CHAPTER 1: INTRODUCTION

1.1 The Biology of the Erythrocyte

Red blood cells or erythrocytes are highly specialised for their function as transporters of oxygen and carbon dioxide to all cells of the body. They have a high surface area to volume ratio suitable for gaseous exchange and the loss of even four percent of their surface area results in cell lysis. The erythrocytes, with diameters of approximately eight micrometers, must distort to pass through narrow capillaries of no more than four micrometers in diameter. This increases the surface area available for gaseous exchange, but the geometry of the elongated erythrocytes is unstable and the more stable biconcave disc shape must be resumed as the capillaries widen. Erythrocytes contain a large quantity of haemoglobin, the oxygen-carrying molecule (Mohandas & Chasis, 1993).

All blood cells derive from pluripotent haemopoietic stem cells which reside in the bone marrow and become specialised through haemopoiesis (blood cell development). The process whereby the progenitor cells are stimulated by erythropoietin to commit to erythroblast formation and ultimately to differentiation into mature erythrocytes is known as erythropoiesis. Immature erythrocytes (reticulocytes) are released from the bone marrow into the blood where their maturation, mediated by cytoplasmic adenosine-5'-triphosphate (ATP)-dependent enzymes, occurs over a period of 24 to 48 hours. Reticulocytes are anucleated but contain a number of polyribosomes with associated ribosomal ribonucleic acids (RNA) that are eradicated as the cell matures (Junqueira *et al*, 1995). In a non-disease state the reticulocytes comprise up to two percent of the circulating erythrocytes (Hoffbrand & Pettit, 1993). This is similar to the number of aged erythrocytes removed from the blood daily (Junqueira *et al*, 1995).

The mature erythrocyte is an anucleated membranous sack filled with a 33 percent haemoglobin solution and enzymes for the Embden-Meyerhof (glycolytic) and hexose-monophosphate-shunt pathways. Both of these pathways are responsible

for the conversion of glucose (the erythrocyte's sole source of energy) to ATP which provides the energy required to preserve the erythrocyte's flexibility, shape and volume (Hoffbrand & Pettit, 1993).

Under normal conditions erythrocytes remain in the blood for approximately 120 days, becoming more rigid over time. This occurs when unsupported portions of the cell membrane are removed (splenic conditioning) during passage through the spleen. The increased rigidity of the erythrocytes stimulates the removal of erythrocytes from the circulation as they become trapped in the narrow sinuses of the spleen, the primary site for erythrocyte destruction. The erythrocytes also begin to display defective oligosaccharide and protein complexes on the membrane surface which signal phagocytosis of erythrocytes by macrophages within the spleen. Within the phagocyte, the cell membrane is recycled while the haemoglobin is transported to the liver where it is broken down (Walensky *et al*, 2003).

1.2 The Molecular Structure of the Erythrocyte Membrane

The erythrocyte membrane comprises the erythrocyte membrane skeleton and the phospholipid bilayer (figure 1.1). The highly flexible membrane skeleton supports the phospholipid bilayer through tight binding to integral membrane proteins found within the bilayer and maintains the increased surface area to volume ratio vital for distortion under pressure and gaseous exchange.

The phospholipid bilayer consists of choline- and aminophospholipids interspersed with cholesterol, glycolipids, proteins and glycoproteins (Palek & Jarolim, 1995). The cholinephospholipids (sphingomyelin and phosphatidylcholine) predominate on the outer membrane surface while the aminophospholipids (phosphatidylserine and phosphatidylethanolamine) are present on the inner membrane surface. This asymmetry, particularly the internalisation of highly negative phosphatidylserine molecules, is believed to be

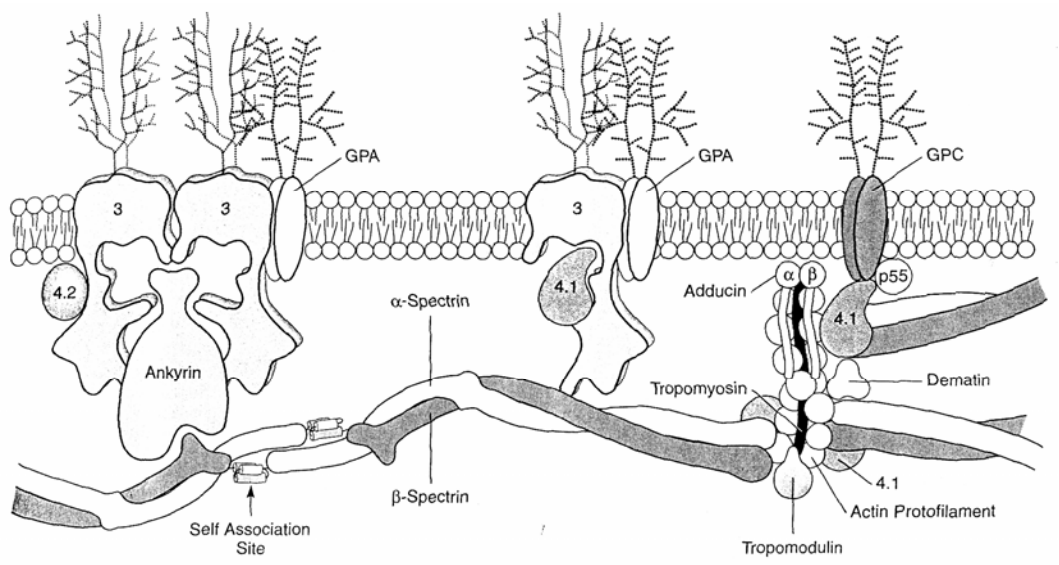

Figure 1.1: Schematic Model of the Erythrocyte Membrane

Band 3 and glycophorin proteins transverse the lipid bilayer of the erythrocyte membrane. These are attached to the spectrin-based membrane skeleton by integral membrane proteins including protein 4.1 and ankyrin. GPA=Glycophorin A, GPC=Glycophorin C, 3=Band 3, 4.1=protein 4.1 (Walensky *et al*, 2003)

important to the function of the erythrocyte (Palek& Jarolim, 1995; Hassoun *et al*, 1998). The phospholipids move constantly between the inner and outer membrane in a characteristic 'flip-flop' fashion controlled by the activities of three enzymes: aminophospholipid translocase, floppase and scramblase. The aminophospholipid translocase is ATP-dependent and actively transports phosphatidylserine and phosphatidylethanol-amine to the inner membrane surface ('flip'). Floppase is also ATP dependent and actively transports sphingomyelin and phosphatidylcholine to the outer membrane surface ('flop') while a cytoplasmic calcium increase causes the scramblase enzyme to rapidly disrupt the phospholipid distribution (Hassoun *et al*, 1998).

Approximately 40 percent of the membrane proteins traverse the membrane (transmembrane proteins), forming transporters and channels for a variety of molecules. These proteins allow the erythrocyte to maintain an internal ionic balance so that it does not burst or shrivel up and also permit gaseous exchange (Walensky *et al*, 2003). The other 60 percent of the cell membrane proteins are attached to the cytoplasmic surface of the membrane and are known as peripheral proteins. The membrane skeleton is composed of peripheral proteins that are connected to the cell membrane via their interactions with various transmembrane proteins and, to a lesser extent, via direct interaction with the lipids of the membrane (Walensky *et al*, 2003). The proteins of the erythrocyte membrane can be solubilised in sodium dodecylsulphate (SDS) and separated by polyacrylamide gel electrophoresis (PAGE). They were originally named by Fairbanks *et al* (1971) according to the distance moved on a polyacrylamide gel, an example of which is shown in figure 1.2.

1.3 The Erythrocyte Membrane Skeleton

The membrane skeleton of the erythrocyte is easily extracted and has been extensively studied, both to understand its importance in the erythrocyte and as a model of other cell skeletons. The interactions of the erythrocyte membrane proteins can be viewed in two dimensions: vertical interactions and horizontal

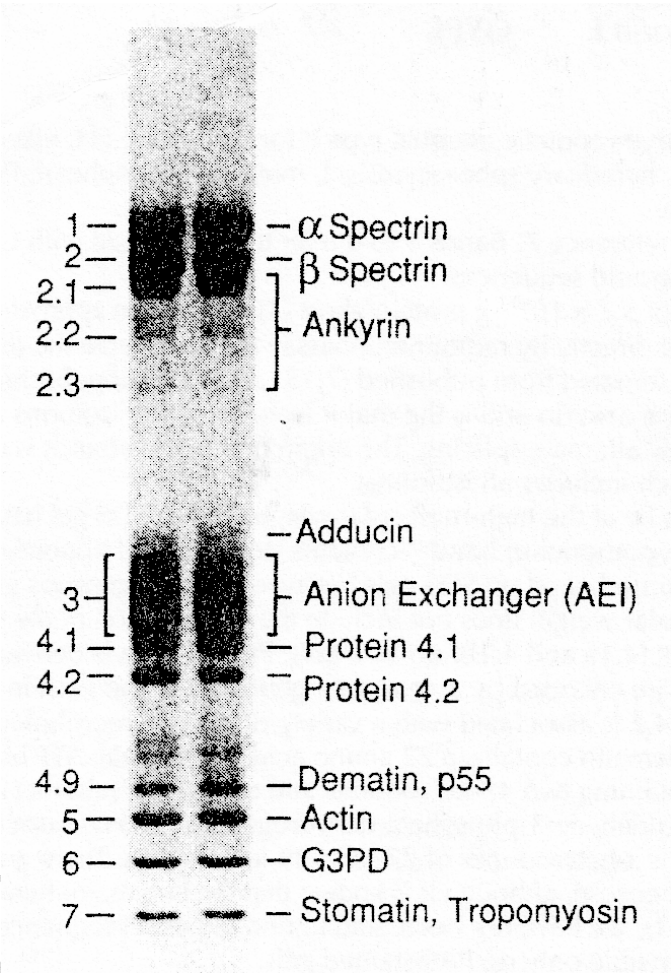

Figure 1.2: Separation of Erythrocyte Membrane Proteins using Fairbanks Polyacrylamide Gel Electrophoresis (Walensky *et al*, 2003)

Using polyacrylamide gel electrophoresis, the proteins of the erythrocyte membrane can be separated based on their molecular weight. The proteins are visualised using Coomassie Blue dye.

interactions. Vertical interactions are those between the membrane skeleton and the integral membrane proteins. Horizontal interactions involve only the membrane skeleton (Walensky *et al*, 2003).

1.3.1 Horizontal Interactions

The membrane skeleton is based on a network of spectrin tetramers bound at their tail ends by actin molecules (Walensky *et al*, 2003). Two spectrin isomers are found in the erythrocyte membrane skeleton: α-spectrin, a 280 kiloDalton (kD) molecule (band 1) encoded on chromosome one, and β-spectrin, a 246 kD molecule (band 2) encoded on chromosome 14. α-spectrin contains 20 complete spectrin repeats and has a calcium-binding site at its carboxy terminal. β-spectrin is composed of 16 repeats, has an ankyrin-binding site within modified repeats fourteen and fifteen and binds both actin and protein 4.1 at its amino-terminal. The 106 amino acid spectrin repeats are folded into triple α-helices maintained through hydrophobic interactions. Spectrin comprises 50 to 75 percent of the erythrocyte membrane skeleton, primarily in the form of approximately 100 000 heterotetramers. The flexible tetramers are composed of two self-associated α-β-spectrin heterodimers. Each tetramer is 194 nm in length but can condense to 76 nm through reversible coiling, an attribute linked to the ability of the erythrocyte membrane skeleton to distort within the microvasculature (Cohen, 1983; Gallagher & Forget, 1993).

Spectrin binds to actin (band 5), the form of which expressed in the erythrocytes is β-actin (Cohen, 1983; Gilligan & Bennett, 1993) which is associated with the formation of short actin filaments. Approximately 30 000 short actin filaments of 30 to 40nm, each containing 12 to 14 actin monomers, are found in each erythrocyte (Gallagher & Forget, 1993). The oligomerisation of actin is stabilised by the binding of two tropomyosin molecules to each chain of actin monomers. Erythrocyte tropomyosin consists of two α-helical polypeptides of 27 kD and 29 kD. These are proposed to associate to form rod-like dimers of 33 to 35 nm, which stabilise and influence the length of the actin filaments and modulate the interaction of spectrin and actin. Tropomodulin, a 41 kD polypeptide, binds to

both tropomyosin and actin. Its amino terminal has affinity for tropomyosin and it is thought that the binding of tropomodulin to both tropomyosin and actin acts to limit the length of the F-actin polymers. It has been postulated that there is a membrane binding site in tropomodulin, given that it remains bound to the membrane even after the removal of spectrin, actin and tropomyosin, but the exact region and the importance of this site have yet to be determined (Walensky *et al*, 2003).

The interaction of spectrin and actin is promoted by adducin, a heterotetramer consisting of two 81 kD α- and two 80 kD β-subunits. The globular, 39 kD, amino head region is protease-resistant and self-associates with other adducin molecules (Gilligan & Bennett, 1993; Walensky *et al*, 2003). It is linked by a 9 kD neck region to a 30 kD, protease-sensitive, hydrophilic tail which interacts with actin to form a spectrin-actin-adducin complex. Actin-adducin interaction is decreased by controlled binding of calmodulin. The stabilisation of the spectrin-actin bond by adducin is transient (Walensky *et al*, 2003) and is ultimately stabilised by proteins 4.1 and 4.9, resulting in the formation of the junctional complex (Cohen, 1983; Gilligan & Bennett, 1993). Protein 4.1 is an 80 kD protein (discussed in greater detail in section 1.5) while protein 4.9 refers to two co-migrating proteins: p55 and dematin (Gilligan & Bennett, 1993). p55 is a palmitoylated 55 kD protein which binds tightly to the lipid bilayer, a result of its palmitoylation, and to protein 4.1 and glycophorin C. It is a membrane-associated guanylate kinase protein homologue and is thus likely to display functional characteristics such as ion channel clustering and signal transduction. Dematin isomers of 48 kD and 52 kD, the latter containing an ATP-binding site, exist in the erythrocyte. The native form of dematin is a trimer (one 52 kD and two 48 kD proteins) that is believed to contribute to membrane stability as adenosine-3', 5'-cyclic monophosphate (cAMP) dependant phosphorylation of the protein inhibits its promotion of actin bundling (Walensky *et al*, 2003).

1.3.2 Vertical Interactions

In addition to the formation of the membrane skeleton, it is important that the membrane skeleton is bound to the lipid bilayer of the membrane. This is achieved through the interactions of the membrane skeleton and the transmembrane proteins. The primary interaction is that of β-spectrin, ankyrin and erythrocyte band 3. Band 3 is a 102 kD multi-pass transmembrane protein (discussed in greater detail in section 1.4) while ankyrin exists as a number of subtypes (forming bands 2.1, 2.2, 2.3 and 2.6 on a SDS-PAGE gel). The most common erythroid form is band 2.1 which is a 206 kD globular protein comprising three domains: an 89 kD amino band 3-binding domain, a central 62 kD spectrin-binding domain and a 55 kD carboxy regulatory domain. The band 3-binding domain is made up of four regions, each of which consists of six 33 amino acid repeats. These crystallise into two antiparallel α-helices preceded by a β-hairpin loop, a structure believed to be important to the binding of band 3. The spectrin-binding domain has a highly conserved core, but it is the more mutable end regions of this domain with which spectrin interacts. The various ankyrin spliceoforms arise from the regulatory domain which displays a high degree of variability. Ankyrin attaches to β-spectrin close to the centre of the spectrin tetramers and then associates with band 3 (Peters & Lux, 1993).

Associated with band 3 is protein 4.2, an 80 kD peripheral membrane protein that binds to band 3 and to the lipid bilayer. The interaction of protein 4.2 and band 3, improved by palmitoylation of protein 4.2 at cysteine 203, is particularly important as it stabilises the proteins of the membrane skeleton: band 3 binds more strongly to the membrane skeleton and shows decreased rotational mobility in the presence of protein 4.2. Total absence of protein 4.2 (either a result of ineffective band 3 binding or the lack of production of protein 4.2) results in extreme membrane fragility (Inoue *et al*, 1998).

Another major interaction occurs at the distal ends of the spectrin tetramers, involving the attachment of spectrin to protein 4.1 and via this to the glycoproteins (notably glycophorins C and D) of the membrane (section 1.5.3)

(Hemming *et al*, 1995). The glycophorins are single-pass, α-helical, transmembrane sialoglycoproteins. The sialic-acid containing carbohydrate group is attached to the extracellular amino terminal of the protein and gives the extracellular surface of the erythrocyte a strong negative charge. Five glycoproteins have been identified and carry a number of erythrocyte antigens. Glycophorins A, B and E show significant homology and are thought to have derived from one original gene. Glycophorin A is the most common sialoglycoprotein of the erythrocyte membrane and is limited to the erythrocyte and the kidney (Chasis & Mohandas, 1992). It has a molecular weight of 36 kD and is thought to enable the translocation of band 3 to the cell membrane (Bruce et al, 1994; Groves & Tanner, 1992). Glycophorins C and D are encoded by a different gene and are distinguished from one another by distinct initiation codons (Chasis & Mohandas, 1992). Both contain homologous protein 4.1 and p55 binding sites (Hemming *et al*, 1995). Glycophorin C, a 32 kD protein with 13 oligosaccharide groups attached to its amino terminal, is the major protein 4.1-binding glycophorin.

1.4 Band 3

1.4.1 The Band 3 Gene

The 20 kilobase (kb) gene encoding band 3 is located at 17q21-qter and comprises twenty exons (Showe *et al*, 1987). The cDNA was cloned from a foetal liver cDNA library and possesses 2736 bases that encode 911 amino acids and a termination codon. The initiation consensus sequence (CCACGCCAUGG) is similar to the accepted eukaryotic initiation consensus sequence: CC(A/G)CC**AUG**G (Tanner *et al*, 1988). A promoter region is located within intron one and the erythroid initiation codon is found in exon 2. The kidney isoform is transcribed from another initiation codon with its own promoter region, both being found in exon 5. A common termination codon is located in exon 20 (Sahr *et* al, 1994).

The band 3 gene has a high percentage of primate *Alu* repeats, nucleotide sequences of approximately 300 base pairs followed by poly-adenine tracts. Sahr *et al* (1994) hypothesised that these *Alu* repeats indicate an increased level of AE1 gene rearrangement which may be linked to disorders such as HS. Except for these repeats, the human band 3 gene shares 80 percent exon homology and 40 percent intron homology with the murine band 3 gene (Sahr *et al*, 1994).

1.4.2 The Band 3 Protein

Band 3 is an anion exchanger protein expressed in the erythrocyte and is alternately termed anion exchanger protein (AE) 1. Band 3 is the major integral membrane protein, comprising approximately 25 percent of the total membrane protein, with up to $1.2x10^6$ copies of band 3 being expressed in the erythrocyte membrane. A shortened isoform of the protein is found in the distal nephron of the kidney (Tanner, 1997). The anion exchangers (AE1, AE2 and AE3) possess similar membrane domains with 70 percent homology, the size and amino acid sequence of the cytoplasmic domains distinguishing the diverse proteins. AE2 is expressed in a wide variety of tissues while AE3 is localised to the brain, the retina and the heart (Reithmeier *et al*, 1996).

Erythrocyte band 3 has a size of approximately 100 kD, forming a diffuse band on a polyacrylamide gel. The apparent variation in size is caused by different numbers of N-acetyllactosamine repeats in the carbohydrate chain attached to asparagine 642 (Jay, 1986). It has two independently functioning domains (figure 1.3) which can be separated by mild proteolysis. The first is the 52 kD carboxy-terminal transmembrane domain which constitutes the pore through which anions can move. The second is the 43 kD amino-terminal cytoplasmic domain that is responsible for binding to the erythrocyte membrane skeleton as well as to haemoglobin and some glycolytic enzymes (Tanner, 1997). The amino acid sequence of band 3 was deduced by Tanner *et al* (1988) from the cDNA sequence.

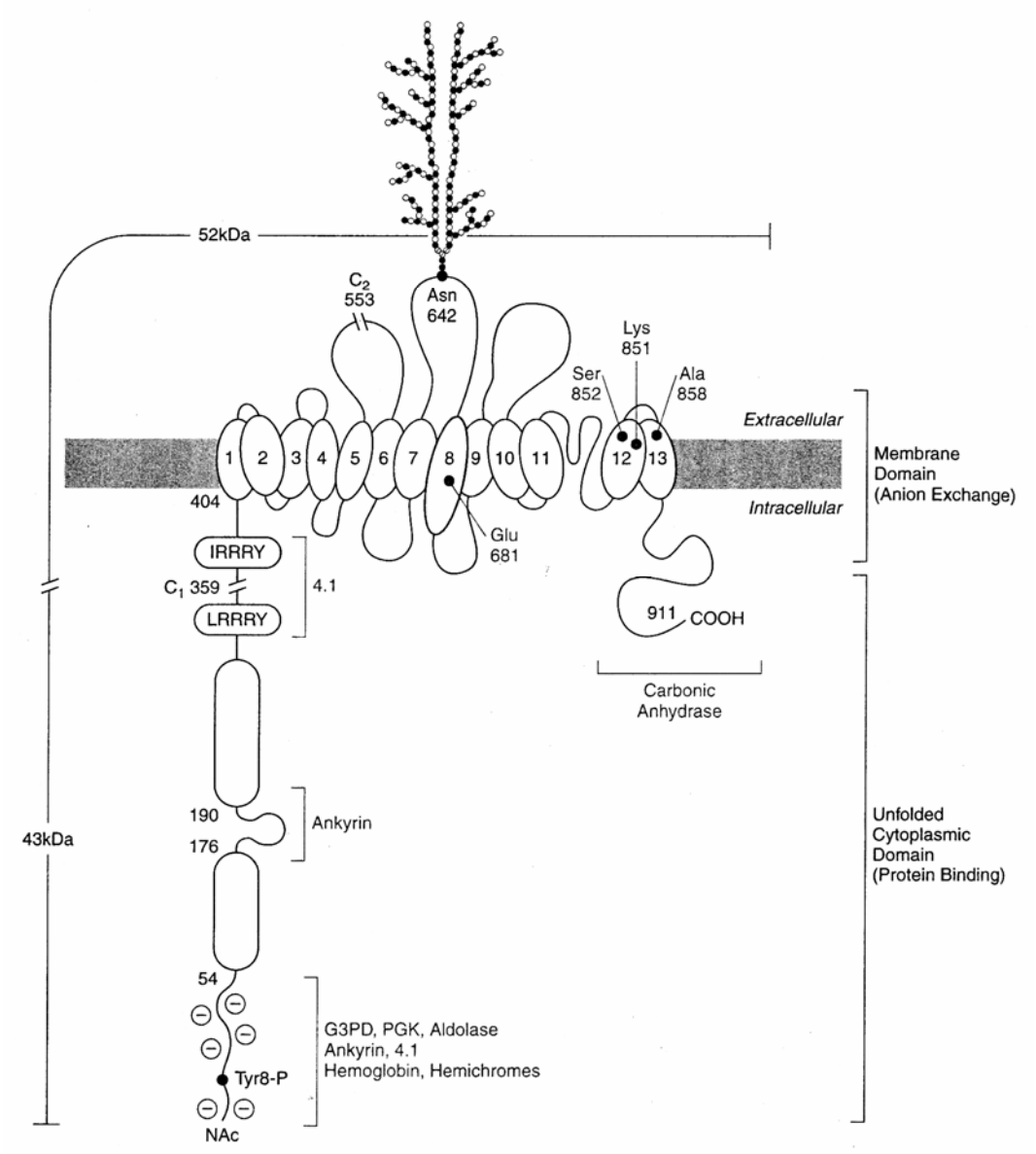

Figure 1.3 Organisational Model of Human Erythrocyte Band 3

(Walensky *et al*, 2003)

The 43kD (cytoplasmic) and 52kD (transmembrane) domains are indicated. The cytoplasmic domain binds to the membrane skeleton as well as to a variety of enzymes. The transmembrane domain forms a pore through which anions are transported and also carries a number of erythrocyte antigens.

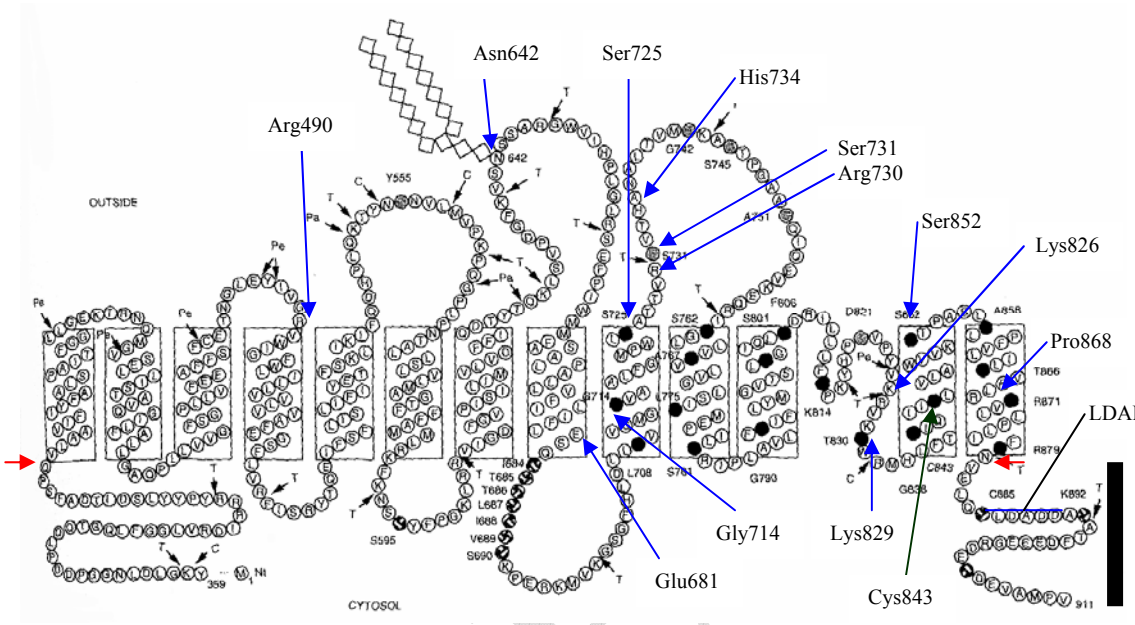

Figure 1.4: Topographical Model of the Membrane Portion of Erythrocyte Band 3

The membrane (anion exchange) portion of band 3 extends from glutamine 404 to asparagine 880 (red arrows). The terminal amino acids of the protein (881 to 911) are believed to be important for the correct insertion of the protein (black line). Arginine 490, glutamic acid 681, glycine 714, serine 725, arginine 730, serine 731, histidine 734, lysine 826, lysine 829, the amino acids between serine 852 and proline 868 and glycosylated asparagine 642 (blue arrows) are important for correct anion exchange function. L^{886}DADD (Blue line) binds carbonic anhydrase II. Palmitoylated cysteine 843 is also indicated (green arrow). The black, grey and hatched circles represent introduced cysteine mutants as determined by sulphydryl-specific chemistry. (Fujinaga *et al*, 1999)

1.4.3 The Band 3 Transmembrane Anion Exchange Domain

The carboxy or anion exchange domain of band 3 (figure 1.4) consists of 552 amino acids extending from glutamine 404 to valine 911, the last amino acid of the membrane domain being asparagine 880 (Fujinaga *et al*, 1999). Amino acids 881 to 911 re-enter the cytoplasm and are believed to be highly important for the correct insertion of band 3 into the erythrocyte membrane (Tanner, 1997). Mutations in the transmembrane region can affect insertion of band 3 into the phospholipid bilayer.

The current model of the membrane domain topology was established using sulphydryl-specific chemistry and describes an anion exchange channel consisting of 13 α-helical, transmembrane segments. A relaxed transmembrane stretch of amino acids is found between helices 11 and 12. The hydrophobic transmembrane regions are linked by hydrophilic loops of varying sizes (Fujinaga *et al*, 1999). Some extracellular loops are modified by glycosylation or contain polymorphisms which give rise to various blood group antigens, notably the I/i and Diego blood groups (Poole, 2000). Two notable post-translational modifications are the N-glycosylation of asparagine 642 and palmitoylation of cysteine 843. The former is important for the conformation of band 3 for optimal anion transport while the purpose of the latter modification is unknown (Tanner, 1997).

The function of the anion exchange domain is to actively transport anions into the cell (figure 1.5), allowing others to flow out of the cell (Tanner, 1997). Glycine 714, serines 725 and 731 and amino acids 852 to 868 have all been identified as having an important role in the regulation of anion transport (Fujinaga *et al*, 1999). Other studies have indicated roles for lysines 826 and 829 and for arginines 490 and 730 (Shayakul & Alper, 2004). Notably, codon 490 has been identified as a mutation hotspot. The primary function of band 3 is to pump chloride into the cell, causing hydrogen carbonate to stream into the plasma so that carbon dioxide is transported away from the tissues by both the erythrocyte and the plasma (Tanner, 1997). Although not part of the transmembrane domain, cytoplasmic L^{886}DADD in the carboxy-terminal binds to carbonic anhydrase II, bringing it into

the proximity of the anion exchange channel and appears to accelerate Cl^-/HCO_3^- exchange (Shayakul & Alper, 2004).

The anion exchange channels operate by a 'ping-pong' mechanism, exchange transport being necessary for conversion between the inward (cytoplasm) and outward (plasma) conformations. Intracellular hydrogen carbonate enters the channel and is translocated into the plasma. The anion channel cannot resume its inward conformation whilst empty and hence chloride anions must bind to the protein and be translocated into the cell (reverse cycle). Extracellular hydrogen carbonate is similarly internalised as chloride is pumped out of the erythrocyte so that carbon dioxide can be exchanged for oxygen at the lungs (Fröhlich & Gunn, 1986). Glutamic acid 681 (transmembrane domain 8) is thought to be bind protons to allow anion transport in an acidic environment. It loses the proton in an alkaline environment which stimulates the cycle where chloride anions are transported into the cell (Shayakul & Alper, 2004). The binding and translocation of hydrogen carbonate is the rate-limiting step of cycle, being slower than that of chloride (Fröhlich & Gunn, 1986). This could be caused by hydrogen bond formation between the aforementioned glutamic acid 681 and histidine 734 (Shayakul & Alper, 2004).

The multitudinous band 3 molecules of the erythrocyte membrane are responsible for the transport of 10^{10} to 10^{11} chloride and hydrogen carbonate anions per second (Tanner, 1997). It has been noted in several papers that dimeric band 3 transports anions with greater efficiency than monomeric band 3, indicating that this dimerisation is important for optimal anion exchange (Walensky *et al*, 2003).

1.4.4 The Band 3 Cytoplasmic Domain and the Membrane Skeleton

The highly acidic stretch of amino acids one to 45 of the cytoplasmic domain binds glycolytic enzymes and is responsible for the regulation of erythrocyte glycolysis (Walensky *et al*, 2003). The binding of the enzymes is regulated by phosphorylation of tyrosine eight (tyr8) of band 3 (highlighted in figure 1.3) (Low *et al*, 1987). The cytoplasmic domain also binds weakly to haemoglobin. When

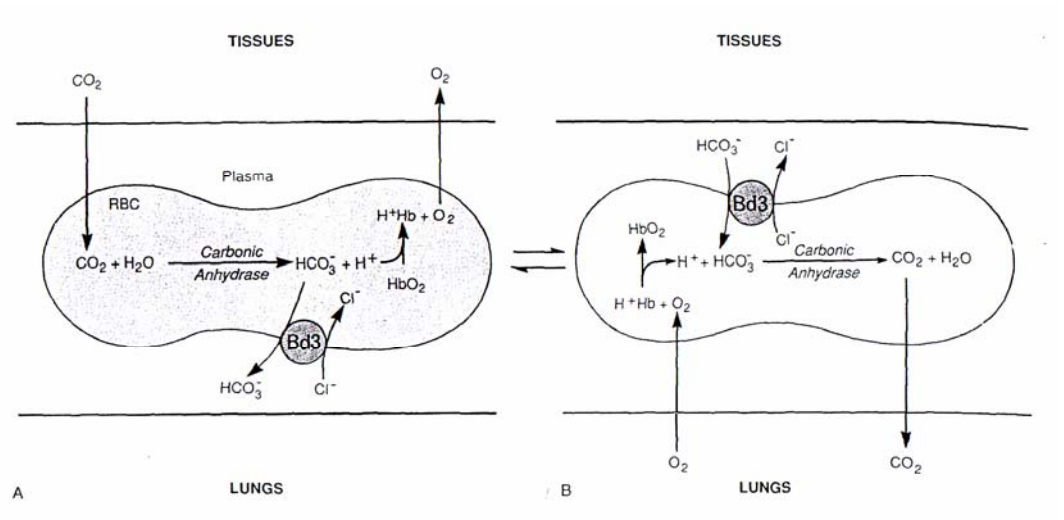

Figure 1.5: Anion Exchange in the Erythrocyte

(A) Oxygen (O_2) enters the tissue from the erythrocyte and carbon dioxide (CO_2) enters the erythrocyte from the tissues. (B) Oxygen enters the erythrocyte from the lungs and carbon dioxide enters the lungs from the erythrocyte. CO_2 is converted to hydrogen carbonate (HCO_3^-) in the erythrocyte. Excess HCO_3^- is carried in the plasma and must revert to CO_2 (in the erythrocyte) before entering the lungs. The transport of HCO_3^- across the membrane is performed by band 3 in exchange for chloride ions (Cl^-). H^+: hydrogen ion (Walensky *et al*, 2003)

haemoglobin is partially denatured, seen in aged and abnormal erythrocytes, it forms hemichromes which bind with greater affinity to band 3. This can cause clustering of band 3 and hence may be linked to the formation of the senescence antigen (Low, 1986).

The cytoplasmic domain of band 3 also binds to the membrane skeleton via ankyrin-1 (section 1.4.5) and protein 4.1. Two protein 4.1 binding sites have been postulated: two regions (L/IRRRY) near to the transmembrane domain and one between amino acids 1 and 91. The protein 4.1 band 3 binding sites are discussed in section 1.5.3.

The band 3 binding site of protein 4.2 (64RRGQPFTIILYF) has been well described. It is found near the beginning of protein 4.2 and a second site is present between amino acids 187 and 211 (Rybicki *et al*, 1995). The protein 4.2 binding sites are not well known but are located within the cytoplasmic domain of band 3. While no precise binding site has been identified, glutamic acid 40 (Rybicki *et al*, 1993), glycine 130 (Inoue *et al*, 1998) and proline 327 (Jarolim *et al*, 1992) have all been implicated as amino acids crucial to this binding, mutations of any of these three amino acids causing decreased protein 4.2 even when band 3 levels are normal or near-normal.

1.4.5 Biogenesis and Processing of Band 3

During the biogenesis of the erythrocyte the levels of the erythrocyte membrane skeletal proteins undergo rapid turnover. In both mammalian and avian erythrocytes the synthesis of band 3 is slow before cell differentiation while spectrin, ankyrin and protein 4.1 are initially produced and degraded at high levels. As the erythrocyte begins to mature band 3 is more rapidly synthesised and lower levels of the other proteins are formed. The degree of degradation is similarly decreased, suggesting that the assembly of band 3 signals the incorporation of spectrin, ankyrin and protein 4.1 into the membrane skeleton (Hanspal & Palek, 1992). Three rate-limiting steps have been defined for the assembly of the membrane skeleton. First is the availability of band 3 which

provides high affinity sites for ankyrin. Second is the availability of ankyrin, particularly as its synthesis decreases during cell differentiation. The third is the rate of synthesis of β-spectrin which attaches to ankyrin after binding to α-spectrin which is produced in excess (Hanspal *et al*, 1998).

Gomez and Morgans (1993) suggested that the erythrocyte membrane proteins (both those of the membrane skeleton and the integral proteins to which they are bound) are assembled as small segments which combine at the phospholipid bilayer. The binding of band 3 to ankyrin could be the most important signal for the formation of the erythrocyte membrane, band 3 being the most abundant erythrocyte membrane-bound protein. Ankyrin seems to aid the movement of band 3 through the Golgi bodies as band 3 processing within the ER is incomplete when ankyrin is lacking, although the mechanism through which this occurs is unclear as band 3 is produced on endoplasmic reticulum-bound ribosomes while ankyrin is produced on free ribosomes (Gomez & Morgans, 1993). The association of band 3 and ankyrin could provide both proteins with protection from degradation within the cytoplasm during their movement to the cell membrane (Hanspal *et al*, 1998).

The alternative hypothesis of Bruce *et al* (1994) suggests that the movement of band 3 to the cell surface is mediated by glycophorin A. Studies on M^kM^k erythrocytes, which are deficient in glycophorins A and B, showed a diminution of sulphate transport as band 3 affinity for anions was reduced. This was attributed to a significant lengthening of the band 3 asparagine 642 glycan chain in the glycophorin A-deficient cells as the number of membrane-bound band 3 molecules did not differ between normal and M^kM^k erythrocytes. It was speculated that glycophorin A associates with band 3 to export the latter from the endoplasmic reticulum and that a diminished quantity of glycophorin A causes a slower rate of band 3 transport. This would result in excess band 3 post-translational modification, hence the lengthened amino terminal glycan chain (Bruce *et al*, 1994). The proposed importance of glycophorin A in the removal of band 3 from the endoplasmic reticulum is supported by the findings of Groves and

Tanner (1992) which indicated that glycophorin A and band 3 expression is linked in *Xenopus* oocytes. In these oocytes the association of glycophorin A with band 3 increased the translocation of band 3 from the endoplasmic reticulum to the cell surface when the concentration of band 3 RNA was low. This association produced no significant change in band 3 transport to the cell membrane when the band 3 RNA concentration was increased. Furthermore, over extended periods of time, band 3 could increase to levels where it moved to the cell membrane without glycophorin A even at low band 3 RNA concentrations. This indicates that glycophorin A alters the time taken for band 3 to be transported to the cell membrane rather than the degree of transport (Groves & Tanner, 1992). Glycophorin A on the other hand is never transported to the erythrocyte membrane if band 3 is absent, as observed in band 3 deficient mice (Hassoun *et al*, 1998a). The weak interaction of band 3 and glycophorin A at the erythrocyte membrane is transient but could delay anion exchange until the erythrocyte is physiologically mature (Groves & Tanner, 1992).

1.4.6 Band 3 Self Association: A Marker of Erythrocyte Age

Band 3 is known to self-associate into dimers and higher oligomers. Dimers do not associate strongly with the membrane skeleton and move freely within the erythrocyte membrane. Tetramers are formed when the cytoplasmic domains of neighbouring dimers undergo cross-linking and become tightly bound to the membrane skeleton through association with ankyrin (Tanner, 1993). Casey and Reithmeier (1991) indicated that band 3 could be used as a marker of cell age by analysing the mixture of band 3 oligomers within the cell membrane. They found that 56 percent band 3 oligomers purified from aged erythrocytes were larger than dimers compared to only 22 percent in younger erythrocytes. Furthermore, band 3 immobilised by the membrane skeleton through interaction with ankyrin was of the higher oligomer type whereas free band 3 tended to be dimeric. They also noted an increase in disulphide cross-linkage in higher order band 3 oligomers when compared to band 3 dimers, apparently a consequence and not a cause of the heightened oligomerisation. This could lead to sulphydryl oxidation of band 3 which Beppu *et al* (1990) linked to the formation of the erythrocyte senescence

antigen. During sulphydryl oxidation, a reduced cysteine residue loses a hydrogen molecule from its sulphydryl group, becoming oxidised. Oxidised cysteine residues interact to form disulphide bonds, altering the structure of proteins and affecting the integrity of the cell (Stryer, 1998). Band 3 flippase activity, believed to be significant in maintaining the membrane integrity of the erythrocyte by actively transporting phosphatidylserine and phosphatidylethanolamine to the inner membrane, is decreased as band 3 aggregates form. This causes the erythrocytes to become 'sticky', increasing the likelihood of coagulation and splenic sequestration (Hassoun *et al*, 1998).

1.5 Protein 4.1

Protein 4.1 constitutes approximately five percent of the membrane skeleton bulk (Cohen, 1983) and is involved in two important membrane interactions. Protein 4.1, spectrin and actin comprise the ternary complex, the primary horizontal interaction of the erythrocyte membrane (Cohen, 1983). Protein 4.1 also binds to glycophorin C, further linking the membrane skeleton to the lipid bilayer (Reid *et al*, 1990).

1.5.1 The Protein 4.1 Gene

Protein 4.1, like many other erythroid proteins, is detected in an erythroid and a non-erythroid form (Conboy, 1993). Unlike the other erythroid proteins, however, a number of widely dispersed protein 4.1 isomers are encoded by a common gene located at 1p32-1pter (Conboy *et al*, 1986). The diversity of protein 4.1 (figures 1.6 and 1.7) exists as a direct result of complex pre-messenger RNA splicing and the use of two different translation initiation (AUG) codons (Conboy, 1993). Erythroid protein 4.1 is encoded from the downstream initiator (AUG-2) codon located in exon 4 of the gene (Tang *et al*, 1988, 1990a).

Tang *et al* (1988, 1990a) indicated that the pre-messenger RNA splicing is linked to five specific stretches of DNA that they named Motifs I to V (Figure 1.6) in order of discovery. Inclusion or exclusion of these motifs appears to be tissue

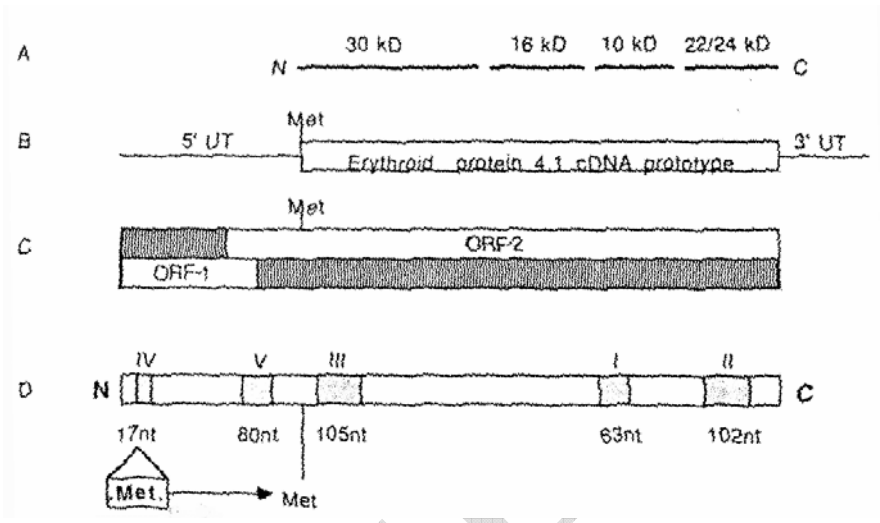

Figure 1.6: The Variable Motifs of Protein 4.1

(A)Structural domains of human erythrocyte protein 4.1. (B) Model of erythroid protein 4.1 cDNA. (C) Open reading frames 1 and 2. (D) Motifs I to V (shaded boxes) are named in order of discovery. The number of nucleotides comprising each motif is indicated. Initiation codons (Met) are located in motif IV and downstream of motif V. (Tang *et al*, 1990b)

specific. Motif I is 63 nucleotides long and encodes 21 amino acids within the highly conserved spectrin-actin binding domain discussed in section 1.5.4. Motif I is only expressed in the erythroid form of protein 4.1 and is often referred to as the erythroid motif. Motif II is found within the 22/24 kD carboxy domain while motif III encodes 35 amino acids located within the 30 kD membrane-binding domain (Tang *et al*, 1988, 1990a, 1990b). These two motifs are found in over 90% of protein 4.1 mRNAs (Tang *et al*, 1990a, 1990b). Motifs IV and V are located within the 5'-untranslated region and are linked to two overlapping open reading frames (ORF). Motif IV is found within ORF1 and contains an initiation codon. Motif V lies within ORF2 and contains a number of termination codons which prevent the use of ORF1 and the initiation codon in motif IV.

Deletion of motif V in conjunction with inclusion of the extreme 5' motif IV results in the use of both ORF1 and ORF2 and the 135 kD nonerythroid form of protein 4.1 is produced from the upstream initiation codon in motif IV. Any other permutation of these two motifs results in the use of an initiation codon downstream of motif V and the 80 kD protein 4.1 (generally only present in erythrocytes) is synthesised (Tang *et al*, 1990a, 1990b).

1.5.2 Erythroid Protein 4.1

The erythroid protein 4.1 isomers (proteins 4.1 a and b) are globular proteins with apparent molecular weights of 78 kD (protein 4.1b) and 80 kD (protein 4.1a), the difference being caused by the deamidation of asparagine 502 to aspartate (Inaba *et al*, 1992). The actual amino acid sequence predicted from the cDNA sequence indicates a 66kD protein (Conboy, 1993). Four structural domains (see Figure 1.7) have been identified on protein 4.1: the 30 kD amino terminal domain, a 16 kD linker domain, the 10 kD spectrin-actin binding domain and the 22/24 kD carboxy terminal domain (Leto & Marchesi, 1984; Conboy, 1993). The 22/24 kD domain undergoes two post-translational modifications: deamidation of asparagine 502 (Inaba *et al*, 1992) and O-linked glycosylation with N-acetylglucoseamine (Holt *et al*, 1987). The precise functions of the 16 kD and 22/24 kD domains are unclear, although the latter is believed to bind a variety of regulatory and

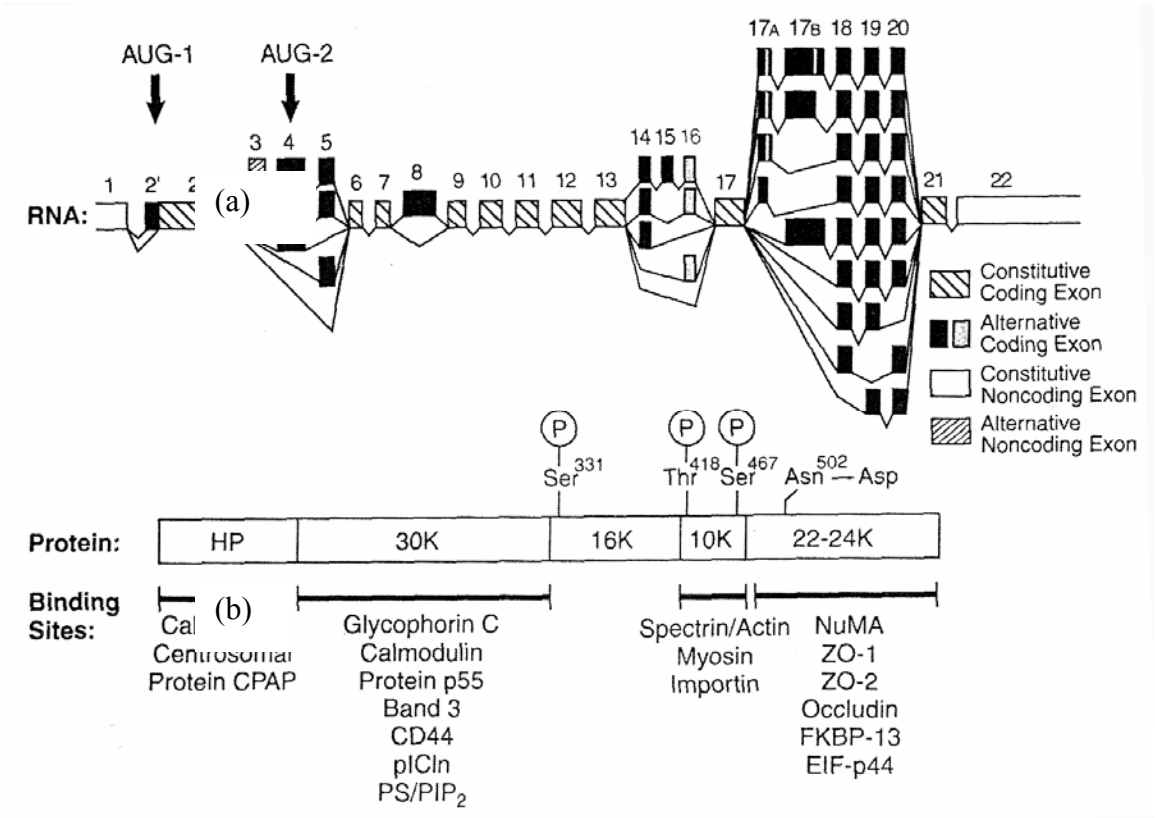

Figure 1.7: Alternative Splicing of Protein 4.1 mRNA and the Resultant Functional Domains of Protein 4.1

(a) Depiction of alternative splicing of protein 4.1 mRNA. (b) The functional domains of protein 4.1 indicating the proteins to which they bind. HP=Domain present in high molecular weight protein 4.1 and not in erythroid protein 4.1. The 10kD domain is found only in erythroid protein 4.1.

(Walensky *et al*, 2003)

junctional proteins (Grimm *et al*, 2002). The 10 kD and 30 kD domains are known to be vital to the maintenance of the erythroid membrane skeleton.

1.5.3 The Protein 4.1 Membrane Binding Domain

The 30kD amino terminal comprises the membrane-binding domain and is homologous with the diverse Ezrin-Radixin-Moesin (ERM) proteins, phosphorylation-activated proteins responsible for the linking actin filaments to the plasma membrane (Tsukita *et al*, 1997). The protein 4.1 membrane-binding domain binds to glycophorin C and also to band 3 to aid in linking the membrane skeleton to the lipid bilayer (Cohen, 1983).

The major membrane attachment site for protein 4.1 is glycophorin C, although it can also bind glycophorin D which is encoded on by the same gene. The binding can be direct (55 percent of protein 4.1) or indirect (28 percent of protein 4.1) where p55 mediates the binding (figure 1.8) (Hemming *et al*, 1995). The remaining 17 percent of erythroid protein 4.1 binds to band 3. The motif of protein 4.1 which interacts with glycophorin C is thought to be L^{37}EEDY, which is negatively charged and binds to the positively charged residues (Y^{85}HRKG) of the cytoplasmic domain of glycophorin C. Protein 4.1 and p55 are believed to interact through the SH3 domain of p55 and a proline-rich sequence of the protein 4.1 membrane-binding domain (Hemming *et al*, 1995).

Protein 4.1 also binds to band 3 through L^{37}EEDY. This binding is regulated by calcium levels in the cells and by the binding of calmodulin by protein 4.1 (figure 1.9) (Nunomura *et al*, 2000). Nunomura *et al* (2000) identified two binding sites for calmodulin: a calcium sensitive binding site encoded in exon 9 (referred to as peptide 9) and an exon 11-encoded calcium-independent site (referred to as peptide 11) with an 80-fold greater affinity for calmodulin. Studies of the amino acid sequence of both these sites indicated a recognized calmodulin binding motif: AKKL<u>X</u>XXXVXX(X)HXXXXL, X denoting any amino acid and (X) indicating the possible presence of an extra unknown amino acid. The essential amino acid (underlined) has been identified as serine 185 (peptide 9) or tryptophan 268

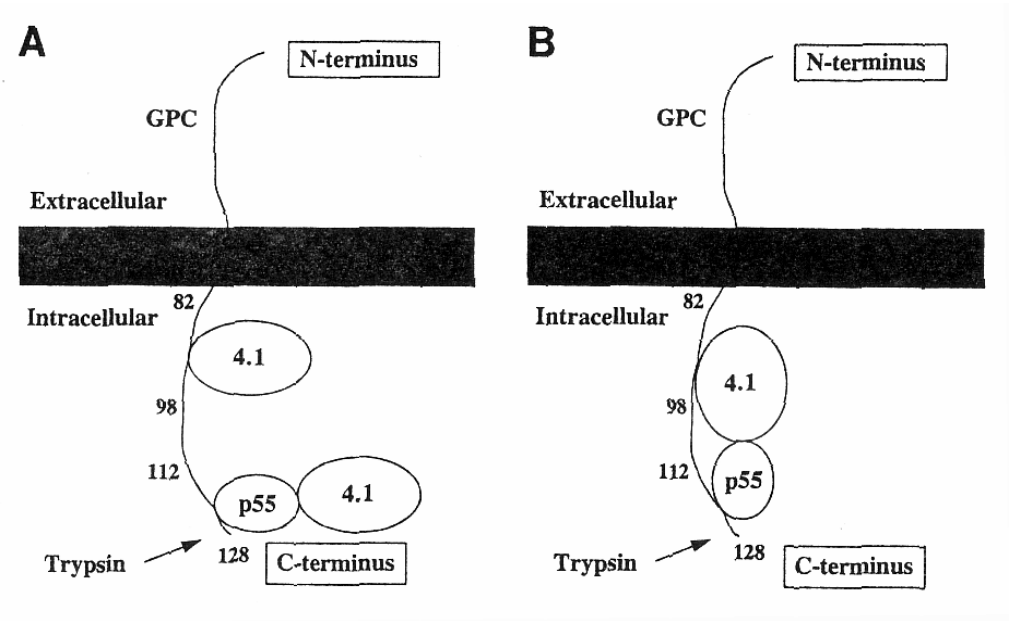

Figure 1.8: Models for Protein 4.1-Glycophorin C Binding

Model A shows the direct and indirect binding of two protein 4.1 molecules while model B shows the indirect binding of only one protein 4.1 molecule through p55. In model B, other glycophorin C molecules would bind protein 4.1 directly as shown in model A but the indirect protein 4.1 binding would not occur.

(Hemming *et al*, 1995)

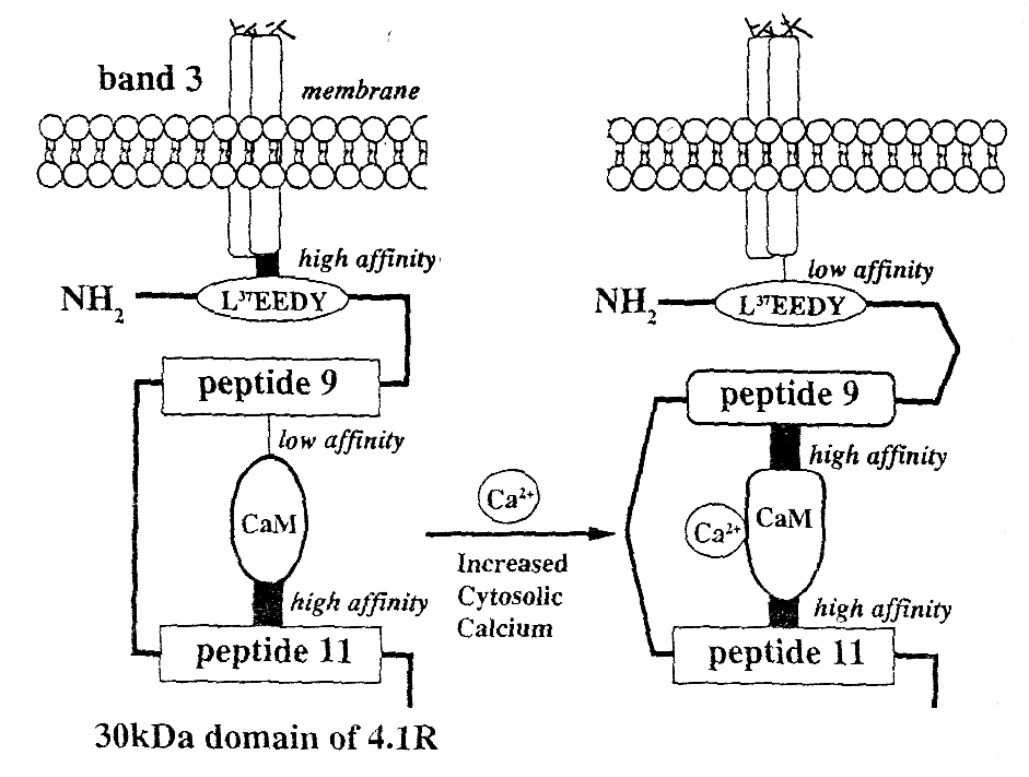

Figure 1.9: Model for Calmodulin Modulation of Protein 4.1-Band 3 Binding

When calcium (Ca^{2+}) concentrations are low, calmodulin (CaM) binds to calcium-independent peptide 11, causing high affinity protein 4.1 binding to band 3. When calcium concentrations increase, calmodulin binds with greater affinity to peptide 9 and the affinity of protein 4.1 for band 3 decreases.

(Nunomura *et al*, 1999)

(peptide 11). At physiological calcium concentrations (less than 1.0μM) peptide 11, with its calmodulin-sensitive tryptophan residue, binds calmodulin and the resultant protein 4.1 conformation has a high affinity for band 3. An influx of calcium increases the affinity of the less receptive serine in peptide 9 for calmodulin and the rearrangement of protein 4.1 causes a conformation not conducive to band 3 binding (Nunomura *et al*, 2000).

As both glycophorin C and band 3 bind to protein 4.1 through L^{37}EEDY, it is likely that there is calcium-regulated competition for this binding site. Blocking of the glycophorin C-binding sites for protein 4.1 results in decreased binding of protein 4.1, spectrin and ankyrin to the cell membrane (Reid *et al*, 1990) whereas the blockage of all protein 4.1-binding sites on band 3 has no apparent effect on the protein 4.1-mediated binding of spectrin and ankyrin to the lipid portion of the erythrocyte membrane (Workman & Low, 1998). This indicates that the protein 4.1-glycophorin C interaction is important for the stability of the erythrocyte membrane skeleton, while the association of protein 4.1 and band 3 may regulate the band 3-ankyrin interaction (An *et al*, 1996). When calcium concentrations in the erythrocyte increase and the affinity of protein 4.1 for band 3 decreases, protein 4.1 is more likely to bind to glycophorin and band 3 is more likely to bind to ankyrin. Both of these interactions stabilise the erythrocyte membrane (Nunomura *et al*, 1999).

1.5.4 The Protein 4.1 Spectrin-Actin Binding Domain

The 10 kD spectrin-actin binding domain is encoded by alternatively spliced exon 16 and constitutive exon 17. It contains three important motifs: spectrin-binding motifs 1 and 2 and the actin binding motif (figure 1.10). The eight amino acid actin-binding domain (LKKNFMES) is located between amino acids 19 and 26 encoded by exon 17. Amino acids one to three (LKK) and seven and eight (ES) are conserved throughout actin binding sites, but are separated by a sequence consisting of three uncharged amino acids (NFM) rather than by a single hydrophobic or uncharged residue as is seen in the typical six amino acid binding sequence (LK[K/R/H][Q/A]E[S/T]). As other actin binding sites only bind actin

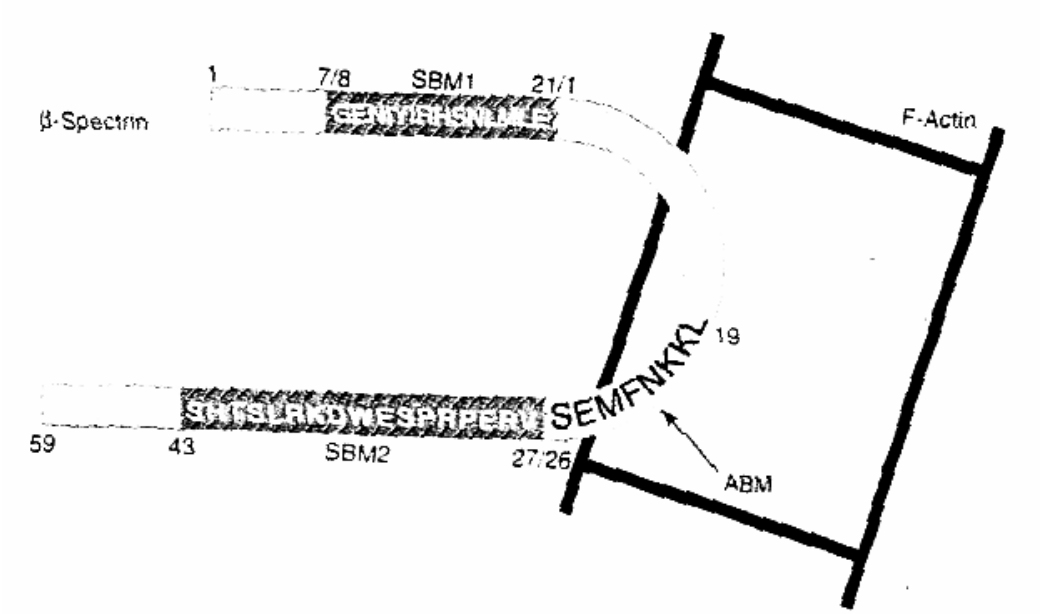

Figure 1.10: Diagrammatic Representation of the Protein 4.1 Spectrin-Actin Binding Domain (Grimm *et al*, 2002)

The hairpin loop represents protein 4.1 (amino acids 1to 59). The black region with white lettering between glycine 8 and glutamic acid 21 indicates the minimal portion of spectrin binding motif 1 (SBM 1) and that between valine 27 and serine 43 depicts spectrin binding motif 2 (SBM2). The amino acids (large black letters) between leucine 19 and serine 26 comprise the actin binding motif (AMB). Actin is depicted by blocks and is labelled F-actin (filamentous actin) and the position of β-spectrin is indicated.

in the presence of actinin and cannot form ternary complexes, this so-called spacer sequence is believed to be important for the correct functioning of erythroid protein 4.1. Nucleotide substitution studies indicate that the uncharged nature of the residues of this site is also important (Grimm *et al*, 2002).

The bipartite spectrin binding site flanks the actin binding site (figure 1.10). Spectrin-binding motif 1 comprises the 21 amino acid peptide (KKRERLDGENIYIRHSNLMLE) encoded by exon 16, although the minimal spectrin-binding motif has been shown to involve only the C-terminal 14 residues (underlined). It is separated from the amino acids of spectrin-binding motif 2 and the actin-binding motif by a loop of amino acids (amino acids 1 to 18 encoded by exon 17) which are believed to be structurally if not functionally important for the formation of the ternary complex. Spectrin-binding motif 2 (VPEPRPSEWDKRL STHS) is located between amino acids 27 to 43 encoded by exon 17 and lies directly next to the actin-binding motif (Grimm *et al*, 2002).

The association of spectrin, actin and protein 4.1 to form a ternary complex is important for the maintenance of membrane elasticity and shape (Cohen, 1983). Phosphorylation of protein 4.1 plays a major role in the inhibition of spectrin-actin binding. The amino acids known to be phosphorylated by phosphokinase A to cause this inhibition are serine 331 (in the 16 kD domain) and serine 467 and threonine 418, which both lie in the 10 kD domain (Horne *et al*, 1993). Horne *et al* (1993) showed that spectrin-binding domain 1 (motif 1) is vital for the formation of the ternary complex of spectrin, actin and protein 4.1. Recombinant protein 4.1 molecules, some lacking motif I, were incubated with spectrin or actin individually or spectrin and actin together and then isolated by centrifugation. Insignificant amounts of protein were present in the pellet unless spectrin, actin and Motif I were present. (Horne *et al*, 1993) These results were corroborated in experiments using recombinant protein 4.1 with mutated or absent actin-binding motifs (Discher *et* al, 1993; Grimm *et al*, 2002)

1.5.5 Conversion of Protein 4.1b to Protein 4.1a: An Indicator of Cell Age

The time controlled conversion of protein 4.1b to protein 4.1a can be used as an indicator of erythrocyte age. Protein 4.1b is common in reticulocytes, whereas the quantity of protein 4.1a is initially low and increases over the lifespan of the erythrocyte. The distinguishing feature of protein 4.1b is the presence of asparagines (polar negative) at positions 478 and 502, both of which are found in the elastic regions of the protein. Protein 4.1a expresses aspartate (polar uncharged) at these positions. The modification is caused by the deamidation of asparagine, a common post-translational modification. (Inaba *et al*, 1992) The deamidation of asparagine is not catalysed by enzymes and occurs when a nitrogen molecule on the amino acid situated at the carboxyl side of asparagine attacks the carbonyl carbon of asparagine, converting it to succinimide. Succinimide is a five carbon ring that is hydrolysed to give rise to aspartate and isoaspartate. The rate of the formation of succinimide is limited by the amino acid to the carboxyl side of the asparagines, since this is the amino acid that initiates the reaction (Inaba *et al*, 1992). Asparagine 478 is rapidly deamidated, amino acid 479 being a glycine which has a highly reactive nitrogen. Asparagine 502 is flanked at its carboxy side by alanine, and is deamidated slowly over the life of the erythrocyte, giving rise to protein 4.1a. The apparent molecular weight difference between protein 4.1a and 4.1b when the erythrocyte membrane proteins are separated using polyacrylamide gel electrophoresis is an artefact caused by these different amino acid residues (Inaba *et al*, 1992). While the significance of this change is undefined, it can be used to determine the approximate age of the erythrocytes. Reticulocytes have more protein 4.1b than protein 4.1a and the protein 4.1a:b ratio can be used as an indicator of reticulocytes, a decreased ratio indicating increased reticulocyte numbers.

1.6 Erythrocyte Membrane Disorders

The scientific advances of the past century have allowed diseases such as haemolytic anaemia to be examined from the clinical to the molecular level. Anaemia is defined as a pathological condition where the blood haemoglobin

concentrations are below normal, haemolytic referring to anaemia caused by the lysis of the erythrocytes (haemolysis). A significant number of haemolytic anaemias occur as a result of physically or functionally deficient erythrocyte membrane proteins (Coetzer & Palek, 1993; Tse & Lux, 1999). The protein abnormalities affect both the vertical and the horizontal protein interactions, causing the erythrocyte membrane to weaken and lose its biconcave shape and deformability. The degree to which the membrane is affected varies greatly, but can lead to premature erythrocyte destruction (Palek & Lambert, 1990). The haemolytic anaemias of interest to this study were hereditary spherocytosis (HS) and hereditary elliptocytosis (HE). Both are clinically and genetically heterogeneous disorders (Palek & Lambert, 1990).

1.7 Hereditary Spherocytosis

1.7.1 Profile of Hereditary Spherocytosis

HS is a haemolytic disorder characterised by osmotically fragile erythrocytes that are spherical rather than biconcave. (Palek & Lambert, 1990; Coetzer & Palek, 1993; Tse & Lux, 1999; Eber & Lux, 2004) It is the primary cause of hereditary haemolytic anaemia in Caucasians (1:5000 to 1:10 000) and has been documented in South African Blacks (Coetzer, 1999). A number of individuals with HS are asymptomatic or have only mild anaemia; others display life-threatening haemolysis and anaemia (Coetzer & Palek, 1993; Eber & Lux, 2004).

1.7.2 Clinical and Molecular Diagnosis of Hereditary Spherocytosis

Patients with severe HS generally present with jaundice and anaemia. Anaemia occurs when the compensation of haemolysis is inadequate (Palek & Lambert, 1990; Tse & Lux, 1999; Eber & Lux, 2004) and jaundice arises when excessive quantities of unconjugated bilirubin (the break-down product of haemoglobin) are deposited in the skin and mucous membranes (Walensky *et al*, 2003). Haemoglobin levels are usually decreased and bilirubin levels are increased as erythrocytes are destroyed. Reticulocyte counts increase as greater numbers of immature erythrocytes are released into the blood stream in an effort to

compensate for the heightened levels of erythrocyte destruction (Palek & Lux, 1983; Walensky *et al*, 2003; Eber & Lux, 2004). Many patients with severe HS are unable to counteract the haemolysis and are given transfusions of packed erythrocytes on a regular basis (Walensky *et al*, 2003; Eber & Lux, 2004).

Spherocytes are osmotically fragile, fragmenting at higher concentrations of salt than normal erythrocytes. Coombs' tests for autoimmune disorders causing haemolysis are negative (Walensky *et al*¸ 2003; Eber & Lux, 2004). Spherocytes and spiculated erythrocytes are visible using a light microscope (Palek & Lux, 1983; Tse & Lux, 1999; Eber & Lux, 2004). Splenomegaly, the distension of the spleen, may be present as spherocytes become trapped in the splenic sinuses for extended periods of time. Severely affected patients generally undergo splenectomy, the removal of the spleen, to minimise erythrocyte destruction (Coetzer & Palek, 1993; Tse & Lux, 1999; Eber & Lux, 2004) This procedure can fully alleviate symptoms of HS in moderately affected patients and partially ameliorate symptoms in patients with severe HS, but is generally only performed as a last resort due to the increased risk of sepsis, infection and ischaemic heart disease. Splenectomised patients display acanthocytosis, the presence of erythrocytes with irregular membrane projections ('thorns'). This anomaly occurs as splenic conditioning (the removal of unsupported portions of the erythrocyte membrane) is no longer possible (Walensky *et al*, 2003; Eber & Lux, 2004).

The protein defect causing the spherocytosis can be detected using quantitative polyacrylamide gel electrophoresis but this labour intensive technique is not used in routine laboratories. The sensitivity of this technique is moderate, very small deficiencies resulting in HS often being overlooked. The genetic mutation underlying the protein defects can be detected using various screening and sequencing techniques but routine genetic screening is not possible unless a specific mutation is known in a kindred as most of the protein defects tend to be familial (personal communication, Prof. Theresa Coetzer).

1.7.3 Pathogenesis of Hereditary Spherocytosis

HS occurs when one or more of the proteins involved in vertical interactions of the membrane skeleton are deficient (figure 1.11) (Coetzer & Palek, 1993). In approximately 75 percent of cases HS is an autosomal dominant disease (Coetzer & Palek, 1993) and homozygous HS is rare. The remaining 25 percent of cases are caused by spontaneous dominant mutations or recessive mutations (Tse & Lux, 1999). The most common membrane protein deficiencies within the well-studied Caucasian population of the USA and Europe are those of spectrin (30%) or ankyrin with a secondary spectrin deficiency (30 to 45%). Band 3 deficiencies are responsible for another 20 percent of Caucasian HS cases and a small percentage of cases are caused by ankyrin deficiencies or, particularly in the Japanese population, by a partial or total deficiency of protein 4.2 (Eber & Lux, 2004; Tanner, 1997; Yawata *et al*, 2000).

Spherocytes and microspherocytes, the markers of the disorder, develop when portions of the erythrocyte cell membrane, unsupported by the membrane skeleton, form lipid vesicles which are removed from the erythrocyte (figure 1.11). This results in a decreased surface area to volume ratio (Coetzer & Palek, 1993; Eber & Lux, 2004; Tse & Lux, 1999). The spherocytes are not able to deform under the shear pressure of the microvasculature and become trapped in the narrow sinuses of the spleen (Coetzer & Palek, 1993; Eber & Lux, 2004; Mohandas *et al*, 1983; Palek & Lambert, 1990). Ultimately the entrapped cells are phagocytosed and digested (Coetzer & Palek, 1993; Eber & Lux, 2004; Tse & Lux, 1999).

1.7.4 The Role of Band 3 in Hereditary Spherocytosis

Studies using band 3 null mice (produced through the targeted inactivation of band 3) showed that only five to ten percent survived the neonatal period and those that survived to adulthood had extremely severe haemolytic anaemia. The high mortality rate was linked to their propensity for thrombosis (Hassoun *et al*, 1998). This is believed to be a result of destabilisation of the lipid bilayer (Hassoun *et al*, 1998), possibly because band 3 flippase activity (section 1.4.6) is lost, causing the presence of highly negative phosphatidylserine on the outer

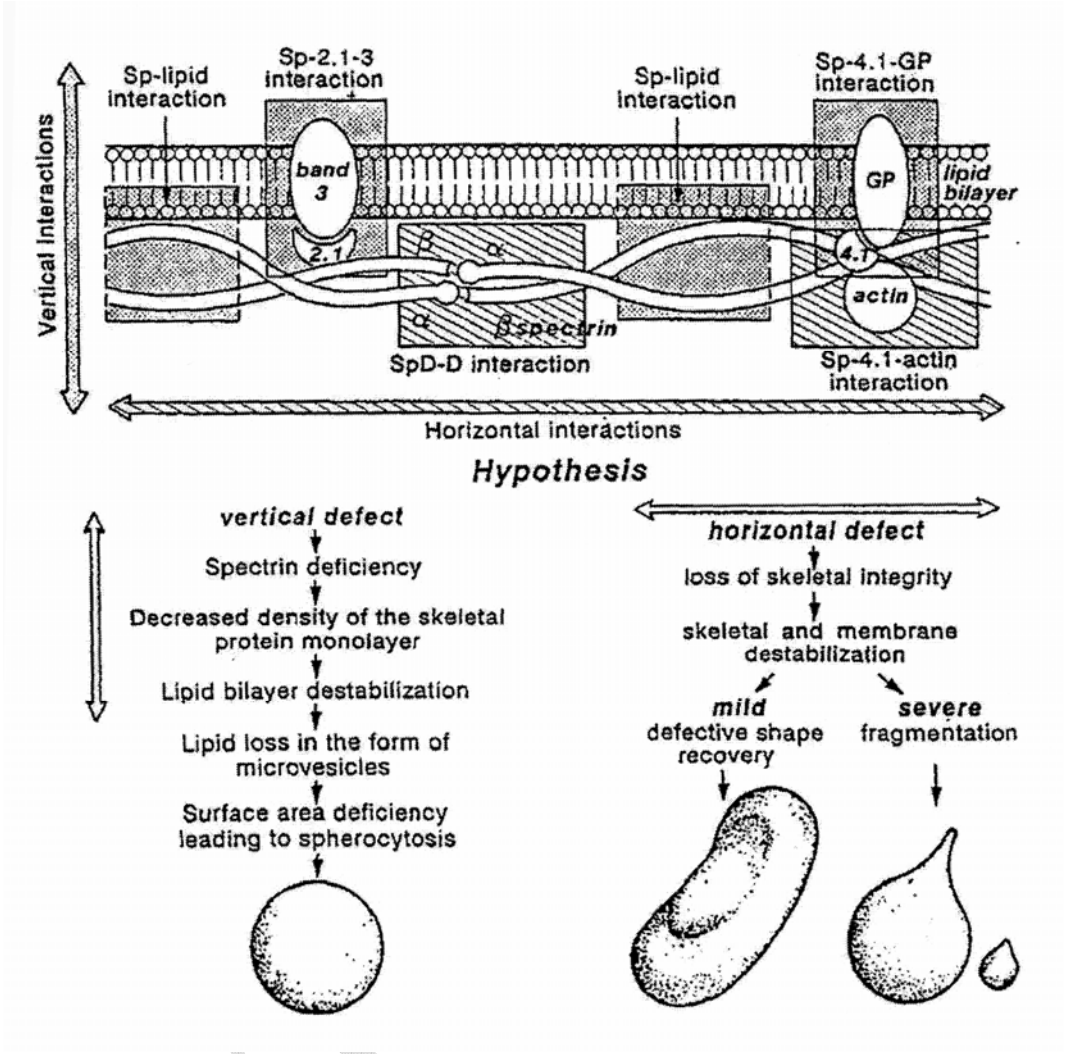

Figure 1.11: Pathogenesis of Erythrocyte Membrane Disorders

(Coetzer & Palek, 1993)

The major vertical (vertical arrow and upper labels) and horizontal (horizontal arrow and lower labels) interactions are shown. When a protein involved in one or more of these interactions is dysfunctional or deficient, the erythrocyte membrane is altered. This results in the production of abnormal erythrocytes and may cause disease. Abnormalities of the vertical interactions (left panel) cause the production of spherocytes while abnormalities of the horizontal interaction (right panel) either lead to the presence of elliptocytes or fragmentation of the erythrocytes.

membrane. This can stimulate the coagulation pathway, producing a many-fold increase in the production of the powerful clotting factor, thrombin. Phosphatidylserine also signals macrophages to phagocytose the erythrocytes (this is usually only seen in aged erythrocytes) (Hassoun *et al*, 1998). The external presence of phosphatidylserine has also been linked to membrane fusion and vesiculation (Hassoun *et al*, 1998) and this, combined with the loss of band 3 membrane support (Jarolim *et al*, 1996), could be responsible for the abnormal budding of the membrane and hence the presence of microspherocytes and a decreased surface area to volume ratio (Hassoun *et al*, 1998).

HS as a result of band 3 deficiencies was documented in the early 1990s and since that time over 50 different band 3 genotypes have been described (see Appendix A). The mutations are named according to the city in which the proband resided. The majority of these mutations are point mutations which alter an amino acid in the band 3 protein, resulting in altered protein binding or insertion into the erythrocyte membrane. Some point mutations create premature termination codons which lead to truncation of the protein and others cause the mRNA to become unstable, degrading before the protein can be synthesised (Tanner, 1997; Tse & Lux, 1999). A number of band 3 mutations are found in so-called hotspots, codons in which a number of mutations have been described, and often occur over CpG dinucleotides.

The majority of band 3-deficienct HS patients have single heterozygous mutations, but compound heterozygotes and homozygotes have been described. The severity of the band 3 deficiency varies from mutation to mutation but compound heterozygotes (different nucleotide changes in either the same or two paired alleles) and homozygotes (identical mutations in both alleles) tend to be more severe. Examples of compound heterozygous mutations include the band 3 Okinawa/Fukuoka and band 3 Cape Town/Prague III combinations. Band 3 Okinawa in *trans* to band 3 Fukuoka results in a severe loss of band 3 (66% band 3 present) and no protein 4.2 (Kanzaki *et al*, 1997). Band 3 Fukuoka (G130R) is normally linked to a mild band 3 deficiency with poor protein 4.2 binding,

indicating the involvement of amino acid 130 in protein 4.2 binding (Ideguchi *et al*, 1994). Band 3 Cape Town and band 3 Prague III together cause a severe phenotype with only 61 percent of normal band 3 and 80 percent of normal protein 4.2 (Bracher *et al*, 2001).

Only three cases of probands with homozygous band 3 mutations have been described in humans. Clinically, two had severe haemolytic anaemia while the third had only mild symptoms. The first case of HS with a homozygous band 3 mutation was presented in 1997 although the full case history and follow-up was only published by Ribeiro *et al* in 2000. This proband had extremely severe HS, requiring monthly blood transfusions, which was linked to homozygous band 3 Coimbra (V488M). Band 3 Coimbra is a missense mutation linked to typical HS with partial band 3 and protein 4.2 deficiencies in the heterozygous state. Amino acid 488 is located near the beginning of the fourth transmembrane region of band 3 and the V→M substitution most likely affects the insertion of the protein into the erythrocyte membrane (Alloisio *et al*, 1997). The infant described by Ribeiro *et al* (2000) had no band 3 and an associated lack of protein 4.2; spectrin, ankyrin and glycophorin A deficiencies were also noted. Given that this was the third child of a consanguineous couple (the first child being stillborn and the second pregnancy being terminated) and that extreme medical intervention was required, it was hypothesised that band 3 mutations resulting in the total absence of band 3 were incompatible with life without comprehensive therapy (Ribeiro *et al*, 2000). These observations were supported by the studies of Perrotta *et al* (1998) who described an infant with homozygous band 3 Neapolis (intron 2, +2 T→C). This splicing mutation was observed in the heterozygous state in a number of family members, including both parents (a consanguineous couple) and generally caused mild, compensated HS, although two individuals were splenectomised. Two protein transcripts were detected: one using a proposed alternative initiation codon and one with a premature termination codon. This shortened protein may be inserted but it was hypothesised that band 3 was completely absent (Perrotta *et al*, 1998). The third proband a homozygous band 3 mutation, band 3 Fukuoka (G130R), was diagnosed with compensated HS at age 29 (Inoue *et al*, 1998). As

indicated in the previous paragraph, heterozygous band 3 Fukuoka is normally associated with extremely mild HS (Ideguchi *et al*, 1994), as was also seen in the parents of this proband who were clinically normal although they were carriers of band 3 Fukuoka. The homozygous proband had 90 percent of normal band 3, indicating that the protein was inserted into the membrane, but only 45 percent of normal protein 4.2. This indicated the involvement of cytoplasmic amino acid 130 in protein 4.2 binding and also demonstrated that partial protein 4.2 deficiencies did not interfere with the erythrocyte membrane integrity (Inoue *et al*, 1998).

1.7.5 Factors Modulating Band 3 Deficiency

In some kindred a more acute deficiency of the band 3 protein is caused by the presence of a polymorphism (rather than a second mutation) in *trans* to the causative mutation. Known modulating factors are band 3 Montefiore, band 3 Mondego and band 3 Genas.

Band 3 Montefiore (E40K) is encoded by exon 4 and a lysine (AAG) replaces a glutamic acid (GAG). It is generally considered to be a polymorphism, although it is linked to decreased protein 4.2 binding with clinically silent HS, indicating that it could be a recessive mutation. Band 3 Mondego is a two-part mutation: E40K (band 3 Montefiore) and P147S. P147S is located in exon 6 and a serine (TCT) replaces proline (CCT) 147. When the mutations of band 3 Mondego are found in *trans* with band 3 Coimbra, the severity of the mild band 3 Coimbra phenotype is significantly worsened (Allosio *et al*, 1997). Band 3 Genas is located within the band 3 promoter region. The nucleotide change (g→a) occurs 62 nucleotides upstream of the initiation codon in exon 2, but when found in *trans* with band 3 Lyon causes a moderate rather than the normal mild band 3 deficiency (Allosio *et al*, 1996).

1.7.6 The Band 3 Memphis Polymorphism

Band 3 Memphis causes ‘slow-moving’ band 3 on polyacrylamide gels. Band 3 Memphis I (K56E) comprises a single nucleotide change at the first position (nucleotide 166) of codon 56 within exon four of band 3. Ultimately lysine (AAG)

56 is replaced by a glutamic acid (GAG) with no apparent effect. As this corresponds with the band 3 amino acid sequence found in mice, rats and rabbits it is believed that the lysine more commonly seen in humans is a more recent form of band 3 (Jarolim *et al*, 1992b). In band 3 Memphis II (P854L) one or both alleles carry a C→T substitution of nucleotide 2561 in addition to the band 3 Memphis I polymorphism. This nucleotide change at position 2561 results in the replacement of proline 854 with a leucine, once again without apparent effect (Bruce *et al*, 1994a). Band 3 Memphis II has been found in *cis* to band 3 Okinawa (Kanzaki *et al*, 1997).

1.8 Hereditary Elliptocytosis

1.8.1 Profile of Hereditary Elliptocytosis

At the level of the membrane skeleton, HE disturbs the horizontal rather than vertical interactions (figure 1.11) (Coetzer & Palek, 1993; Tse & Lux, 1999). This decreases the ability of the erythrocytes to resume the biconcave shape lost under the shear forces of the microvasculature resulting in elliptical rather than biconcave erythrocytes (Coetzer & Palek, 1993; Palek & Lux, 1983). The elliptocytes are generally mechanically unstable, lysing at lower shear pressure than normal erythrocytes, and may also be thermally unstable, being destroyed at a lower temperature than normal erythrocytes (Palek & Lambert, 1990; Palek & Lux, 1983).

1.8.2 Clinical and Molecular Diagnosis of Hereditary Elliptocytosis

HE is more clinically heterogeneous than HS, three different classes being identified. The first and most common is simple HE in which the biconcave erythrocytes appear elliptical. The second is spherocytic HE in which both elliptocytes and spherocytes are seen and the last is stomatocytic HE (Coetzer & Palek, 1993; Palek and Lambert, 1990; Walensky *et al*, 2003).

No definitive clinical picture exists for common heterozygous HE, even within separate families. The silent carrier state is often identified only through sensitive

tests for membrane stability or genetic analysis of kindred where a mutation has been identified. Typically patients with common heterozygous dominant HE have more than twenty percent elliptocytes in a peripheral blood smear, although the haemolysis varies from minimal to moderate. Homozygotes and compound heterozygotes may display more severe, even life-threatening, anaemia and haemolysis with poikilocytes (erythrocytes with a tear-drop shape), pincered cells (erythrocytes with a mushroom-like shape), spherocytes and fragmented erythrocytes being visible in addition to the elliptocytes. They tend to have splenomegaly, jaundice and aplastic crises. Reticulocytes are prominent and the erythrocytes show increased osmotic fragility and a decreased mean cell volume as a result of cell fragmentation (Palek & Lux, 1983; Walensky *et al*, 2003).

Spherocytic HE is also dominantly inherited and is characterized by the presence of osmotically fragile spherocytes, microspherocytes, microelliptocytes and elliptocytes that are rounder than those seen in simple HE. As with the more severe cases of common HE, the patients suffer from anaemia, jaundice and aplastic crises, reticulocytes indicating compensated haemolysis. This hybrid disease is very rare in Caucasian populations and has not been reported in other racial groups (Palek & Lux, 1983; Walensky *et al*, 2003).

Stomatocytic HE or Southeast Asian Ovalocytosis (SAO) is predominantly found in Asia (five to 25 percent of native Asians) but has also been described in the Cape Coloured population (Coetzer *et al*, 1996). This is the only genetically homogenous form of HE and all patients carry a heterozygous 72bp deletion within the band 3 gene. The homozygous form is lethal, but in heterozygotes haemolysis and anaemia are uncommon. The erythrocytes are unique because of their rigidity and the presence of a distinct bar traversing the elliptocyte. SAO is believed to provide some protection against the invasion of the erythrocyte by the malaria parasite (Palek & Lux, 1983; Walensky *et al*, 2003).

Clinically, patients with mild HE may show some compensated haemolysis, increased reticulocytes and on rare occasions a palpable spleen. Patients with

homozygous or compound protein defects show more severe anaemia and splenomegaly, many being transfusion dependent. As mentioned above, the elliptocytes are mechanically and thermally unstable. Treatment of common mild HE is seldom necessary except for folic acid supplements. Patients with severe HE respond well to splenectomy, elliptocytes experiencing the greatest shear pressure and risk of phagocytosis within the spleen (Walensky *et al*, 2003).

The proteins primarily implicated in HE are spectrin, protein 4.1 and glycophorin C (Palek & Lambert, 1990; Palek & Lux, 1983; Tse & Lux, 1999). Spectrin mutations usually result in functional rather than quantitative defects. An increase in the quantity of spectrin dimers and a decrease of spectrin tetramers is the most common cause of HE (Coetzer & Palek, 1993; Palek & Lux, 1983), leading to a lessening in membrane deformability. Spectrin deficiencies have been noted in 20% of Caucasian and 40% of Black patients suffering from HE (Walensky *et al*, 2003). Protein 4.1 mutations are usually quantitative (Coetzer & Palek, 1993; Conboy, 1993; Palek & Lux, 1983), leading to weakened spectrin-actin association (Palek & Lux, 1983; Tse & Lux, 1999).

1.8.3 The Role of Protein 4.1 in Hereditary Elliptocytosis

Hereditary elliptocytosis has been linked to protein 4.1 deficiencies and to non-functional protein 4.1 (Coetzer & Palek, 1993; Palek & Lambert, 1990; Palek & Lux, 1983). Protein 4.1 mutations are responsible for up to 40% of HE cases in the European and Arab populations (Delaunay *et al*, 1999) but have not been reported in the Black or Asian populations, although a number of Black patients with protein 4.1 deficiencies have been identified in studies in the Red Cell Membrane Unit, University of the Witwatersrand (personal communication, Prof. Theresa Coetzer).

Heterozygous mutations causing alterations to protein 4.1 generally cause mild HE, some patients appearing asymptomatic and only being identified through an increased presence of elliptocytes in a blood smear. In the homozygous form however the clinical symptoms can be extreme and elliptocytes, poikilocytes and

severe haemolysis are present (Tse & Lux, 1999). Low molecular weight forms of protein 4.1 are associated with increased clinical severity when compared to the high molecular weight form that has the spectrin-binding domain (Palek & Lambert, 1990).

Only seven protein 4.1 mutations linked to HE have been identified worldwide. These are protein 4.1 Algeria, protein 4.1 Annecy, protein 4.1 Aravis, protein 4.1 Hurdle-Mills, protein 4.1 Lille, protein 4.1 Madrid and an unnamed variant (Walensky *et al*, 2003). Unlike band 3 disorders, which are generally caused by point mutations, the identified protein 4.1 disorders tend to be linked to nucleotide deletions, four of the seven identified mutations being thus caused. Furthermore, the mutations described for protein 4.1 only affect the initiation codon or the spectrin-actin binding domain whereas band 3 mutations are dispersed throughout the gene.

Protein 4.1 Algeria was the first homozygous protein 4.1 mutation to be identified (Tchernia *et al*, 1981) and is caused by the abolition of 318 nucleotides within the portion of the protein 4.1 RNA, resulting in the loss of the downstream initiation codon and hence loss of mRNA translation (Conboy *et al*, 1986). Protein 4.1 Annecy is similarly caused by a large nucleotide deletion; in this case a 70kD fragment between exons two and twelve is removed. Once again the initiation codon is lost and the protein cannot be translated (Dalla Venezia *et al*, 1998). The unnamed mutation is an 80 amino acid deletion extending from lysine 407 to glycine 486 and causing the loss of the spectrin-actin binding domain. This removes much of the mechanical strength of the membrane skeleton although the resultant HE observed is mild (McGuire *et al*, 1988). Protein 4.1 Aravis, caused by the deletion of codon 447, also affects the spectrin-actin binding domain. This codon usually encodes one of two lysine residues (Lys447-Lys448) within the highly important spectrin-actin binding domain, probably altering the binding of spectrin or actin through conformational changes in the protein (Lorenzo *et al*, 1994). Protein 4.1 Hurdle-Mills also occurs within the spectrin-actin binding domain with amino acids 407 to 529 being duplicated. The HE observed is mild

and the membrane stability doesn't appear to be greatly influenced. The protein 4.1 isoform observed on the polyacrylamide gels is 95kD rather than 80kD in size (McGuire *et al*, 1988).

Single nucleotide changes have only been found in the initiation codon and two substitutions have been identified: Protein 4.1 Madrid and protein 4.1 Lille. The protein 4.1 Madrid missense mutation (ATG→AGG) causes the downstream (erythrocyte) initiation codon to be replaced by an arginine codon meaning that protein 4.1 mRNA translation does not take place. This mutation is silent in the heterozygous state but is linked to severe HE when found in the homozygous form (Dalla Venezia *et al*, 1992). The identification of protein 4.1 Madrid was soon followed by that of protein 4.1 Lille, another mutation affecting the erythroid initiation codon. In this case the methionine codon would be replaced by a threonine (ACG) codon and once again protein 4.1 would not be produced (Garbarz *et al*, 1995). Neither protein 4.1 Madrid nor protein 4.1 Lille appear to affect the nonerythroid protein 4.1 isoforms, all of which use the upstream initiation codon, although it was hypothesized by Dalla Venezia *et al* (1992) that mutations affecting both the upstream and downstream initiation codons or creating a termination codon between the two could be nonviable as protein 4.1 is dispersed throughout the body.

1.9 Aim of Study

The aim of this project was to determine the genetic mutations underlying band 3-deficient HS and protein 4.1-deficient HE in South African kindred.

1.9.1 Band 3 Strategy

- Extraction and amplification of band 3 deficient HS patient DNA
- Restriction enzyme digestion analysis for selected known mutations
- Single strand conformational polymorphism analysis of band 3 exons
- DNA sequencing of selected patient DNA

1.9.2 Protein 4.1 Strategy

- Western blot analysis of protein 4.1-deficient HE patient erythrocyte membrane proteins
- Extraction and amplification of protein 4.1-deficient HE patient DNA
- Restriction enzyme digestion analysis for selected known mutations