, residues 4-32 and 186-213, exhibit complete agreement with a consensus fingerprint sequence that predicts the occurrence of a βαβ-fold (Rossmann-fold) characteristic of a domain that binds the ADP (Adenosine diphosphate) moiety of dinucleotide factors (Phillips *et al.,* 1995). Residue region 4-32 is the FAD-binding domain whereas residue region 186-213 is the NADPH-binding domain. These domains are present at the same relative positions within the primary structure of all mammalian FMOs. Examination of the internal organisation of the human FMO3 gene reveals that the entire fingerprint sequence involved in binding the ADP moiety of FAD is encoded by exon 2 (Dolphin *et al.*, 1996; Kubo *et al.*, 1997)

The FAD-binding domain is known to be of the sequence GxGxxG. The three glycines, especially the second and third glycines, in the FAD-binding domain are necessary for FMO to show catalytic activity (Dolphin *et al.*, 1996; Kubo *et al.*, 1997).

The NADPH-binding domain sequence is GxGxxG/A. These sequences are not only found in FMO enzyme proteins, but in many other nucleotide-binding proteins including flavoenzymes, pyridine-nucleotide-dependent enzymes and protein kinases (Kubo *et al.*, 1997).

Three-dimensional analysis of several dinucleotide pyrophosphate-containing proteins demonstrates that (Kubo *et al.*, 1997):

- The first glycine is necessary for a tight turn of the main chain with special angles.
- The second glycine associates with dinucleotide pyrophosphate moiety.
- The third glycine is important to provide space for close interaction between the beta-sheet and the alpha-helix.

Conversion of the first glycine into alanine produces slight effects on the FAD binding site and the conformation change between beta-sheet and alpha-helix. This is because the physical properties of glycine might be similar to that of alanine. Since the 4α-hydroperoxyflavin with a substrate is required for the oxidation of the substrate, the substitution of the first glycine to alanine may prevent the substrate approach to 4α-hydroperoxyflavin (Kubo et al., 1997). The substitution of the second glycine to alanine in the FAD binding domain causes the inability of the protein to bind to FAD. Alteration of the third glycine to valine changes the conformation between beta-sheet and alpha-helix (Kubo *et al.*, 1997)

2.2.2.3. Other conserved sequences.

Conservative amino acid changes on the upstream of the FAD-binding domain of FMO3 results in an enzyme that retains activity The N-terminus of FMO3 does not have any discernible signal peptide sequence. The presence of the FAD-binding region towards the N-terminus suggests that this region is not a functional insertion domain (Rettie *et al.*, 1994).

Sequences known to be unique for mammalian FMOs include residues 321 – 339 (motif FATGY) as well as codons 325 – 400 which is highly conserved (Akerman *et al.*, 1999b). Codons 153 – 158 (PXXP motif) are implicated in protein folding.

The sequence and location of the PXXP motif are conserved in all known mammalian FMOs with the exception of FMO5, which contains an alanine at this position and exhibits markedly reduced catalytic activities towards typical FMO substrates (Akerman *et al.*, 1999b). Most of the FMO cDNAs sequenced to date contain the highly favourable translation initiation site, XXATGG, where X is cytosine. Coding region changes in the nucleotide region 45 – 690, which is equivalent to amino acid residues 15 – 230 cause dramatic differences in the physico-chemical properties of FMOs as observed on SDS-PAGE (Rettie *et al.*, 1994).

In FMO1 – 3, the initiation amino acid methionine is not present and the following amino acid is N-acetylated (Dolphin *et al.*, 1997b). Residues such as alanine or glycine near the N-terminus apparently promote the removal of methionine during FMO protein saturation. With the exception of FMO4, all FMOs sequenced to date have at least a single putative consensus N-glycosylation site with the following amino acid sequence: Asn-Xaa-Ser/Thr (Zhang *et al.*, 1996).

FMO isoforms possess very strong membrane association properties. While all FMO isoforms possess a number of hydrophobic regions, only the C-terminus (residues 318 – 533) is sufficiently hydrophobic to be a membrane-insertion sequence (Rettie *et al.*, 1994). However, truncation of C-terminus (26 amino acids) from rabbit FMO2 did not lead to loss of membrane associated properties when the remaining FMO cDNA is expressed in *E. coli.* This fact suggests that FMO membrane association is not a passive event that is dictated exclusively by hydrophobic C-terminal amino acid residues, rather, the information for active FMO membrane association is probably encoded in an integral sequence proximal to the N-terminus (residues 15 – 230) (Rettie *et al.*, 1994).

N-linked glycosylation, as shown in figure 3.1, often occurs at residue 61 (Asn^{61}) within the conserved motif with the following sequence: Asn-Xaa-Ser/Thr. O-linked glycosylation occurs at residues 29, 241 and 381 ($Thr^{29/241/381}$), respectively, and the conserved motif has the following sequence; Xaa-Pro-Xaa-Xaa, where at least one Xaa is threonine (Zhang *et al.*, 1996). However, glycosylation status does not have a significant influence on the catalytic functions. This was shown to be the case in FMO3 enzyme derived from the baculovirus expression system (Zhang *et al.*, 1996).

2.2.3. Mutations in the FMO3 gene.

It has yet to be established whether the mutation(s) responsible for trimethylaminuria affect the expression of the gene or the activity or stability of the encoded protein (Dolphin *et al.*, 1997a). The effects of specific mutations on the overall activity of the FMO3 enzyme have been characterised as shown in Appendix V. The relative positions of common mutations in relation to the FMO3 exons are shown in figure 2.4.

Fig. 2.4: **The FMO3 gene model with its common mutations**. The mutations are shown in red (severe trimethylaminuria) and green (mild trimethylaminuria) letters to designate their severity. The numbers in the bold lined (blue) boxes indicate the relative size of each exon. The arrows indicate the position of each mutation relative to the exon. The exon names are exon 1 through to 9, from left to right. The red lines indicate the size of the translated region.

2.2.3.1. The M66I mutation

It is possible that the methionine of codon 66 may constitute an important secondary starting point for translation mechanism, which would be perturbed by the isoleucine substitution. This is essentially a missense mutation. It may also be possible that this substitution may result in abnormal protein folding, resulting in an inactive protein. Since this mutation falls in the nucleotide region 187 – 213, which is a βαβ-fold that binds either the ADP moiety of NADP or FAD, the substitution of methionine by isoleucine may possibly interfere with the affinity of either cofactor towards the binding site and thus incapacitates the FMO3. This mutation causes trimethylaminuria (Cashman *et al.*, 1997).

2.2.3.2. The P153L mutation

P153L is a missense mutation on the FMO3 gene in codon 153, which substitutes C→T corresponding to nucleotide 551 in the cDNA in exon 4. It is known to abolish activity of FMO3 enzyme (Dolphin *et al.*, 1997b; Basarab *et al.*, 1999).
This mutation replaces proline with leucine at position 153. The marked effect of the Pro-Leu substitution on both the N- and S-oxidation activities of FMO3 clearly identifies Pro 153 as important for the structure and function of the enzyme (Dolphin *et al.*, 1997b).

Proline 153 is the first residue of a PXXP motif located between the fingerprint sequences identifying the βαβ-folds of the FAD- and NADP- binding domains, 33 residues upstream of the latter sequence (Akerman *et al.*, 1999b; Phillips *et al.*, 1995). With respect to the three dimensional structure, this position may be strategically located to aid in the direction of the binding sites to the proximity of NADP and FAD of the FMO3. This may be the reason why the catalytic activity of FMO3 is abolished (Cashman *et al.*, 1997). The PXXP motif is highly conserved and the consequent abolition of enzyme activity underlies the importance of P153 to enzyme function (Basarab *et al.*, 1999).

2.2.3.3. The E158K mutation

This mutation occurs at codon 158 and involves the replacement of glutamine with lysine and its effect towards FMO3 activity is negligible (Dolphin *et al.*, 1997b). However, its effect in combination with other mutations is yet to be established. The glutamic acid 158 and lysine 158 alleles represent a common FMO3 polymorphism. *In vitro* expression studies on this polymorphism indicates that the K158 form of the protein is a poorer TMA N-oxygenator than the E158 form *(*Akerman *et al.,* 1999b*)*.

In some studies, 30% trimethylaminuria probands exhibited labile hypertension, which is suggestive of disordered biogenic amine metabolism (e.g. migraine, disordered metabolism of tyramine) *(*Akerman *et al.,* 1999b; Brunelle *et al.*, 1997*)*.

The sequence surrounding the FMO3 substitutions P153L and E158K involves a PXXP motif, which is implicated in protein folding. The sequence and location of the PXXP motif are conserved in all known mammalian FMOs with the exception of FMO5, which contains an alanine at this position and exhibits markedly reduced catalytic activities towards typical FMO substrates *(*Akerman *et al.,* 1999b*)*. Lys^{158} and Glu^{158} are both active against the N- and S-heterocentre substrates although the extent of substrate oxygenation shows considerable sensitivity to the particular polymorph used (Cashman *et al.*, 1997).

Physical and chemical properties are identical for both polymorphs when expressed in *E. coli*, although they have different capabilities to discriminate between the N- and S-containing substrates. There is increased relative activity towards the S-substrate for Glu^{158}, which is 20 times more than that of the N-substrate. For the Lys^{158}, there is almost 5 times more activity towards the S-substrate than for N-oxygenation (Lang *et al.*, 1998).

2.2.3.4. The R492W mutation

The R492W mutation is found in the hydrophobic domain (300 - 500). It involves a hypermutable CpG site and is a non-conservative change in a highly conserved region of the FMO3 gene. Arginine, at codon 492, is conserved in all FMO isoforms (Akerman *et al.*, 1999b). Its apparent conservation may be vital for some structural function since this position does not seem to be part of the catalytic site.

2.2.3.5. The E314X mutation

The E314X mutation is located in the hydrophobic region (300 - 500), which is thought to be partly responsible for FMO3 membrane binding capability. This is essentially a nonsense mutation, resulting in production of an incomplete protein product and thus functionally inactive. This nonsense mutation reduces the size of the FMO3 enzyme by 36 amino acids.

Loss of activity is strongly predicted from truncation of the FMO3 protein at codon 314. This is due to the fact that deletion of the final 30 amino acids of this 532 residue protein does ablate function *in vitro (*Akerman *et al.,* 1999b*)*.

2.2.3.6. A52T and R387L

The A52T mutation is positioned between the codon 32 XProXX motif and the codon 61 AsnXaaSer/Thr motif. It is a missense mutation. Its proximity to the XProXX domain, which binds the FAD, may possibly reduce the affinity with which the enzyme binds to FAD. On the other hand, mutation R387L is also a missense substitution.

It is found in the conserved region (325 - 400). The A52 and R387 residues appear to be highly conserved in the FMO genes hence satisfying the criteria for designation as disease-causing *(*Akerman *et al.,* 1999b*)*.

2.2.3.7. The E308G mutation

This is a polymorphism that may also indicate other variations of drug and chemical detoxification. It is apparently always linked to E158K (Zschocke and Mayatepek, 2000). It is located in the silent region (242 - 318) in which no mutations or functional domains have been described with respect to enzyme activity alterations.

2.2.3.8. The V143G mutation

This is an exon 4 missense polymorphism. It involves the substitution of valine by glutamic acid at codon 143 (Basarab *et al.*, 1999). It is known to reduce overall activity, although by a relatively small margin. This mutation together with mutation E158K may result in trimethylaminuria because both codons fall under the critical domain region (Sasche *et al.*, 1999). The critical domain region wherein the entire essential binding domains are found spans from amino acid 15 to 230.

2.2.3.9. Summary

Mutations appear to cluster in exon 7, with three of the four identified changes in this exon occurring within a ten amino acid 'hotspot'. They are in close proximity to the FATGY signature found in an area highly conserved among FMOs, strongly suggesting that this substitution affects protein function *(*Akerman *et al.*, 1999b*)*.

Mutations 158K, 153L, 199T and 475D are described as mild trimethylaminuria-causing substitutions. Mild FMO3 deficiency may be expected to slow breakdown of biogenic amines and may constitute a risk factor for hypertension and cardiovascular disease (Zschocke and Mayatepek, 2000). Differences in N-oxidation capacity may contribute to the inter-individual variations that are observed in the metabolism and kinetics of nicotine. Impaired breakdown of certain drugs, including substances used for chemotherapy, may explain drug toxicity and unwanted side effects observed in some patients (Zschocke and Mayatepek, 2000).

With the number of possible mutations in the FMO3 gene as well as their effect on the enzyme's activity, it is clear that FMO3 deficiency is not merely a rare recessive disorder. It is rather a spectrum of phenotypes of transient or mild malodour depending on environmental exposure (Zschocke and Mayatepek, 2000).

2.2.4. Regulation of the FMO gene expression

Like many other genes, FMO gene expression is undoubtedly under some sort of regulation. The mechanism of regulation warrants further research, yet the following facts illustrate the extent to which gene expression is regulated:

1. Evidence for induction of FMO2 in rabbit lung as a consequence of plasma hormone levels has been obtained although the physiological role for the increased FMO2 has not been ascertained. De-induction has also been observed (Rettie *et al.*, 1994). Induction of different forms of the FMO enzyme in specific tissues by hormones may affect metabolism of numerous drugs.

   - Lung FMO in rabbits increases at least five-fold during pregnancy (Cashman, 1995). This is more evident at days 15 and 28 – 31, which correlate with progesterone and corticosterone plasma concentration peaks (Rettie *et al.*, 1994).
   - Gender-related differences for mouse liver FMO appear to be due to testosterone repression of the hepatic enzyme. Rat liver FMO levels are apparently positively regulated by testosterone and repressed by estradiol (Rettie *et al.*, 1994).
   - Dietary xenobiotics influence rat liver FMO. In the absence of dietary xenobiotics, rats maintained on total parenteral nutrition for seven days, a 75 – 80% decrease in FMO activity was observed (Rettie *et al.*, 1994).
   - Trimethylaminuria has been reported to be exacerbated at puberty and steroid hormones have been shown to influence FMO activity in rodents and man. The activity of aberrant forms of the enzyme in such individuals may be more susceptible to hormonal modulation causing transient trimethylaminuria (Falls *et al.*, 1997).

2. The gene encoding FMO1 is expressed in several foetal tissues, albeit to different extents. During development, the expression of the FMO1 gene is switched off in some tissues, notably liver, but is maintained in the kidney. The lack of expression of the FMO1 gene in adult human liver is in marked contrast to the situation in other mammals such as pig and rabbit, in which FMO1 constitutes a major form of the enzyme in the liver of the adult animal (Phillips *et al.*, 1995).

3. Expression of the FMO3 gene is selectively increased in the liver during development. Thus, FMO1 and FMO3 genes of man are both subject to marked developmental and tissue-specific regulation (Whetstine *et al.*, 2000; Haining *et al.*, 1997).

4. Due to the pronounced differences in tissue-specific levels of mRNA, protein levels, and FMO activity in rabbit, it is possible that FMO gene expression is regulated at the transcriptional stage. However, post-translational modifications of FMO do not appear to play a role in modulating FMO1 – 3 activity (Rettie *et al.*, 1994).

In contrast to genes encoding cyt P450 monooxygenases, FMO genes display relatively little inter-individual variation in their expression (Phillips *et al.*, 1995). As FMO does not respond to the types of inducing agents associated with other monooxygenases, it is likely that FMO is not co-ordinately regulated with these other enzyme systems (Rettie *et al.*, 1994).

FMO expression is both tissue- and species-dependent. Therefore, in a given tissue it is probably the particular profile of FMO isoforms present that determines enzyme activity, with the dominant isoform responsible for substrate specificity and stereoselectivity (Rettie *et al.*, 1994; Kawaji *et al.*, 1997).

2.3. Mutation detection techniques.

2.3.1. Introduction to mutation detection methods

This study is primarily concerned with developing a protocol that can be used for effective mutation detection in the FMO3 gene. It is therefore essential that a brief but detailed background of the mutation detection techniques be discussed before the procedural methods are outlined.

Mutation detection is vital to the whole of biology including medicine. In medical research, mutation detection is fundamental to disease gene isolation, mutation spectrum studies and diagnosis (Forrest, 1998). It is in the context of disease diagnosis (severe or mild trimethylaminuria) that mutation detection will be discussed in this study.

Inherited trimethylaminuria is mainly caused by mutations in the FMO3 gene that result in the impaired oxygenation of the prime substrate, namely, trimethylamine (Ziegler, 1993). In an attempt to characterise the mutations that cause this disorder, cost effective and efficient protocols need to be established and developed.

Generally, mutation detection is time-consuming and expensive. There is still no perfect method developed yet in terms of cost effectiveness, simplicity and 100% mutation detection.

Therefore, the choice of a mutation detection method becomes awkward and cumbersome when one does not have a full understanding of the principles involved in each of the available methods. Various applicable methods will be discussed so as to give a brief background of each. Sue Forrest (1998) classified mutation detection methods into two groups. The first is the scanning mode, which searches for unknown mutations and the second is the diagnostic mode, which tests for known mutations in a known nucleotide sequence. The prime considerations in any approach to mutation detection, irrespective of the type of the technique used, include sensitivity (proportion of the mutations that can be detected) and specificity (absence of false positives).

The choice of a specific method depends mainly on the mode desired. In the case of the scanning mode, the following factors are critical:

- Simplicity and versatility (Hayashi *et al.*, 1998)
- Ability to detect near 100% of mutations (Forrest., 1998)
- Ability to scan longer lengths (Taylor, 2000)

Although the above factors are major criteria for method selection, budget requirements, laboratory equipment availability and percent detection required are limitations that often dictate the type of method to be employed. Some of the scanning mode methods frequently used for mutation detection include:

- Single stranded conformation polymorphism (SSCP)
- Heteroduplex analysis (HA)
- Denaturing gradient gel electrophoresis (DGGE)
- Mismatch cleavage analysis

Mutations that can lead to mild or severe trimethylaminuria have not yet been fully exhausted in terms of detection and characterisation. For this reason, scanning methods take precedence to diagnostic methods in this study and it is the former technique that will be discussed in more detail. However, it is imperative to perform diagnostic tests in suspected trimethylaminuria patients to confirm absence or presence of the already known disease causing mutations.

For the detection of known mutations, restriction enzyme analysis stands prominent as a tried and tested method although amplification refractory mutation system (Taylor, 2000) and solid phase mini-sequencing have been widely used. Electrophoresis in general presents one of the best methods for detection of mutations. All of the scanning techniques mentioned in this chapter are based in part on the electrophoresis principle. The term electrophoresis is applied to the movement of small ions and charged macromolecules in solution under the influence of an electric field (Grierson, 1985).

The rate of migration depends on the size and shape of the molecule, the charge carried, the applied current and the resistance of the medium (Helms, 1990; Olckers, 1999).

There are many modifications of this technique to suit different types of mutation detection protocols. The common methods for detecting single base changes in genomic DNA involves restriction fragment length polymorphism, abbreviated RFLP (Gedil, 2000), SSCP, DGGE and mismatch cleavage techniques (Taylor, 2000). The RFLP method, however, is ineffective for screening new random mutations.

2.3.2. Single stranded conformation polymorphism (SSCP) and Heteroduplex analysis (HA).

2.3.2.1. Introduction

SSCP and HA are two widely used techniques due to their simplicity, either in combination with each other or separately (Hayashi, 1996; Olckers, 1999; Macek, 2001). SSCP and HA stand out as well established methods of choice in mutation detection (McQuaid, 1997).

In SSCP, single-base substitutions can be detected due to differential, sequence specific, electrophoretic mobility of the single strand DNA. The general methodology includes DNA segment amplification, denaturation using heat and formamide followed by separation by non-denaturing electrophoresis with distinct sieving properties (Macek, 2001).

Single strand conformation polymorphism and DGGE methods were developed as new alternatives for detecting and localising single base changes (King, 1999).

Of all the methods mentioned in this study, SSCP is the closest competitor to DGGE with respect to sensitivity and specificity. Its main advantages being simplicity and versatility (Macek, 2001).

2.3.2.2. Single stranded conformation polymorphism

SSCP relies on the formation of secondary structures by single-stranded DNA. The formation of the secondary structure is the result of interaction between respective nucleotides within the single-stranded molecule. The interactions themselves are determined by the DNA sequence, such that a single base substitution may result in an alternative secondary structure. The three-dimensional structures that result from intra-strand interactions contain hairpin-like and looped regions (Taylor, 2000).

Therefore, similar single stranded-DNA folds differently from another if it differs by a single base and mutation-induced changes of tertiary structure of the DNA results in different mobilities for the two strands in polyacrylamide gel (McQuaid, 1997). This forms the basic principle of SSCP. Mutations are thus detected as the appearance of new bands or new band positions on autoradiograms (Gerrard and Dean, 1998).

Electrophoretic mobility of a single-stranded DNA molecule in a gel matrix is sensitive to size, shape and charge of the molecule (Olckers, 1999). The success of SSCP electrophoresis or detection resolution depends on whether the substitution affects the three-dimensional folding pattern of the conformer and/or the effect it presents on the electrophoretic mobility of the conformer. Under suitable electrophoresis conditions mobility changes caused by sequence modifications can be resolved with sensitivities ranging from 60 to over 95% in SSCP (Taylor, 2000; Macek, 2001). Hayashi (1996) estimates the sensitivity of SSCP to be of the order 80% in a single run for fragments shorter than 350bp.

Disadvantages and limitations of SSCP include the following:

- One cannot predict conditions for sequence variants in a particular sequence context, especially when the mutation has not been identified before. There is also an apparent low detection in SSCP analysis (Olckers, 1999).
- There are no empirical rules with regard to detectability of mutations in general. For fragments up to 350bp, sensitivity is estimated at 80% in a single run. With careful optimisation, over 90% of mutations can be detected in the DNA fragment of suitable size (Macek, 2001; Taylor, 2000).

- Low detection rate of G → C transitions compared with other detection techniques. These changes are therefore likely to be missed (MRC course, 2000).

The main draw back to this method is the need for a radioactive protocol (Hayashi, 1996; Kurihara *et al.*, 1999). The lack of objective predictability in SSCP, which means that, non-appearance of a band-shift does not prove the absence of a mutation, is another disadvantage in direct contrast to DGGE (Macek, 2001).

Some of the advantages of SSCP are:

- The low cost of experimental equipment.
- General simplicity of the method (Claustres, 2001). SSCP is rapid to perform and can be carried out using equipment available in most molecular biology laboratories (Gerrard and Dean, 1998).

2.3.2.3. Heteroduplex Analysis

Heteroduplex is double-stranded DNA in which the two DNA strands do not show perfect base complementarity (Dean, 1996; McQuaid, 1997). When DNA is denatured, the two strands are separated. On renaturation or annealing, complementary DNA strands reassociate and form homoduplex. On the other hand, PCR products from a heterozygote form heteroduplexes with complementary strands from a wild type allele. This forms the basis of mutation screening of heterozygous individuals using HA (Glavac and Dean, 1996).

The electrophoretic mobility of the heteroduplex in a polyacrylamide gel is not the same as that of the homoduplex, and one of the two bands will move slower in comparison to the other and thus detect the difference in conformation of the nucleotide fragments concerned. In fragments of under 200bp, insertions, deletions and most single base substitutions can be detected (McQuaid, 1997). Heteroduplex analysis is mostly used in combination with SSCP.

The sensitivity and specificity of SSCP is known to improve when applied in combination with HA. This is achieved by running longer gels that allow detection of residual double strand heteroduplexes in a given lane. When properly optimised, this combination is highly sensitive and many different mutations within a DNA fragment can often be distinguished on the same gel (Gerrard and Dean, 1998).

After denaturation of the sample prior to loading, there is often the reformation of a significant amount of double-stranded DNA which will appear in the lower position from the single-stranded products on the SSCP gel. Since heteroduplexes can be resolved from homoduplexes, the appearance of the heteroduplexes can give additional information on the presence of variants (Gerrard and Dean, 1998; Olckers, 1999). This essentially constitutes the basic principle of PCR-SSCP/HA (polymerase chain reaction-single stranded conformation polymorphism/heteroduplex analysis) technique. In this way, variations in the nucleotide sequence of a specific fragment are observed both in the upper single strand region and the lower double strand portion of the gel. The number of possible conformations is so large that there is no theoretical basis for choosing experimental conditions. The more experimental conditions run the greater the sensitivity.

Even when the two techniques are combined, lack of a specific band shift does not necessarily advocate the absence of a mutation in a given fragment. This is one of the major drawbacks of this method as one can not ascertain the complete absence of a mutation, theoretically or practically (Macek, 2001). To verify the location of a mutation in a fragment, direct sequencing is required. This is a general problem with most scanning techniques.

An improvement on HA has been established in the form of conformation sensitive gel electrophoresis (CSGE). In this improved version of HA, mild denaturing conditions are introduced. These conditions aid in creating changes (bends) in double-stranded DNA and thus increase differential migration patterns of homo- and heteroduplexes (Bio-Rad manual, 1996).

One of the major advantages of HA is that conditions do not have to be optimised, as conditions are constant for the majority of fragments and time for the optimal separation of different sized fragments can be predicted (Claustres, 2001).

2.3.3. Denaturing gradient electrophoresis

2.3.3.1. Introduction and theoretical background

The identification and characterisation of single nucleotide variations in DNA represents an ongoing challenge in the detection of DNA damage as well as in the genetic analysis of inherited disorders. In search for higher resolution, gels containing a gradient of increasing denaturing power can be used (Sealey and Southern, 1985). Denaturing gradient gel electrophoresis is a generic designation of a family of related mutation scanning methods that includes parallel and vertical DGGE, TTGE (temporal temperature gel electrophoresis), CDGE and TSGE (temporal sensitive gel electrophoresis).

These methods are based on the reduction of mobility in denaturing gel matrices due to gradual, co-operative, melting of distinct DNA domains within the analysed fragment as part of the double helix unravels (Lerman and Beldjord, 1998; Muyzer, 1999; Macek, 2001).

The electrophoretic mobility of DNA in polyacrylamide gels is sensitive to the secondary structure of the molecule with respect to its helicity, partial melting, or complete melting and dissociation of the strands (Lerman *et al.*, 1984). The DGGE protocol (Fodde and Losekoot, 1996) allows the identification of point mutations, which alter the melting behaviour of the DNA fragment to be analysed. In DGGE, DNA fragments of the same size are separated by their denaturation profile. DGGE is based on electrophoretic mobility of a double–stranded DNA molecule through polyacrylamide gel containing a linearly increasing concentration of denaturing agents such as formamide and urea (Muyzer, 1999). The theory behind DGGE principle shows that the two strands of DNA separate, or melt, when heat or a chemical denaturant is applied.

The temperature at which the a DNA duplex melts is influenced by two factors (King, 1999):

- The hydrogen bonds formed between complimentary base pairs, GC-rich regions melt at higher temperatures than AT-rich regions.
- The attraction between neighbouring bases of the same strand or stacking.

In practice, the denaturants used are heat (a constant temperature of 60^0C) and a fixed ratio of formamide (ranging from 0 – 40%) and urea (ranging from 0 – 7M) (Fodde and Losekoot, 1996; Helms, 1990). A fragment starts moving as a double-stranded DNA molecule but, as it migrates through the gel, it reaches a point where the denaturing agent concentration is sufficient to cause the melting of AT-rich segments.

This partial melting reduces the mobility of the fragment as it moves further into the denaturant gradient and the retardation becomes greater (Gedil, 2000). This results in the fragment becoming stationary (Muyzer, 1999), in the form of a partially double-stranded molecule containing one or more large internal loops and/or terminal single-stranded tails that become entangled in the gel matrix (King, 1999; Lerman *et al.*, 1984).

The point in the gel where a partially double-stranded molecule occurs depends on the composition and sequence organisation of the fragment (de Wachter and Friers, 1985). In the absence of a high melting domain (sequence), the increasing denaturing environment may completely denature the DNA fragment into two single strands (Helms, 1990).

DGGE is a powerful method in separating same size DNA fragments or PCR products based on sequence. The melting profile of a given DNA duplex is predominantly determined by its base sequence with greater GC content resulting in higher melting temperatures (Hepburn and Miller, 1996). The mobility of a given fragment drops sharply at the point in the gel where complementary strands begin separating.

The final fragment position is therefore sequence-dependent and, in combination with size separation, this makes two-dimensional separation of DNA fragments possible (Sealey and Southern, 1985). The physical denaturing of the DNA fragment does not proceed in a zipper-like manner, but rather in a step-wise process. As the DNA fragment enters the denaturing conditions, discrete portions of the fragment will suddenly denature into single-stranded DNA within a narrow range of denaturing conditions. This discrete portion of the DNA fragment constitutes a melting domain (Helms, 1990). A melting domain is a region within the fragment in which all of the base pairs melt at approximately the same temperature in contiguous groups of about 50 - 400bp in length (Lerman and Beldjord, 1998).

The melting temperature (T_m) is the temperature at which each base pair of a DNA duplex is in perfect equilibrium between the denatured and helical state. Since stacking interactions between adjacent bases have a significant influence on the stability of the double helix, the T_m of any given DNA molecule is largely dependent on its nucleotide sequence.

Therefore, when DNA fragments differing by a single nucleotide change in their lowest melting domain are electrophoresed through denaturing gradient gels, branching and consequent retardation of their mobility will occur at different positions along the gel allowing their separation (Fodde and Losekoot, 1996; Muyzer, 1999). The ability of DGGE to detect sequence alterations is based on the differential melting characteristics of homoduplex DNA versus heteroduplex DNA. As heteroduplex DNA migrates through the denaturant gradient the areas of non-homology melt at a lower temperature than the comparable homoduplex region. This results in an area of decreased mobility within the fragment, retarding its progress through the gel. This reduction in mobility results in a separation of homoduplex from heteroduplex fragments thereby identifying a region of sequence alteration (Hepburn and Miller, 1996). The PCR-DGGE combination is extremely powerful when applied for the detection of heterozygous nucleotide variants: continuous denaturation and reannealing of single strand molecules during PCR allows the formation of homoduplexes as well as heteroduplex molecules. The presence of a single mismatch within the latter greatly decreases the melting temperature allowing separation from the homoduplexes and an easier visual detection of the mutants.

A total of four bands will be observed in a typical DGGE analysis of a DNA sample heterozygous for a single alteration: two homoduplexes and two heteroduplexes (Bio-Rad lab manual, 1996).

PCR-DGGE is an invaluable technique applicable in research and diagnosis, especially in the analysis of inherited conditions caused by the heterozygous mutation spectra, frequent *de novo* mutations or polymorphisms (Fodde and Losekoot, 1996; Whatman, 1999). Finally, the mobility of a DNA molecule on a polyacrylamide gel determines the mutation detection resolution.

The DNA molecule mobility in a polyacrylamide gradient gel subjected to a voltage is determined by the nucleotide chain length, net charge, and conformation of the DNA on the one hand, and by the gel composition (mainly gel concentration, but also the degree of cross linking) on the other. The DGGE protocol is non-destructive to the sample (Lerman and Beldjord, 1998; Macek, 2001).

2.3.3.2. Computer simulation, predictions and primer design.

Although the theory and methodology of DGGE are relatively simple, a significant amount of preparative work must be undertaken before using the technique to screen mutations in a particular gene (King, 1999). Primers must be designed and carefully chosen so that the region to be screened for mutations has one or at most two discrete melting domains.

Full sequence data must be available so that a melting map of the molecule can be constructed, and primers can be designed to amplify a single melting domain region. The optimal gradient and running conditions must also be established. Lerman and co-workers (Hepburn and Miller, 1996) have reduced the prediction of melting domains within a DNA fragment of known sequence to a computer algorithm. The melting profiles can be performed using an adaptation of Lerman's programme created by Bio-Rad, called MacMelt™ software. This programme allows preliminary examination of the properties of the DNA fragment to be analysed.

The analysis provides (Fodde and Losekoot, 1996; King, 1999):

- Melting map.
- Optimal gel running time.
- Expected effects of virtually any base change on the T_m map.
- Base change consequences in terms of gradient displacement.

MacMelt™ is a trdemark of Bio-Rad Laboratories, Hemel Hempstead HP2 7TD, UK.

Computer simulation can save both time and money. It allows the identification of the different melting domains within the fragment and their specific T_m values. Based on this type of information, one can design PCR primers to encompass one or two melting domains within 100 – 1000bp (Bio-Rad lab manual, 1996). Thermal stability patterns of DNA fragment of choice can be altered and made more favourable by end-modification of fragments (Lerman and Beldjord, 1998).

There are, however, problems associated with the DGGE technique:

- More than two domains within one PCR product should be avoided. This is because two domains may result in a decreased sensitivity of mutation detection especially within the regions with the highest melting temperatures. More than two domains with similar melting temperatures or within a range of 5°C can still be manipulated with relatively higher sensitivity.
- Significant T_m differences between two melting domains within the same PCR fragment should also be avoided. If allowed, denaturation and branching of the domain with the lowest T_m might cause retardation of the partially melted molecule to such a degree that it will not reach the position along the gel where the denaturant concentration is high enough to denature the second T_m domain. This phenomenon is called "the dragging effect".

In circumstances wherein two or less domains are unavoidable, one can circumvent the problems by performing one of the following:

- Analysis of the same fragment on two different gradient gels.
- Digestion of the PCR product with specific restriction enzymes.

The restriction enzyme used for the large DNA molecules should be suitable for that specific fragment; that is, the GC-clamp should be left undigested.

The suspected mutation site should be located on the fragment attached to the GC-clamp because the sensitivity of mutation detection on this fragment will be ~100% whereas further else where on the fragment, sensitivity may be considerably reduced (Fodde and Losekoot, 1996). The preliminary computer simulation saves a lot of practical experimentation, however it is still, necessary to determine the optimal voltages and running times and to confirm the optimal denaturing gradient chosen (King, 1999). The aim is to obtain well-separated bands. When optimised gel running conditions have been established the method can be used for mutation screening.

2.3.3.3. GC-clamps

In normal experimental conditions (linear gradient), DGGE can only resolve a limited fraction of all possible point mutations. The majority of the DNA fragments of approximately 300 – 1000bp will encompass more than one melting domain (Bio-Rad lab manual, 1996).

The juxtaposition of GC-rich and AT-rich regions commonly seen within the human genome often result in complex melting profiles with adjacent domains demonstrating widely divergent melting temperatures as shown in figures 5.8, 5.10 and 5.12.

Such multiple melting domain fragments do not permit the resolution of sequence alterations in the higher melting temperature domains due to the more rapid melting of the lower temperature domains (Hepburn and Miller, 1996).

Base variants located within the highest melting domain will not be resolved by DGGE due to loss of sequence-dependant migration on complete strand dissociation. This problem can be overcome by introducing into the fragment to be analysed a GC-rich domain (GC-clamp) whose high T_m will prevent complete dissociation of the two strands (Lerman and Beldjord, 1998; King, 1999). The presence of the GC-clamp on either side of the DNA fragment might have profound effects on the T_m map and thus affect the percentage of detectable changes positively. This is shown in figure 5.8 through to 5.13

Since a GC-clamp as short as 40bp can efficiently serve as a high melting domain for the DGGE analysis of most DNA fragments, one of the two PCR primers used to amplify the target DNA sequence can be designed with a 5'-GC-clamp tail. This GC-sequence will incorporate at one of the ends of the resulting PCR product.

The introduction of the GC-clamp, which aid the formation of uniform low melting domains, increases the percentage of mutations detectable by DGGE to close to 100% (Fodde and Losekoot, 1996; King, 1999; Macek, 2001).

Table 2.1 gives a list of commonly used GC-clamps for eukaryotic cells (Li *et al.*, 1996)

Table 2.1: **GC-clamps for eukaryotic cells.**

Name	Size	Sequence
GC1	30mer	5'- CGC CCG CCG CGC CCC GCG CCC GTC CCG CCC -3'
GC2	38mer	5'- CGC CCC GCG CCG CCG CCC CGC CCC CGC CCG TCC CGC CC -3'
GC3	35mer	5'- CGC CCC GCC GCG CCC CGC GCG CCC GGT CCC CGC GC -3'
GC4	40mer	5'- CGC CCG CCG CGC CCC GCG CCC GTC CCG CCG CCC CCG CCC G -3'
GC5	40mer	5'- CCC CGC TCC CCG CCC CCC TCC CCG CCC CGC CCC TCG CCG C -3'

2.3.3.4. Unknown nucleotide sequence analysis.

Before performing the DGGE technique, the availability of the complete nucleotide sequence of the fragments to be analysed is desirable although not strictly essential. It is still possible to experimentally determine an unknown nucleotide sequence melting behaviour by means of perpendicular DGGE. The influence of the denaturing agent concentration on the migration distance can be examined by applying a DNA fragment mixture as a long band across a gel containing a horizontal denaturing gradient.

A uniform electrical field is also applied perpendicular to the direction of the gradient (Lerman *et al.*, 1984). This is called perpendicular DGGE (Bio-Rad Lab manual, 1996). In this case, the electrophoretic migration is perpendicular to the denaturant gradient and the sample is applied along the entire width of the gel (Gedil, 2000). In this way, each DNA molecule will travel along a path of constant denaturation concentration at a constant electrophoretic rate.

The resulting curve will provide an indication of (Fodde and Losekoot, 1996):

- The number of melting domains
- The percentage of the denaturant agents at which denaturation of each domain occurs

Perpendicular DGGE produces a characteristic sigmoidal curve. This characteristic shape is a result of two factors. Firstly, at one end of the gel wherein the denaturant concentration is very low, the DNA fragments are separated according to their length, as is the case in conventional gel electrophoresis. Secondly, the apparent mobility of the fragments is much smaller at the edge of the gel with very high denaturant concentration due to the partially denatured fragments (Lerman *et al.*, 1984). The details of the sigmoid curve differ reproducibly depending on the length and sequence of the DNA fragment.

By comparing the melting behaviour of the polymorphic DNA fragments side-by-side on denaturing gradient gels, it is possible to detect fragments that have mutations in the first melting domain. The actual detection is achieved by analysing the S-shaped curve, the characteristic denaturant concentration at which the first domain melts can be determined (Helms, 1990).

Each fragment will show a sharp change in mobility at a characteristic denaturant concentration. The changes in mobility shift to a higher denaturant concentration when the gel is run at a lower temperature. The changes in mobility across the gradient correspond to the extent of loss of helical structure attributable to the denaturing equilibrium (Lerman *et al.*, 1984).

Disadvantages associated with the perpendicular DGGE method may lead to complications that subsequently reduce its sensitivity:

- DNA strands may dissociate completely. This means that the basic principle of DGGE is not applicable.
- The melting temperature differences between domains may be over 8°C. The higher temperature range difference leads to the dragging effect which will cause mobility retardation such that the high melting domain is never exposed to the relevant higher denaturing gel concentrations.
- More than two domains may be present. Three or more peaks are difficult to resolve and may consequently complicate further analysis.
- The site of GC- clamp attachment may have unfavourable effect.

Information on the relative position of each domain within a fragment is not provided by perpendicular DGGE. Hence, the effect of the presence of a GC-clamp on either side of the molecules cannot be visualised. The combined effect of the disadvantages listed above results in a significant reduction of sensitivity.

2.3.3.5. Known nucleotide sequence analysis.

Testing for the presence or absence of a known mutation is relatively simple compared to scanning an unknown gene sequence for presence of unknown mutations. In addition to DGGE, techniques that can be used for such experimentation include restriction enzyme digestion, allele-specific hybridisation and ligation or non-ligation of adjacent probes (Claustres, 2001).

Only DGGE based method will be discussed in this section. In DGGE, mutations are analysed after the target sequence has been amplified by PCR. Parallel DGGE implies conventional vertical electrophoresis where DNA molecules migrate through linearly increasing concentrations of denaturant agents, commonly urea and formamide (Gedil, 2000).

The length of the molecule, as in conventional electrophoresis largely determines initial velocity, but when the molecule reaches a depth in the gradient corresponding to the abrupt decline in mobility seen in perpendicular gradients, there is little further travel (Lerman *et al.*, 1984). Denaturant gradient gels are made with polyacrylamide and poured with conventional gradient mixers. In order to allow reproducibility of the electrophoretic runs, DGGE is generally performed at a constant temperature of 60°C. This temperature was empirically chosen to exceed the T_m of an AT-rich DNA fragment in the absence of denaturants. For GC-rich sequences, temperatures of up to 75°C can be employed (Fodde and Losekoot, 1996). Once the necessary computer simulations have been performed, primers can be designed to amplify the target sequence and the first denaturing gradient gel can be run.

The choice of the denaturant range is made based on the T_m of the domain to be analysed: the parallel gel should be initially chosen within a 25 – 30 % difference in denaturant concentration centred around the T_m of the domain. The conversion factor between T_m and percentage denaturant for gels run at 60°C is given by the empirical formula:

$$\textbf{\% denaturant} = (3.2 \times T_m) - 182.4$$

For example, if the DNA fragment has a single domain with $T_m = 75^0$C, it should be ideally analysed on a gradient between 43 and 73% denaturant. Narrower denaturant gradients (10 – 15% difference) can also be employed with satisfactory results. If two distinct domains are present within the same fragment, two different gradients should be made for their optimal mutation analysis (Fodde and Losekoot, 1996). Sequence differences can be easily detected in DNA fragments when nearly identical fragments are electrophoresed in the same direction as that of the denaturing gradient (Helms, 1990). These parallel gels permit the simultaneous comparison of as many sets of fragments as there are lanes on the gel, unlike the perpendicular gels (Helms, 1990; Gedil, 2000).

2.3.3.6. Summary

DGGE compares favourably to other gene scanning techniques with respect to detection rate, robustness and size of analysed fragment. The PCR-DGGE combination approach has a major limitation in that unpredictable electrophoresis behaviour of the mutant strands cannot be pre-determined; hence different conditions have to be optimised to achieve maximum sensitivity. In spite of this draw back, this approach is one of the most feasible methods as it reduces labour intensity and the cost for sequencing negative samples. An important advantage of DGGE is the possibility to design primers and gel conditions to maximise the detection of alterations in advance. However, as recent as 1996, it was generally accepted that mutations less than 50bp from the GC-clamped primer are less detectable than those located at a greater distance (Hepburn and Miller, 1996).

The relative simplicity and discrimination of the DGGE technique, together with its capability for resolving complex mixtures of molecules, generates a wealth of experimental data. Detailed comparisons among large sets of closely related molecules can be made in a single experiment, for example, in parallel DGGE. Since the physical separations achieved in DGGE are determined by the sequence of the DNA fragment rather than its length, and the sensitivity can be adjusted over a broad span, various otherwise difficult or inaccessible mutation detection becomes relatively easy and straightforward (Lerman *et al.*, 1984).

DGGE can be used to detect ~100% (Bio-Rad lab manual, 1996) of all possible single base changes in DNA fragments up to ~1000bp. Furthermore, DGGE can be used to detect polymorphisms not identifiable by techniques such as SSCP, RFLP (Gedil, 2000) and heteroduplex analysis. In addition to the detection of missense, nonsense, deletion and insertion mutations, one of the major advantages of using the Bio-Rad mutation detection system is that synonymous mutations (Fraser and Mayo, 1975) can also be successfully detected. Thus, DGGE seems an ideal choice for mutation detection in the clinical diagnostic laboratory.

In general, some of the shortcomings of DGGE analysis include:

- Laborious and time-consuming preliminary work prior to actual analysis of fragment.
- Costly synthesis of relatively long (~60 nucleotides) PCR primers.
- Limited size of the largest DNA fragment (~1000bp) that can be efficiently analysed.
- Genes that are exceptionally GC-rich are not easily analysed due to difficulty of converting the different melting domains to a single melting region.
- Method involves use of formamide, a known carcinogen.

On the other hand, the advantages associated with DGGE analysis can far outweigh the limitations stated. The advantages include:

- High sensitivity (King, 1999) and detection rate (Fodde and Losekoot, 1996).
- Improved heterozygote detection (Lerman and Beldjord, 1998).
- Optimisation of analysis through computer simulation (Bio-Rad manual, 1996).
- Non-radioactive and non-destructive protocol (King, 1999; Lerman and Beldjord, 1998).
- Easy isolation of the mutant allele for its subsequent sequence determination (King, 1999).
- Combination of a high sensitivity of mutation detection with a relatively less labour-intensive protocol (Lerman *et al.,* 1984).

The DGGE protocol mutation detection success has been proven to be equal to or more effective than RNA protection and hydroxylamine and osmium tetroxide (HOT), chemical cleavage and/or single strand conformation polymorphism (SSCP) (Fodde and Losekoot, 1996).