MOLECULAR GENETICS

OF

HUNTER SYNDROME

A thesis submitted to The University of Adelaide,
for the degree of

DOCTOR OF PHILOSOPHY

by

PETER JOHN WILSON, B.Sc. (Hons.)

Lysosomal Diseases Research Unit
Department of Chemical Pathology
Adelaide Children's Hospital
(A Division of the Women's and Children's Hospital)
NORTH ADELAIDE SOUTH AUSTRALIA 5006.

Department of Pathology
School of Medicine
The University of Adelaide,
GPO Box 498,
ADELAIDE SOUTH AUSTRALIA 5006.

May 1992

I dedicate this thesis to my Father and Mother

TABLE OF CONTENTS

TABLE OF CONTENTS

CHAPTER 2

MATERIALS AND METHODS

TABLE OF CONTENTS

CHAPTER 2

TABLE OF CONTENTS

CHAPTER 3

CHARACTERISATION OF AN IDS GENOMIC CLONE

TABLE OF CONTENTS

CHAPTER 4

ISOLATION AND EXPRESSION OF FULL-LENGTH IDS cDNA

TABLE OF CONTENTS

CHAPTER 5

MOLECULAR ANALYSIS OF MPS II MUTATIONS

TABLE OF CONTENTS

CHAPTER 6

THESIS SUMMARY

The research presented in this thesis was directed primarily towards the isolation and characterisation of the gene encoding iduronate-2-sulphatase (IDS). When this objective was achieved, the study of the mutations in the IDS gene responsible for MPS II were commenced. The following results were obtained:

1. A 49-mer oligonucleotide probe was designed from the N-terminal amino acid sequence of a 42 kDa polypeptide present as a major species in purified IDS.

2. Several genomic and cDNA libraries were probed with the 49-mer without success. The IDS gene has been localized to the Xq27/q28 boundary. Using this information it was thought that an X-chromosome-enriched genomic library may prove to be more successful. This appeared to be the case, as multiple positive clones were obtained in the first round of screening.

3. Two of the X-chromosome genomic clones were studied in detail by restriction enzyme analysis. A Southern blot with these restriction digests was probed with the 49-mer oligonucleotide to identify the smallest 49-mer positive fragment. This was found to be a 1.3-kb *Hin*dIII genomic fragment which was purified and sub-cloned to enable further characterisation.

4. The 1.3-kb *Hin*dIII genomic fragment was oligolabelled and used to probe a Southern blot containing DNA derived from somatic cell hybrids with various portions of the X-chromosome (Xpter-Xq26 and Xq26-Xqter). The probe hybridized to the DNA derived from the correct region known to contain the IDS gene.

5. The 1.3-kb *Hin*dIII fragment was labelled with tritium and used to probe metaphase chromosomes. This experiment confirmed the result obtained from

the Southern of somatic cell hybrids and also indicated that the genomic fragment derived from the X-chromosome clone was present only on the distal portion of the X-chromosome. Again, this was consistent with what was known about the IDS gene.

6. The 1.3-kb *Hin*dIII genomic fragment was sequenced and used to screen a cDNA library. The DNA sequence obtained from the genomic fragment identified the 49-mer binding region which was co-linear with the 42-kDa N-terminal amino acid sequence. Multiple cDNA clones were isolated with the 1.3-kb fragment and the largest ot these clones was found to contain a 1.5-kb *Eco*RI insert.

7. The 1.5-kb cDNA clone was sequence and revealed the possibility of the presence of another polypeptide species. A 14-kDa polypeptide species observed in the purified preparation of IDS was N-terminal amino acid sequenced and this sequence was found to be co-linear with sequence near the 3'-end of the 1.5-kb cDNA clone. No termination codon was observed, indicating that the 1.5-kb cDNA clone was incomplete. A restriction fragment from the 3'-end of the 1.5-kb cDNA clone was used to screen an endothelial cDNA library. A 2.3-kb cDNA clone was identified which was sequenced and found to contain the complete coding region of the potential IDS gene. The 1.5-kb cDNA clone was oligolabelled and used to probe a Northern blot. The probe detected major RNA species of 5.7, 5.4, 2.1 and a minor species of 1.4-kb in human placental RNA.

8. A cDNA fragment derived from the 2.3-kb cDNA clone was subcloned into the expression vector, pRSVN.07. The expressed product was observed to degrade specific IDS substrates, demonstrating that the cDNA encoded IDS and was full-length.

9. The IDS sequence has strong homology with other sulphatases (such as sea urchin arylsulphatase, human arylsulphatases A, B and C, and human glucosamine-6-sulphatase), suggesting that the sulphatases comprise an evolutionarily related family of genes that arose by gene duplication and divergent evolution.

10. The 1.5-kb cDNA fragment was oligolabelled and used to probe Southern blots containing DNA derived from 23 unrelated MPS II patients. Seven individuals had structural alterations of the IDS gene, including 2 with entire deletions and 5 with partial deletions or rearrangements. All the patients who demonstrated alterations of the IDS gene were clinically severe. The patients with gross deletions of the entire IDS gene were studied in detail and their deletions were found to extend beyond the IDS locus. It is unknown whether the extent of these deletions encompass other genes, however the patients were clinically severely affected. The DNA derived from a MPS II female patient with an X:5 balanced translocation was probed with the 1.5-kb cDNA and it was demonstrated that the translocation did in fact disrupt the IDS gene.

11. The Southern blots of MPS II patients revealed that the majority of patients had mutations that were beyond the detection limits of this procedure. It therefore became an objective of this study to determine the molecular nature of some of these MPS II mutations. Polymerase chain reaction was used to amplify the coding region of the IDS gene from reverse transcribed mRNA isolated from patient skin fibroblasts. Chemical cleavage of mismatched heteroduplexes between patient and normal control DNA revealed many unique cleavage patterns. Direct PCR sequencing was used to characterize the nature of the mutation in six unrelated MPS II patients. Three of these mutations were single base pair (bp) substitions, i.e. patient SF1779 had a C to G change (Arg_{469} to Gly), patient SF2069 had an A to C change (Lys_{228} to Gln) and patient SF811 had a C to T change (Glu_{375} to nonsense). Patient SF532

was found to have a 1 bp deletion within Leu_{15}, and patient SF3362 was found to have a 7 bp insertion (GGAACTA) between His_{335} and Gly_{336}. Both of these mutations result in a translation frameshift, predicting truncated IDS proteins. The mutation of a Glu_{375} codon to a nonsense codon in patient SF811 also predicts a truncated IDS protein. Patient SF635 was found to have a 60 bp deletion from his cDNA sequence, however, the sequences surrounding this mutation are suggestive of a mRNA splicing junction, indicating the need for the intron/exon structure of the IDS gene to determine the exact nature of this mutation. The involvement of the 2 amino acid substitutions in MPS II is unknown. Some of the MPS II patients screened with the chemical cleavage method revealed normal patterns, indicating the need to refine this technique to ensure 100 % mutation detection.

STATEMENT

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university. To the best of my knowledge it contains no material that has been previously published by any other person, except where due reference is made in the text. I consent to this thesis being made available for photocopying and loan.

Peter J. Wilson

ACKNOWLEDGEMENTS

I am extremely grateful to my supervisor, Dr. C. Phillip Morris for his careful guidance through the course of the work in this thesis. It has been an enjoyable and enlightening experience. I would also like to extend my gratitude to my other supervisor, Professor John J. Hopwood for his advice, criticisms and for encouraging me to undertake the research described in this thesis.

I also wish to thank my colleagues in the molecular biology section for their advice, assistance and for providing an enjoyable working environment: Dr. Don Anson (IDS DNA analysis and advice for IDS expression), Dan Robertson, Tom Litjens (setting-up the chemical cleavage method), Teresa Occhiodoro (collaboration with IDS sequence collection and Northern analysis), Annette Orsborn (oligonucleotide synthesis), Hamish Scott, Cathy Meany, Jill Durrant, and Xiao-Hui Guo. I would also like to extend my thanks to Dr. Greg Harper for providing me with many stimulating discussions on other aspects of the lysosome.

I would like to thank Julie Bielicki for providing me with the N-terminal amino acid sequence without which this project would not have commenced. Also I thank her for the IDS enzyme assays.

I extend my gratitude to Sophie Lazenkas for her help and her time in printing and formatting this thesis. Thank you.

I am also appreciative of the collaborative research effort with the Department of Cytogenetics and Molecular Genetics, and in particular Dr. Graeme Suthers for the work on the female MPS II patient. I would also like to acknowledge Dr. Ed Wraith for supplying us with MPS II DNA samples which were used to confirm the identity of the cDNA clone. Thanks also to Paul Nelson for the collaboration with the Southern analysis of MPS II patient DNA samples.

I was supported during the course of this work by an Adelaide Childrens Hospital Scholarship in Molecular Biology. The work was supported by grants from the National Health and Medical Research Council and the Adelaide Childrens Hospital Research Fund. The work described in this thesis was carried out at the Adelaide Childrens Hospital and the Degree was administered by the Department of Pathology, School of Medicine, University of Adelaide. I am also grateful to Professor Barrie Vernon-Roberts for enabling me to carry out this degree through the Department of Pathology.

CHAPTER ONE

INTRODUCTION

1.1 PRELIMINARY COMMENTS

Hunter syndrome (mucopolysaccharidosis type II; MPS II) is one of a number of lysosomal storage diseases that are known as the mucopolysaccharidoses (MPS). This group of lysosomal storage disorders are characterised primarily on the basis of their clinical phenotype and the type of glycosaminoglycan (GAG) excreted in the urine. The study of the MPS has been the subject of intensive research over the last thirty years. Although each individual disorder is relatively rare in the human population, the study of each disorder has improved our understanding of the disease state and has given insights into the biochemistry of GAG degradation.

The aim of the work presented in this thesis was to isolate and characterise the iduronate-2-sulphatase (IDS) gene, the gene known to be defective in MPS II patients. With the primary structure of the IDS gene product known, efforts could be directed towards the elucidation of the defects in the IDS gene. This work provides a foundation for future work which may result in the eventual treatment or cure of MPS II patients.

This introduction is intended to provide a clinical, biochemical and genetic background to MPS II. A brief explanation of how the IDS enzyme is synthesised and modified, and how it is transported to the lysosome will be included. Also a brief description of how the other MPS enzymes interact with IDS to degrade the same substrate will be included. It is not intended to review what is currently known about the lysosome or to provide detail of the molecular genetics of other lysosomal storage diseases.

The primary role of the lysosome is the degradation and recycling of cellular constituents. Over recent years the number of known lysosomal enzymes has increased to more than fifty and in most cases a dysfunctional enzyme results in a pathophysiological condition. The macromolecular substrates for these enzymes are transported to the lysosome from both the interior and exterior of the cell, where the substrates are sequentially degraded to monomers. These monomers are then transported out of the lysosome to either the biosynthetic machinery of the cell or are exported to the extra-cellular medium. When any one of these enzymes is defective a build-up of partially degraded substrates occurs which are unable to be transported from the lysosome. This causes the lysosome to swell until it can no longer function efficiently and eventually the cytosol is overrun with lysosomes, some of which are exocytosed, releasing partially degraded substrates. This along with the disruption to the cytosolic environment is thought to be the likely cause of pathology in lysosomal storage diseases.

In recent years many of the enzymes and genes responsible for the degradation of the mucopolysaccharides, one of the many complex substrates degraded in the lysosome, have been isolated. This provides an excellent opportunity for the study of structure and function relationships between these lysosomal enzymes and provides a detailed knowledge of one particular subset of lysosomal enzymes. One approach to understanding this relationship is to study the mutations that affect enzyme function and lead to patient pathology.

To date a very small number of mutations have been characterised in MPS I, MPS VI and MPS VII therefore allowing limited analysis of the structure and function relationships of these enzymes to be made. Scott *et al* (1992a) demonstrated a premature stop codon in 31% of MPS I alleles studied which results in a severe clinical phenotype. They also demonstrated 2 other mutant alleles that resulted in a base substitution (3%) and another premature stop codon (15%), these mutations also result in a severe clinical phenotype (Scott *et al.*, 1992b). The nature of these mutations, i.e. 2

nonsense and 1 missense, reveal preliminary information concerning the importance of certain regions of the protein but it is difficult to make speculation to exactly how these mutations affect the enzyme in terms of structural or functional changes. With further information on the levels of protein expressed and measurement of the specific activity of the mutant enzyme it may be possible to determine regions of the protein that are both important for substrate binding and catalysis.

Mutations have also been identified in MPS VI, including a single base pair substitution in the ASB gene which is in a conserved area of the sulphatases (Wicker *et al.*, 1991) that results in an intermediate clinical phenotype, the mutation causes a greatly reduced level of protein expressed but the low level of protein has significant enzyme activity which may explain the clinical phenotype. Jin *et al.*, (1991) reported mutations which resulted in both mild and severe clinical phenotypes. These results indicate genetic heterogeneity and help to identify regions of the ASB gene that are both critical and of less importance to overall enzyme function. In MPS VII mutations have also been identified which occur in conserved regions of the β-glucuronidase enzyme (Tomatsu *et al.*, 1991). However, whether they are directly involved in specific binding or catalytic interactions or are essential for maintaining protein integrity is unknown at present. They do however, provide a basis for further studies including site-directed mutagenesis which may clarify the regions involved in both catalysis and substrate binding.

Scott, H.S., Litjens, T., Hopwood, J.J. and Morris, C.P. (1992a). A common mutation for Mucopolysaccharidosis type I associated with a severe Hurler syndrome phenotype. *Hum. Mut.* 1: 103-108.

Scott, H.S., Litjens, T., Nelson, P.V., Brooks, D.A., Hopwood, J.J. and Morris, C.P. (1992b). α-L-Iduronidase mutations (Q 70X and P 533R) associate with a severe Hurler phenotype. *Hum. Mut.* 1: 333-339.

Wicker, G., Prill, V., Brooks, D., Gibson, G., Hopwood, J.J., von Figura, K. and Peters, C. (1991). MPS VI: An intermediate clinical phenotype caused by substitution of valine for glycine at position 137 of ASB. *J. Biol. Chem.* 266: 21386-21391.

Jin, W-D., Jackson, C.E., Desnick, R.J. and Schuchman, E.H. (1992). MPS VI: Identification of three mutations in the ASB gene of patients with the severe and mild phenotypes provides molecular evidence for genetic heterogeneity. *Am. J. Hum. Genet.* 50: 795-800.

Tomatsu, S., Fukuda, S., Sukegawa, K., Ikedo, Y., Yamada, S., Yamada, Y., Sasaki, T. *et al.* (1991). MPS VII: Characterisation of mutations and molecular heterogeneity. *Am. J. Hum. Genet.* 48: 89-96.

1.2 THE MUCOPOLYSACCHARIDOSES

The MPS are a family of heritable disorders caused by deficiency of lysosomal enzymes needed to degrade mucopolysaccharides or GAG. The undegraded or partially degraded GAG are stored in lysosomes (Figure 1.1) and excreted in the urine. Each of the known MPS involves the deficiency of one of 10 enzymes that are required for the step-wise degradation of the GAG dermatan sulphate, heparan sulphate, or keratan sulphate, individually or in combination. In some cases the degradation of chondroitin sulphate may also be affected. These disorders are chronic and progressive, and often display a wide spectrum of clinical severity within one enzyme deficiency (Neufeld and Meunzer, 1989). The following sections will discuss in detail the characterisation of MPS II, one of the MPS.

1.3 MPS II : AN HISTORICAL PERSPECTIVE

In retrospect the first definitive report believed to describe the MPS were by Hunter (1917) who described two brothers now presumed to have had the X-linked form of mucopolysaccharidosis, now referred to as Hunter syndrome or MPS II. The second MPS reported was by Hurler (1919), who described two patients presumed to have the autosomal recessive MPS, now termed Hurler syndrome or MPS I. MPS II was confirmed to be an X-linked recessive by Nja (1946) who used extensive analysis of MPS II family pedigrees to demonstrate an X-linked recessive pattern of inheritance. This work plainly delineated MPS II from the autosomal recessive MPS, MPS I.

Originally, MPS I and MPS II were classified together under the term 'gargolylism', obviously an unsatisfactory term for patient description. The general term 'lipochondrodystrophy' was mistakenly introduced in the 1930s as, at first, it was

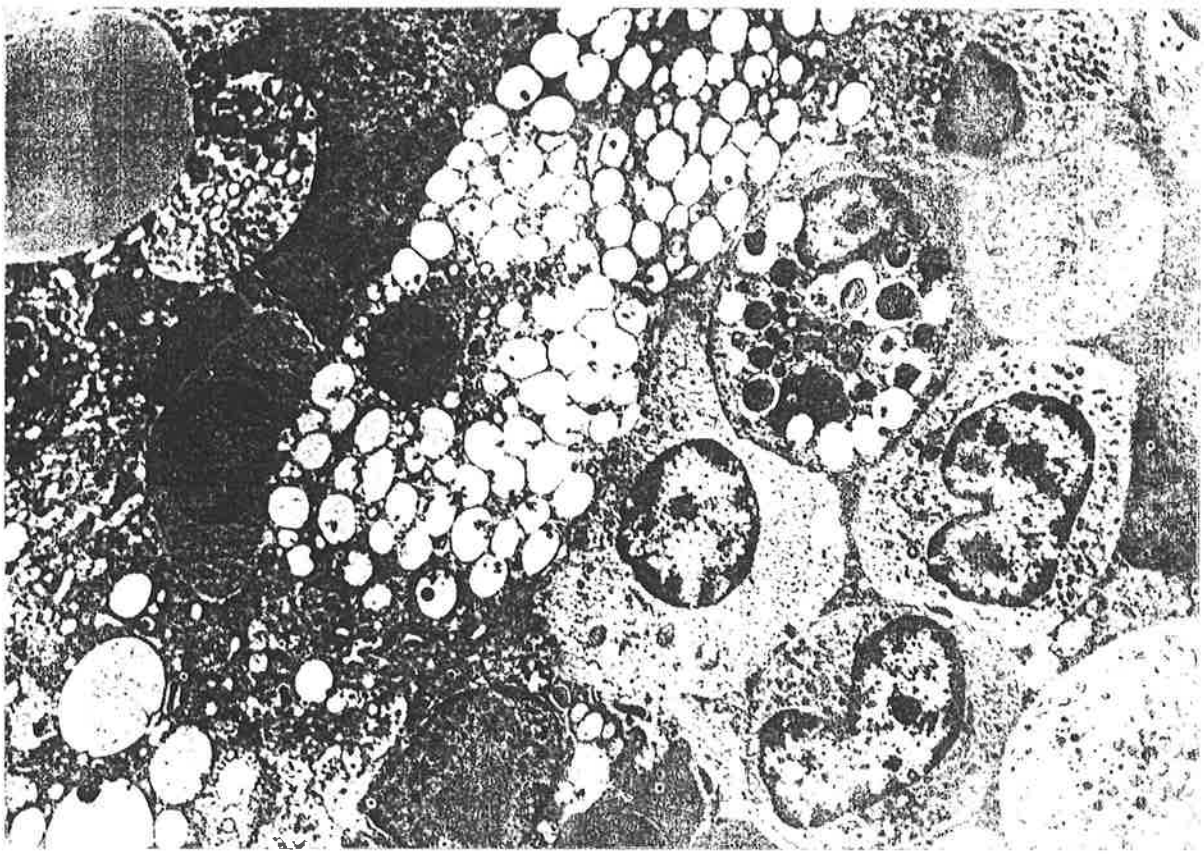
believed the storage material was lipid in nature. It was not until Brante (1952) that the term MPS was introduced when he demonstrated by metachromic staining of the inclusion material in tissues fixed in basic lead acetate, that the stored material was in fact an acid mucopolysaccharide. Dorfman and Lorincz (1957) reported the increased urinary excretion of mucopolysaccharides, which is still the standard test used when an MPS is suspected. Van Hoof and Hers (1964) were the first to demonstrate large vacuoles containing granular material in the liver of a MPS I patient and suggested that the partially degraded mucopolysaccharides were distending the lysosome (Figure 1.1).

In 1966, McKusick (1966) systematised the classification of the MPS into six eponymic types based on the types of mucopolysaccharides excreted in the urine, the mode of inheritance and clinical features. In 1968, Fratantoni *et al.* (1968) demonstrated that fibroblasts from MPS I and MPS II patients when compared to normal controls accumulate excessive amounts of radioactive mucopolysaccharide when incubated in medium containing radiolabelled sulphate. They showed that when the radiolabel is removed from the medium the MPS I and MPS II patient cells were only able to remove the stored material at a reduced rate when compared to normal control fibroblasts. Fratantoni *et al* (1969) were able to demonstrate that the abnormal catabolism of mucopolysaccharides in cultured fibroblasts from MPS I and MPS II patients could be corrected if cells of these two types were co-cultivated with each other or with normal control fibroblasts. They determined that the 'corrective factor' was excreted into the culture medium.

The next stage in the understanding of these diseases was the definition of specific degradative lysosomal enzymes. The enzyme that was found to be deficient in MPS II patients was determined independently by Bach *et al.* (1973), Coppa *et al.* (1973) and

Figure 1.1: **LYSOSOMAL INCLUSIONS**

Bone marrow cells from a cat affected with Maroteaux-Lamy syndrome (MPS VI) (Diagram provided by Prof. John Hopwood).



Sjoberg *et al.* (1973) and was named L-idurono-sulphate sulphatase but is now known as iduronate-2-sulphatase or IDS. Once the enzyme deficiency was identified an intensive effort was made to understand the biochemistry and pathological detail for each disorder which led to the eventual purification of each MPS enzyme and many of the genes (for a review see, Hopwood and Morris, 1990).

1.4 BIOCHEMISTRY OF MPS II

1.4.1 BIOSYNTHESIS AND PROCESSING OF IDS

The pathway of biosynthesis, receptor-mediated transport and post-translational processing of lysosomal proteins is well understood (for reviews see, von Figura and Hasilik, 1986; von Figura *et al.*, 1987; Kornfeld and Mellman, 1989). The lysosomal enzymes are synthesised with an N-terminal endoplasmic reticulum (ER) signal peptide that is recognised by a signal recognition particle as it emerges from the ribosome. After the binding of this complex to a membrane receptor, the lysosomal protein is translocated across the membrane of the ER into the lumen (Erickson *et al.*, 1983). Once the polypeptide enters the ER, pre-formed oligosaccharides are transferred from dolicol carriers to some of the asparagine residues which are part of the protein consensus N-glycosylation amino acid sequence Asn-X-Thr/Ser (see Figure 1.2, compartment A; and also Chapter 4, Section 4.3.4). These oligosaccharides are subjected to extensive modification, including phosphorylation at C-6 of some of the mannose residues when the polypeptide reaches the Golgi complex (Figure 1.2, compartment A; Kornfeld and Kornfeld, 1985). This phosphorylation prevents further processing of the phosphorylated mannose residues, but other branches of the same oligosaccharide may be trimmed or further modified, for example, the addition of sialic acid residues and/or sulphate (Hasilik and von Figura, 1981; Braulke *et al.*,

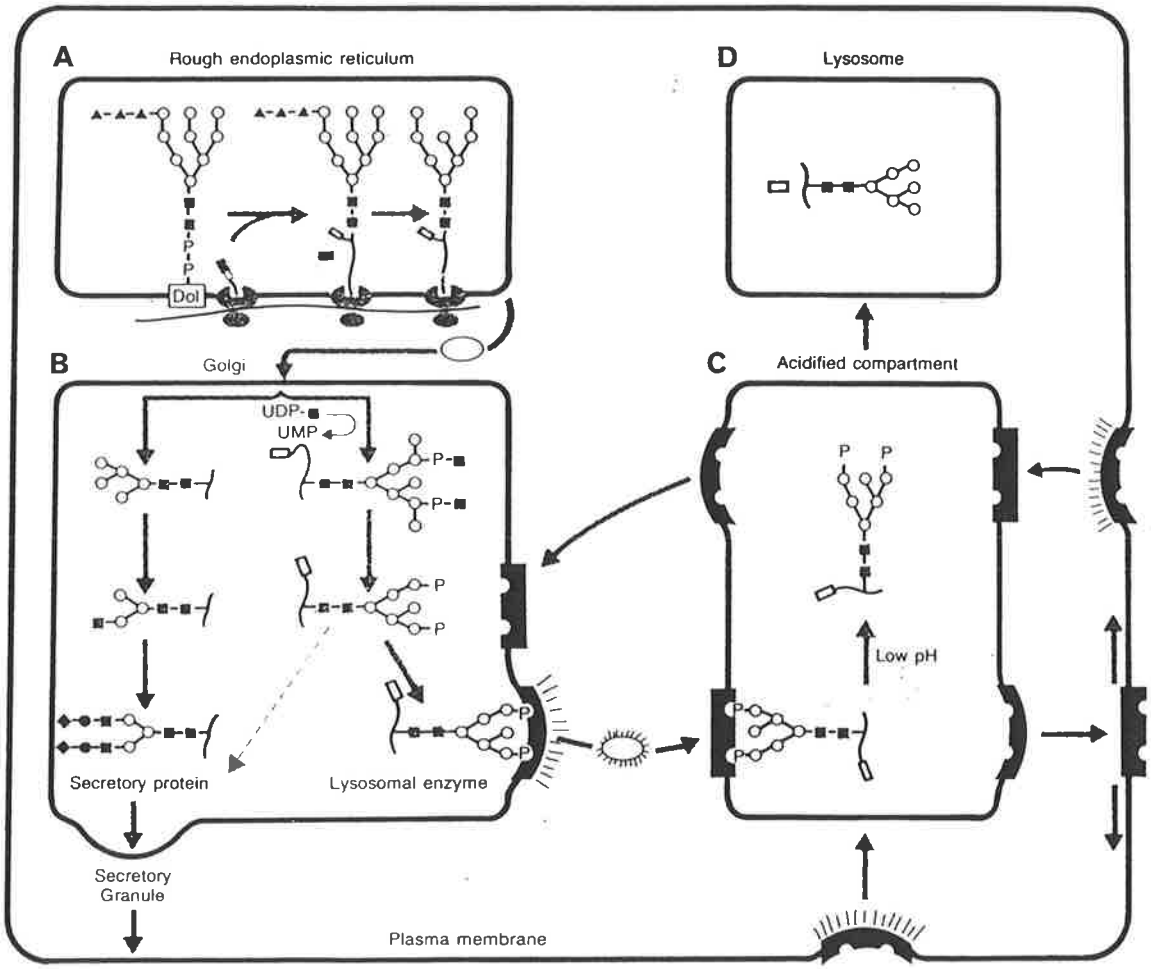
Figure 1.2: BIOSYNTHESIS AND TRANSPORT OF LYSOSOMAL ENZYMES (diagram from Neufeld and Meunzer, 1989).

A: In this compartment newly synthesised proteins enter the rough endoplasmic reticulum (ER) were N-linked complex oligosaccharides are transferred from dolichol carriers, the ER signal peptide is cleaved and some processing of the oligosaccharide occurs.

B: The enzymes are actively transported to the Golgi where further oligosaaccharide processing and some of the mannose groups are phosphorylated. The enzymes that are modified in such a way interact with a specific mannose-6-phosphate (M6P) receptor.

C. The enzyme attached to the M6P receptor is actively transported to the acidic endosome where the M6P receptor dissociates from the lysosomal enzyme and recycles back to the Golgi to pick up other enzymes with M6P groups.

D: The enzyme is transported to the lysosome where it is functional toward its specific substrate.


A
Rough endoplasmic reticulum
Dol
D
Lysosome
B
Golgi
UDP-
UMP
P
P
P
P
P
P
Secretory protein
Lysosomal enzyme
C
Acidified compartment
P
P
Low pH
P
P
Secretory
Granule
Plasma membrane

1987). Mannose-6-phosphate (M6P) residues function as recognition signals for receptors that allow the targeting of enzymes to the lysosome rather than the secretory pathway. The two types of M6P receptor (cation-dependent and -independent) occur throughout the Gogi complex and the trans-Golgi network. The cation-independent M6P receptor-lysosomal enzyme complexes are packaged into clathrin-coated vesicles and enter the endocytic route to the lysosomes via endosomal and other prelysosomal compartments (Lemansky *et al.*, 1987). The acidic environment within these vesicles (Figure 1.2 compartment C) dissociates the receptor-protein complex and the receptors are recycled back to the Golgi complex. This process may be leaky and a fraction of the newly synthesised lysosomal protein follows the pathway for secretory proteins (Figure 1.2, compartment B). These secreted lysosomal enzymes may then interact with M6P receptors found on the plasma membrane in clathrin-coated pits. When the lysosomal enzyme associates with the receptor the coated pit is endocytosed and the enzymes follow the endocytic pathway to the lysosome. Once the lysosomal enzymes reach the lysosome they become active in the acidic environment and begin to degrade their specific substrates (Figure 1.2, compartment D).

After the removal of the ER signal peptide, the newly synthesized lysosomal enzymes may be subjected to further limited proteolytic processing which includes removal of N- and C-terminal sequences and intra-chain cleavages. The products of the cleavage may remain associated with each other through disulphide bonds or non-covalent forces. The function of the limited proteolysis is unknown. It has been shown that most lysosomal precursors have normal enzyme activity, for example secreted β-N-acetylhexosaminidase A (Hasilik *et al.*, 1982). Additional support for this argument was obtained when it was demonstrated that secreted recombinant IDS was shown to be catalytically active towards specific IDS substrates (Section 4.3.7b). Lysosomal enzymes which are not targeted to the lysosome are secreted instead as higher M_r

precursor forms. Since the processing of lysosomal enzymes may take several days during transport to the lysosome and several weeks within the lysosome, it is perhaps not surprising that multiple forms of lysosomal enzymes (which differ in their native protein M_r and subunit compositions) have been isolated from the different tissue sources used in their purification (see the following section).

1.4.2 PURIFICATION OF IDS

The purification of IDS has been an extremely difficult task and has been the goal of several laboratories over the last two decades. Cantz *et al.* (1972) were the first to report a 120-fold purification of IDS (the 'Hunter corrective factor') from human urine, with an estimated molecular weight of 65-kDa from SDS-PAGE. Isoelectric focusing revealed several polypeptide bands. However, they were unable to recover Hunter corrective factor from the gel and could not determine which polypeptide band was reponsible for activity, indicating that they did not have a homogenous preparation of the 'Hunter corrective factor'. Wasteson and Neufeld (1982) reported a 50,000-fold purification of the IDS enzyme from human plasma that on SDS-PAGE revealed a single polypeptide of 80-kDa. Lissens *et al.* (1984) reported a 155-fold purification of IDS with two forms from human urine. One of the forms they studied had a calculated molecular weight of 66-kDa from SDS-PAGE run under reducing conditions. Di Natale and Danielle (1985) reported the presence of two forms of IDS in human placenta after a 30,000-fold purification that in SDS-PAGE run under reducing conditions revealed a single polypeptide species between 80 and 90-kDa. Western blot analysis of various IDS preparations demonstrated subunit molecular masses of 72-kDa for placenta, 60-kDa for fibroblasts and two polypeptides in serum of 75 and 83-kDa (Daniele and Di Natale, 1987). No amino acid sequence was reported from any of these studies.

The purification of the MPS enzymes has proved to be a singularly difficult process due to their low abundance and low pH optima, and the problems associated with obtaining a homogeneous protein preparation. For example, in the case of α-L-iduronidase (MPS I), seven polypeptides were identified after a 170,000 fold purification (Clements *et al.* 1989), this is likely due to the aggressive proteolytic environment of the lysosome which results in the removal of protease susceptible residues and the appearance of multiple polypeptide fragments of the primary gene product. This makes the identification of suitable probes for gene isolation difficult as there is uncertainty as to which polypeptides are derived from the enzyme of interest. The enzyme N-acetylgalactosamine-4-sulphatase (MPS VI) was purified to homogeneity and revealed a 43 kDa and 8 kDa polypeptide linked by a disulphide bond (Gibson *et al.* 1987).

The enzymes that are involved in the degradation of heparan and dermatan sulphate show absolute specificity towards their natural substrates. Including the 7 sulphatases that are involved in this process. However, some sulphatases, the arylsulphatases for example exhibit the capacity to recognize and cleave certain artificial sulphated substrates, e.g. 4-methylumbelliferyl sulphate, whereas the other sulphatases in this group will not cleave the sulphate residue of this substrate. This functional simularity is underscored by protein sequence comparison which reveals that the arylsulphatases exhibit greater sequence conservation, implying that they are evolutionarily more closely related then the non-arylsulphatases.

Bielicki *et al.* (1990) reported a 500,000-fold purification of IDS from human liver. Chromatofocusing led to the separation of two major forms (A and B) from human liver, lung, kidney and placenta. Form A and B from human liver were shown to consistently contain a 42-kDa polypeptide species. The N-terminal amino acid sequence of the 42-kDa polypeptide was determined and was found to have homology with the N-terminal amino acid sequence of other sulphatases (see Chapter 3 and Chapter 4, Section 4.3.6). The differing molecular sizes obtained for purified human IDS by the various investigators can most likely be attributed to the different conditions used in each experiment. For example, the behaviour of the IDS enzyme on gel-permeation chromatography was dependent upon the nature of the gel matrix and the pH and ionic strength of the eluting buffer (Bielicki *et al.*, 1990). Under conditions of low ionic strength and neutral pH, IDS was eluted with a higher molecular weight, suggesting aggregation of the enzyme (Bielicki *et al.*, 1990). This phenomena has also been reported for the lysosomal enzyme α-L-iduronidase (Clements *et al.*, 1985). The low abundance and low recovery of the IDS enzyme also makes comparison difficult because there may be selective removal of some forms of IDS (A or B) (Bielicki *et al.*, 1990).

1.4.3 HEPARAN SULPHATE DEGRADATION

IDS (EC 3.1.6.13) is a lysosomal exosulphatase that hydrolyzes the C2-sulphate ester bond from non-reducing-terminal iduronic acid residues of the GAG heparan sulphate (HS) and dermatan sulphate (DS). The form A preparation of IDS was found to have an acidic pH optimum and hydrolyzed both HS and DS derived oligosaccharides (Bielicki *et al.*, 1990).

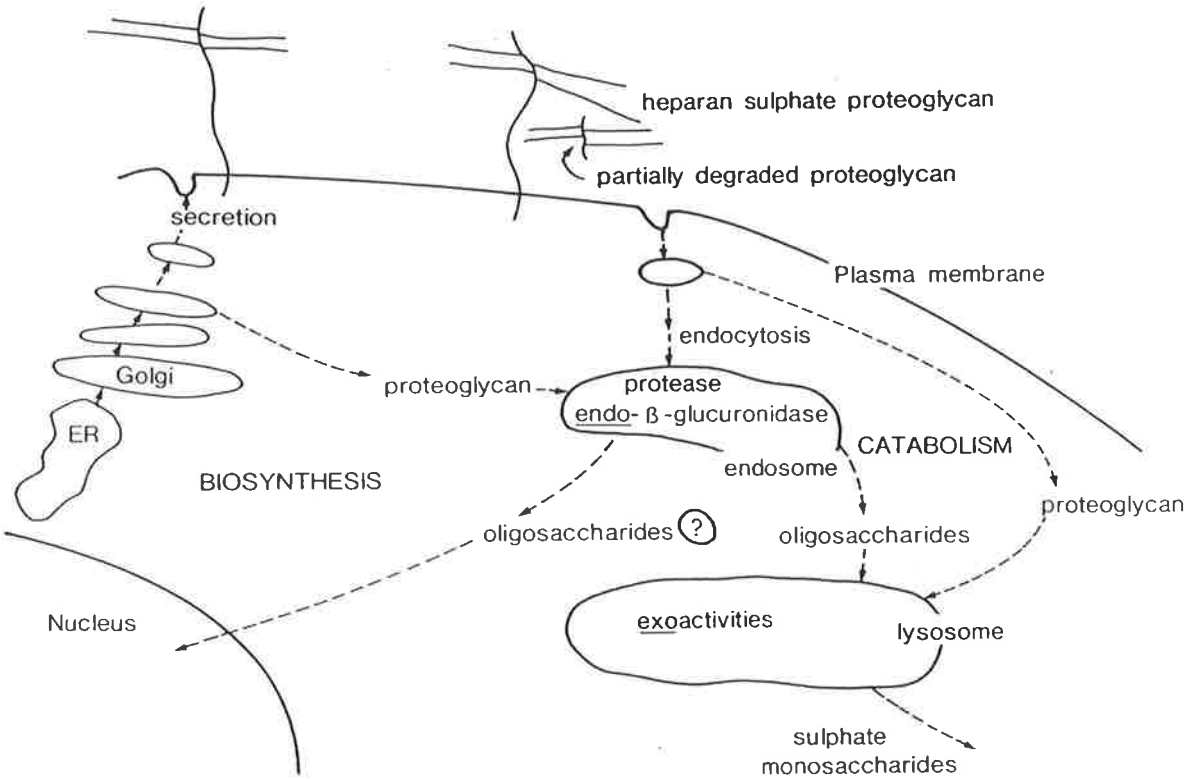
The degradation of GAG is a sequential process involving a number of other lysosomal enzymes. All of these lysosomal enzymes when deficient lead to lysosomal storage disorders, except glucuronate-2-sulphatase and glucosamine-3-sulphatase, where patients with these enzyme deficiencies have yet to be identified.

HS is synthesised as a proteoglycan with two or more HS chains covalently linked to a protein core, and is found to be associated with the extracellular matrix and plasma membrane (see Figure 1.3; Fransson, 1989). The exact function(s) of HS still remains speculative, but the control of cell adhesion, migration, differentiation, and proliferation have been implicated (Poole, 1982; Stigson and Kjellen, 1991). Figure 1.3 represents a schematic of the pathways of synthesis and catabolism of HS.

HS contains uronic acid residues, alternating with α-linked glucosamine residues (Roden, 1980; Fransson, 1989; reviewed in Hopwood, 1989). The uronic acid may be β-D-glucuronic acid or α-L-iduronic acid residues, which are sometimes sulphated and the glucosoamine residues may be N-sulphated or N-acetylated, 6-sulphated and 3-sulphated. Some HS proteoglycan is intergrated into the plasma membrane before being endocytosed and transported to the lysosome for degradation (Figure 1.3). It has been proposed that the first stage of HS degradation occurs in early endosomes, where proteolysis produces single HS chains of approximately 30-kDa (reviewed in Freeman and Hopwood, 1992), which are then hydrolyzed by endo-β-glucuronidase activities to oligosaccharides with a molecular weight of approximately 5-kDa (for review see Hopwood, 1989; Freeman and Hopwood, 1992). These HS oligosaccharides are transported to the lysosome to be processed from their non-reducing ends by the sequential action of up to three glycosidases, five sulphatases and an N-acetyltransferase (Figure 1.4). This degradation process is extremely rapid with the monosaccharides and inorganic sulphate, generated in this degradation process,

Figure 1.3: **BIOSYNTHESIS AND DEGRADATION OF HEPARAN SULPHATE**

Diagram (from Hopwood, 1989) represents a simplified scheme detailing the processes involved in the biosynthesis and catabolism of heparan sulphate (HS) The left side of the diagram represents the events involved in the biosynthesis of HS in the Golgi, secretion from the cell and insertion into the plasma membrane. The right side of the diagram represent the steps involved in the endocytosis and catabolism of HS.


heparan sulphate proteoglycan
partially degraded proteoglycan
secretion
Plasma membrane
endocytosis
Golgi
proteoglycan
protease
endo- ß -glucuronidase
ER
BIOSYNTHESIS
CATABOLISM
endosome
proteoglycan
oligosaccharides ?
oligosaccharides
Nucleus
exoactivities
lysosome
sulphate
monosaccharides

actively transported from the lysosome to possibly be re-used in the cells biosynthetic processes.

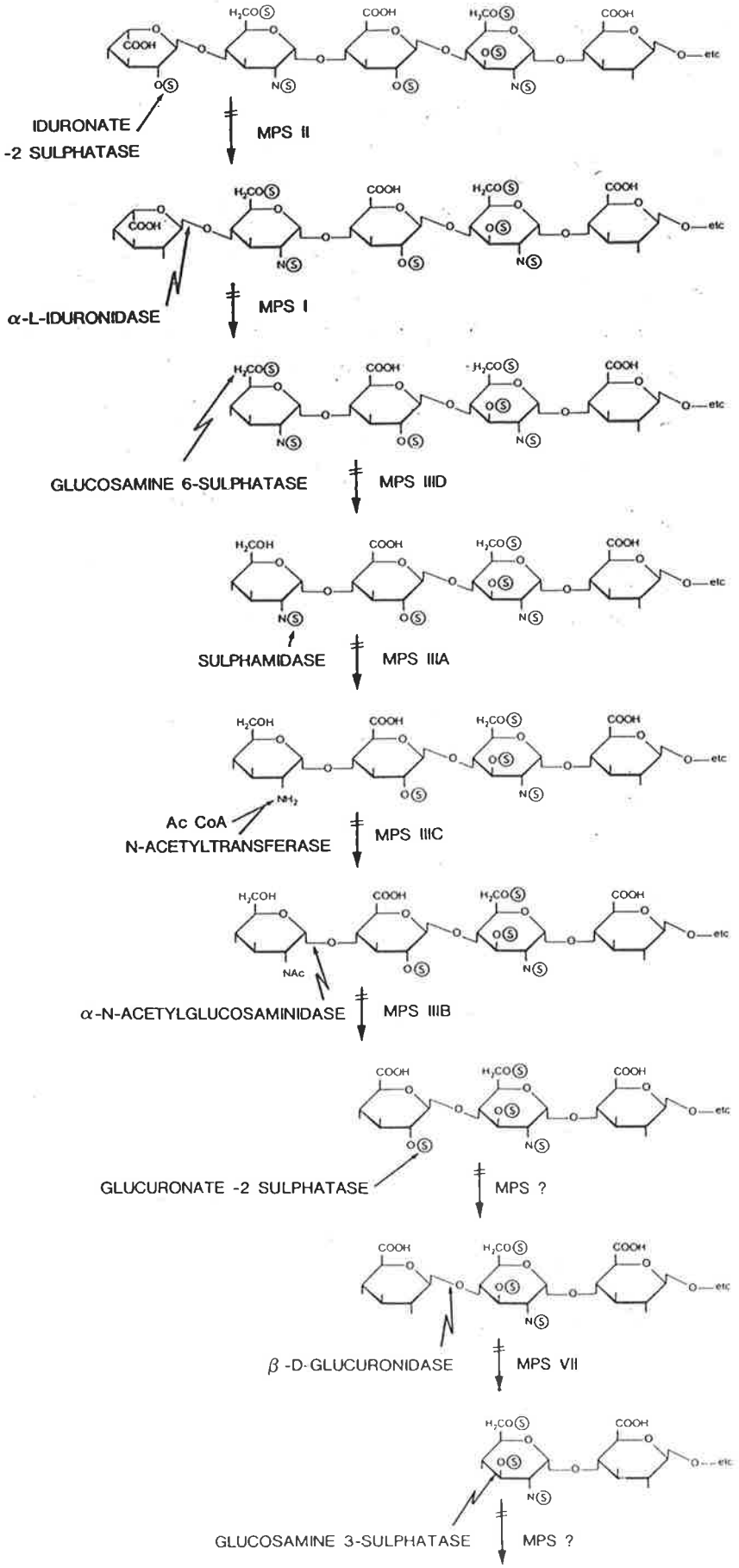
IDS, which when deficient results in MPS II, hydrolyses terminal iduronate-2-sulphate esters in HS (Figure 1.4). The next enzyme involved in the sequential degration of HS is α-L-iduronidase which leads to MPS I when deficient. α-L-Iduronidase acts on the oligosaccharide product of IDS to produce iduronic acid and the substrate for the next enzyme in the pathway glucosamine-6-sulphatase (Figure 1.4), which when deficient leads to MPS IIID. Sulphamidase, which is deficient in MPS IIIA, acts next on the oligosaccharide product of glucosamine-6-sulphatase.

After the de-N-sulphation of the non-reducing end of glucosamine-N-sulphate (GlcNS) residues by sulphamidase, a biosynthetic step is required to N-acetylate this GlcN product before degradation can proceed (Klein *et al.*, 1978). This reaction involves the transfer of the acetyl group from acetyl-CoA a substrate present in the cytoplasm to the GlcN substrate present in the lysosome. This enzyme is the only enzyme involved in the degradation of HS that is an integral membrane protein (Freeman *et al.*, 1983). This step in the process is unusual in that a bond is formed in an overall process aimed primarily at breaking bonds, and it has been found that this process has its counterpart in N-deacetylase a hydrolyse involved in heparin biosynthesis located in the Golgi (Lindahl *et al.*, 1986). A deficiency of N-acetyltransferase leads to MPS IIIC. The next enzyme to act, on the acetylated oligosaccharide product is α-N-acetylglucosamidase which when deficient results in MPS IIIB.

Glucuronate-2-sulphatase is shown as the next enzyme in this degradative pathway and is required for the de-sulphation of non-reducing end glucuronate-2-sulphate residues (Figure 1.4). No patients have been described with a deficiency of this enzyme, but the

Figure 1.4: **DEGRADATION OF HEPARAN SULPHATE**

Schematic representation of the step-wise degradation of HS oligosaccharides in the lysosome. The diagram shows all the bonds that are modified in HS by the action of 9 exoenzymes. The positions of the glucuronate-2-sulphate and glucosamine-3-sulphate serve only as a guide.


IDURONATE
-2 SULPHATASE
MPS II
α-L-IDURONIDASE
MPS I
GLUCOSAMINE 6-SULPHATASE
MPS IIID
SULPHAMIDASE
MPS IIIA
Ac CoA
N-ACETYLTRANSFERASE
MPS IIIC
α-N-ACETYLGLUCOSAMINIDASE
MPS IIIB
GLUCURONATE -2 SULPHATASE
MPS ?
β -D-GLUCURONIDASE
MPS VII
GLUCOSAMINE 3-SULPHATASE
MPS ?

clinical phenotype may be expected to be similar to MPS III (A-D) (Freeman and Hopwood, 1991). The next enzyme shown, β-D-glucuronidase which when deficient leads to MPS VII, is found in both the lysosome and microsomes. Glucosamine-3-sulphatase desulphates the product from β-D-glucuronidase, namely glucosamine-3-sulphate residues, again patients with a deficiency of this enzyme have yet to be reported (Figure 1.4).

1.4.4 DEGRADATION OF DERMATAN SULPHATE

IDS is also required for the degradation of DS. DS consists of α-L-iduronic acid residues with some being C-2-sulphated, and β-D-glucuronic acid residues, alternating with β-linked N-acetylgalactosamine residues that may be C-4 or C-6-sulphated (for a review see Roden, 1980). DS is synthesised as a proteoglycan and is mainly found in the matrix of connective tissues. Proteolysis of endocytosed DS proteoglycan produces single DS chains which may be reduced to DS oligosaccharides by the action of endoglycosidase activities present in the endocytic pathway to the lysosome. Lysosomal degradation of the DS oligosaccharides occurs in a step-wise manner with the enzymes shown in Figure 1.5.

IDS is required for the desulphation of the C-2-sulphated iduronic acid to produce the substrate for α-L-iduronidase (Figure 1.5). When either of these enzymes is deficient both HS and DS accumulate in MPS I and MPS II patients. The action of α-L-iduronidase produces the substrate for N-acetylgalactosamine-4-sulphatase, which when deficient of this enzyme causes accumulation of DS in tissues and urine and results in MPS VI.

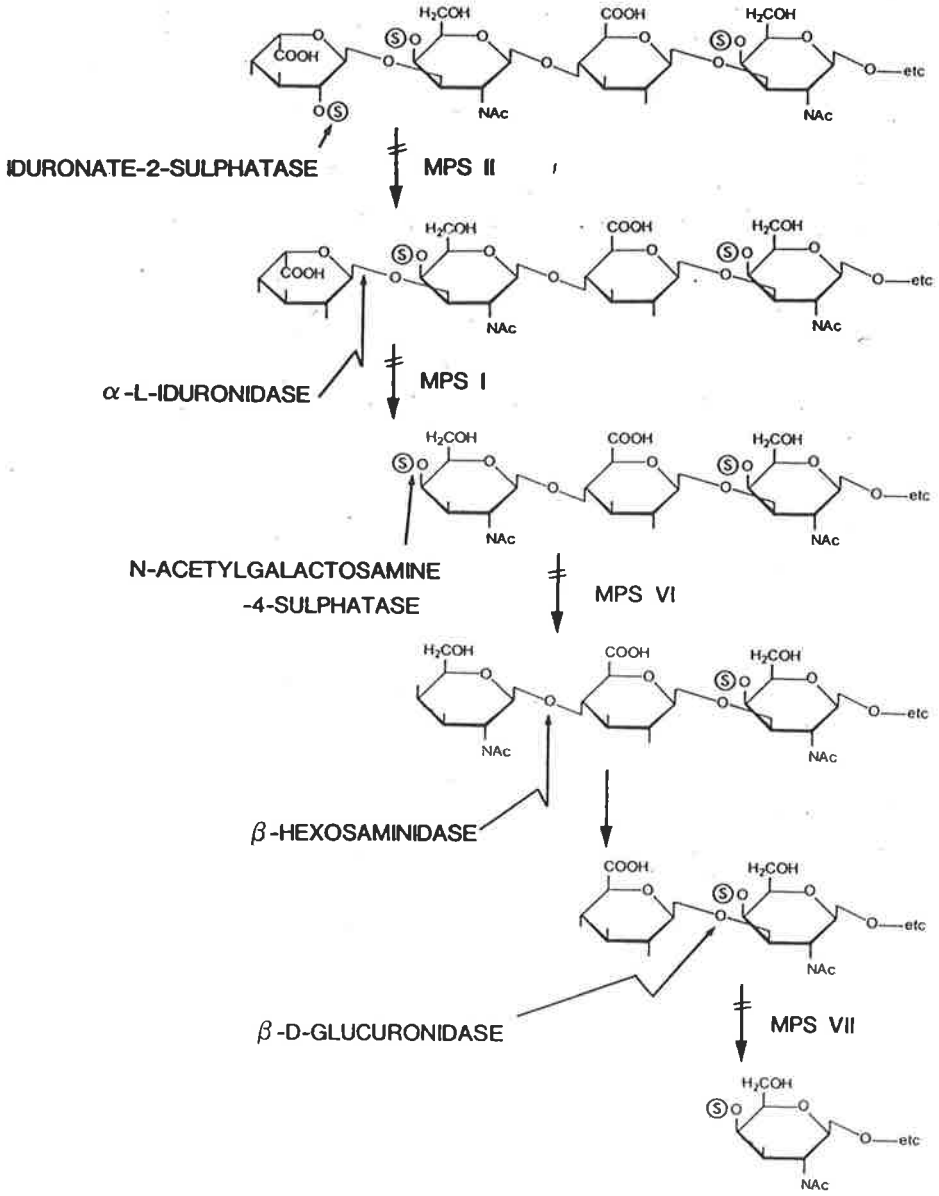
Dermatan sulphate proteoglycan is believed to play an important part in regulating collagen fibrillogenesis. In addition, it has been shown to bind (via the core protein) to transforming growth factor-β (TGF-β) which results in the inhibition of the growth stimulating activity of TGF-β (Yamaguchi *et al.* 1990). Dermatan sulphate proteoglycans have also been shown to bind fibronectin and prevent cell attachment which suggests that these molecules may be involved in mediating cell-substrata interactions (Rosenberg, 1992).

Yamaguchi, Y., Mann, D.M. and Ruoslahti, E. (1990). Negative regulation of transforming growth factor b by the proteoglycan decorin. *Nature*. 346: 281-284.

Rosenberg, L.C. (1992). Structure and function of dermatan sulphate proteoglycans in articular cartilage. In Articular Cartilage and Osteoarthritis, (eds. KE Kuettner, R. Schleyerbach, JG Peyron and VC Hascall), pp 45-63. Raven Press.

Figure 1.5: **DEGRADATION OF DERMATAN SULPHATE**

A scheme for the step-wise enzymatic degradation of DS.


IDURONATE-2-SULPHATASE
MPS II
α-L-IDURONIDASE
MPS I
N-ACETYLGALACTOSAMINE
-4-SULPHATASE
MPS VI
β-HEXOSAMINIDASE
β-D-GLUCURONIDASE
MPS VII
ALTERNATE PATHWAY
β-HEXOSAMINIDASE

The N-acetylgalactosaminide residue is thought to be hydrolysed by the action of the ß-hexosaminidase isoenzymes A or S (Figure 1.5). β-D-Glucuronidase is the final enzyme required for the degradation of DS (Figure 1.5). This enzyme is also required for HS degradation and as a result patients with a deficiency of this enzyme accumulate both HS and DS in their tissues (MPS VII). An alternative pathway for DS degradation has been proposed where β-hexosaminidase A or S is able to hydrolyse the β-linked N-acetylgalactosaminide-4-sulphate residue to produce N-acetylgalactosamine-4-sulphate (Hopwood and Elliott, 1985). This monosaccharide sulphate ester substrate can be hydrolysed by N-acetylgalactosamine-4-sulphatase to produce inorganic sulphate and N-acetylgalactosamine. The role this alternative pathway has in the normal degradation of DS is at present unknown.

1.5 GENETICS OF MPS II

1.5.1 INHERITANCE

MPS II is the only known MPS that is inherited as an X-linked recessive. Thus, MPS II is expected to occur only in hemizygous males, however there have been a number of reports identifying MPS II in females (Neufeld *et al.*, 1977; Mossman *et al.*, 1983; Broadhead *et al.*, 1986; Clarke *et al.*, 1990). One of the two females in Neufeld *et al.* (1977) was later identified as having multiple sulphatase deficiency, an autosomal recessive disorder, the basis for the disorder in the other MPS II female was undetermined. Broadhead *et al.* (1986) described a female MPS II patient in whom the symptoms resulted from inactivation of the normal X in a heterozygote. Mossman *et al.* (1983) reported an MPS II female patient with a balanced X-autosome translocation. The authors presumed that the translocation disrupted the IDS locus. This proposal was confirmed when an IDS cDNA was used to probe a Southern blot

containing DNA derived from this patient (see Chapter 5, Section 5.3.3). The translocation was also informative as to the location of the IDS gene, and for the first time placed the gene at Xq26-27. Upadhyaya *et al.* (1986) and Roberts *et al.* (1989) were able to refine the cytogenetic data placing IDS at Xq27-28. This localisation was used to aid the isolation of the IDS gene (see Chapter 3).

MPS II heterozygotes have two populations of cells, depending on whether the maternal (MPS II) or paternal (normal) X-chromosome is active. This is in agreement with the Lyon hypothesis (Lyon, 1962). Clones from single cells that have been derived from a heterozygote either display an MPS II phenotype or a normal phenotype as determined by their metachromasia (Danes and Bearn, 1967), by their radiolabelled mucopolysaccharide accumulation or by their specific enzyme deficiency (Capobianchi and Romeo, 1976; Migeon *et al.*, 1977). The normal clinical phenotype of heterozygotes can be explained by the presence of IDS in the normal cells which can diffuse into the mutant cells and correct the storage of GAG.

1.5.2 INCIDENCE OF MPS II

The MPS are rare genetic disorders. The incidence of MPS II was found to be 1 per 150,000 live births in British Columbia (Lowry and Renwick, 1971), 1 per 260,000 live births in Great Britain (Young and Harper, 1982) and 1 per 67,500 live births in Israel (Schaap and Bach, 1980). The incidence of MPS II in Australia was found to be close to 1 in 50,000 live births (Professor John Hopwood, personal communication). The Israeli figures possibly represent complete ascertainment, whereas the British and Canadian surveys may be incomplete. It was suggested that the higher incidence of MPS II in the Jewish population could be attributed to either genetic drift or founder effect (Chakravarti and Bale, 1983) or to selection of the MPS II allele in Ashkenazi

Jews (Zlotogora *et al.*, 1985). According to theory, since the reproductive fitness of X-linked diseases such as MPS II is close to zero, approximately one-third of cases will result from new mutations (proposed by Haldane, 1935) (see Chapter 5).

Young *et al.* (1982a,1982b) suggested that there were two distinct clinical phenotypes for MPS II, a severe and a mild clinical phenotype. However, McKusick and Neufeld (1983) suggested that the existence of two separate and distinct clinical phenotypes of MPS II is misleading, and that the available data indicated that a much broader and perhaps continuous clinical spectrum was more probable. There was no detectable difference in the enzyme levels between MPS II patients with or without severe mental retardation, since the residual activity of IDS was minimal in both cases (Liebaers and Neufeld, 1976). It has been postulated that the different clinical heterogeneity seen in this disorder reflects different mutations at the IDS locus (Yatziv *et al.*, 1977). This hypothesis is supported from the results obtained in this thesis (see Chapter 5).

1.6 CLINICAL PHENOTYPE OF MPS II

1.6.1 MPS II IS CHARACTERISED BY A BROAD CLINICAL SPECTRUM

Clinically, severely affected patients with MPS II are characterised by coarse facial features, short stature, skeletal deformities, joint stiffness and mental retardation (Neufeld and Muenzer, 1989). The first signs of the disease are usually detected between the ages of 2-4 years of age with progressive neurologic and somatic involvement (Young and Harper, 1983), whereas mildly affected MPS II patients have no or minimal neurological pathology and their somatic pathology takes a slower course (Young and Harper, 1982b). Unlike MPS I there is no corneal clouding

(Topping *et al.*, 1971), with one reported exception (Spranger *et al.*, 1978). Electroretinography indicates retinal dysfunction which varies according to the severity of the disease (Gills *et al.*, 1965; Caruso *et al.*, 1986). Chronic papilloedema has also been reported in mild and severe patients without raised intracranial pressure (Beck, 1983; Beck and Cole, 1984). It is possible this is due to deposition of GAG within the sclera, with compression of the optic nerve at the intrascleral level (Goldberg and Duke, 1967).

Chronic diarrhoea due to autonomic nervous system involvement and possibly mucosal dysfunction are problems faced in the more severe forms of MPS II. Persistent ear infections and gradual loss of hearing are also common features in both the severe and mild forms (Peck, 1984; Zechner and Moser, 1987). In the severe forms of the disease the central nervous system deterioration is much more marked, although it is generally not as serious as MPS I. There is often communicating hydrocephalus increasing the intracranial pressure and excacerbating the nervous system degeneration (McKusick, 1972; Ballanger *et al.*, 1979). Obstructive airway disease (Sasaki *et al.*, 1987; Ginzberg *et al.*, 1990) and cardiac failure due to valvular dysfunction, myocardial thickening, pulmonary hypertension, coronary artery narrowing and myocardial infarction (reviewed in Pyeritz, 1983) are generally responsible for the increased mortality in the severe form of MPS II. The milder forms of MPS II also suffer from breathing difficulties and cardiac failure, but often survive into adulthood (Young and Harper, 1982a,1982b).

1.6.2 DIAGNOSIS OF MPS II

Until recently the tests performed on MPS II patients were based on detection of specific GAG excreted in the urine. These tests are quick, cheap and serve as a useful

guide for preliminary evaluation, but have a high percentage of false-negative and false-positive results. The Adelaide Childrens Hospital screens all patients suspected for MPS conditions by testing for MPSuria. The method used is straight forward and does not have a false negative/positive rate (in Hopwood and Harrison, 1982). Definitive diagnosis of MPS II and the other MPS is established by specific enzyme assays (Hall *et al.*, 1978; Kresse *et al.*, 1982; Hopwood and Morris, 1990). In general, cultured fibroblasts are used but leucocytes, serum or plasma have been used for most MPS.

The level of IDS activity in fibroblasts from female carriers of the MPS II condition is complicated by cross-correction of MPS II cells by enzyme from normal cells *in vitro*, e.g. in the culture of skin fibroblasts derived from a potential MPS II heterozygote the cells with a normal IDS activity secrete IDS which is endocytosed by MPS II cells that corrects the MPS storage resulting in a population of cells that appear to be normal by the radiolabelled mucopolysaccharide accumulation test. Cloning of cells can reveal the MPS II cells and define the carrier status, however this procedure is costly, time consuming and technically demanding (Capobianchi and Romeo, 1976; Migeon *et al.*, 1977).

Another test that has been developed to detect carriers of MPS II and other X-linked diseases is the measurement of the activity of the defective enzyme in individual hair roots (Nwokoro *et al.*, 1977; Fluharty *et al.*, 1977; Harrison, 1982). This assay depends on the phenomenon of random inactivation of one of the pair of X-chromosomes present in every female cell (Lyon, 1962). If a female is a carrier of MPS II she will have both normal and mutant cells, each expressing either normal IDS or mutant IDS enzyme activity. The proliferation in hair roots of a few progenitor cells results in near-clonal populations in individual hair roots thus making it possible to detect the

MPS II cell population (Harrison, 1982). This method is very accurate but it is labor intensive and requires a large sample of hair roots for the assay to ensure the confidence of the result.

The development of recombinant DNA techniques may improve or simplify the methodology for detecting carriers of MPS II. The ability to determine mutations in IDS at the nucleotide level may also enable the correlation of patient phenotype with specific mutations. However, as it appears that the mutations resulting in MPS II are heterogeneous this correlation will be complicated (Wilson *et al.*, 1991; also see Chapter 5). If a correlation can be made a more accurate prognosis of the disease course may be possible. The nature of the mutations determined in MPS II patients may also influence the selection of patients for therapy protocols when they become available (Hopwood and Morris, 1990).

1.7 FUTURE AVENUES FOR PATIENT THERAPY

The discovery that the MPS were lysosomal storage diseases led to the hope that they may be treated by addition of the deficient enzyme, which would be endocytosed and transported to the lysosome via the M6P transport pathway (see Section 1.4.1) (Baudhuim *et al.*, 1964; Hers, 1965). When storage was shown to be corrected in cell culture by the addition of the missing enzyme to the culture medium this avenue of treatment was thought to be a genuine possibility (Fratantoni *et al.*, 1969). Experiments were performed to test the efficacy of purified enzyme or plasma and leucocytes when administered to patients with MPS (reviewed in Neufeld and Muenzer, 1989). The results obtained were disappointing and no influence on pathology could be detected, in retrospect probably because the amount of enzyme

administered was insufficient and was not of a form readily 'taken-up' by the process of receptor-mediated endocytosis (see Section 1.4.1).

The poor results obtained with enzyme-replacement encouraged other treatment procedures to be developed. Freidman and Robin (1972) suggested that it was feasible to introduce a normal gene into patient cells *in vitro* and reintroduce the cells to the patient and correct or replace the function of a defective gene. This proposal has seen much development over the last two decades (Anderson, 1984; Freidman, 1989). Bone marrow stem cells were identified as the best candidate for carriers of corrective genes, after the limited success with the clinical use of bone marrow transplantation (BMT) to treat various MPS (mostly MPS I) disorders (Hobbs *et al.*, 1981; Krivit and Paul, 1986). Cells derived from hematopoietic progenitors of the donor (circulating leucocytes and tissue macrophages) may donate lysosomal enzymes by secretion or direct cell-cell interaction to the deficient cells in most tissues of the host.

With the development of recombinant DNA techniques the prospects for novel therapy protocols has improved. It is generally accepted that before any of these protocols can be used on MPS patients the proposed therapy should be evaluated in cell culture and then in animal MPS models. There are several animal models (dog, goat, cat, mouse) available for the MPS (I, IIID, VI and VII) respectively, which bear identical biochemical and clinical features to their human counterparts (Jezyk *et al.*, 1977; Haskins *et al.*, 1979, 1984; Shull *et al.*, 1982; Spellacy *et al.*, 1983; McGovern *et al.*, 1985; Kyle *et al.*, 1990; Prof. Margaret Jones, personal communication (MPS IIID goat)).

MPS II involves somatic (bone and tissue) as well as some neurological symptoms, presenting two sets of problems for the access of gene products to the tissues of

pathology and also adds to the complication of evaluation of therapy regimes. BMT has been shown to be effective in preventing the neurological degeneration in fucosidosis dogs (Taylor *et al.*, 1989) which may be a model for the neurological degeneration seen in the MPS. This indicates that the corrective enzyme is able to cross the blood-brain barrier. This is encouraging for the future therapy prospects for MPS II as well as the other MPS with neurological involvement. The improved level of expression of recombinant enzymes in cell culture should also enable the evaluation of enzyme-replacement therapy to proceed in the available animal models (see Chapter 6 for further discussion).

Bone marrow cells appear to be the most favoured system to engineer therapy because of their capacity to proliferate and colonize the host. BMT has been performed and evaluated in MPS I dogs and MPS VI cats, and was shown to reduce the clinical severity in both systems (Breider *et al.*, 1989; Gasper *et al.*, 1984). To date MPS I, MPS II, MPS VI and MPS VII cultured cells and an MPS VII mouse model have been corrected for lysosomal storage, using retroviral vectors or transgenesis in the MPS VII mouse (Anson *et al.*, 1991,1992; Bielicki *et al.*, 1992; Wolfe *et al.*, 1990; and Kyle *et al.*, 1990).

Several factors need to be overcome to ensure the progress of gene therapy. Firstly, the transport of the active enzyme to the site of pathology is a major problem, i.e. to the growth-plate cartilage and or to the brain via the blood-brain barrier. The transport of enzyme past the blood-brain barrier may not prove to be a problem for the MPS with little neurological involvement, however, in the severe forms of MPS II, mental retardation is a major symptom and whether this will prove to be an obstacle or not will need to be established. It may be that, as in the fucosidosis case (Taylor *et al.*, 1989), the normal IDS enzyme will be able to cross the blood-brain barrier and

correct/prevent the MPS II neurological pathology. Currently there is no MPS II animal model to evaluate these procedures, however, it may be possible to create a mouse model via the use of homologous recombination of embryonic stem cells (Capecchi, 1989). However, a larger animal model system would be preferable to obtain conditions similar to humans.

1.8 AIMS OF THE PROJECT

The research presented in this thesis was aimed at the isolation and characterisation of the IDS gene. The project was commenced when the N-terminal amino acid sequence of an IDS 42-kDa polypeptide became available from work within the Adelaide group (Bielicki *et al.*, 1990). The specific aims were as follows.

1. To use the N-terminal amino acid sequence to design an oligonucleotide to use as a probe to screen cDNA and genomic libraries to isolate the IDS gene.

2. To use an IDS cDNA clone as a probe to study MPS II patients at the Southern analysis level.

3. To determine the molecular nature of mutations in MPS II patients.

4. To look for a correlation between the patient clinical phenotype and the molecular nature of each gene defect.

5. To gain insights into areas of the IDS protein that are critical for enzyme function by the study of mild and severe MPS II patient mutations.

CHAPTER TWO

MATERIALS AND METHODS

2.1 MATERIALS

2.1.1 ELECTROPHORESIS

Acrylamide	Biorad, Richmond, California, USA
Agarose, type I, low EEO	Sigma Chemical Co., St. Louis, Missouri, USA
Agarose, low gelling temperature	Sigma
Ammonium persulphate	Ajax, Auburn, NSW, Australia
Bromophenol blue	B.D.H. Chemicals Ltd., Poole, Dorset, England
Ethidium bromide	Boehringer Mannheim, Germany
Formamide	Ajax
NuSieve GTG Agarose LMP	FMC Bioproducts, Rockland, USA
Sequagel sequencing system	National Diagnostics, New Jersey, USA
N',N',N',N', -tetramethylethenediamine (TEMED)	Biorad
Urea	Ajax
Xylene cyanol FF	Tokyo Kasei, Tokyo, Japan

2.1.2 ENZYME

All the enzymes used in this study were obtained from the following companies:

Calf intestinal phosphatase	Boehringer Mannheim
E.coli DNA-polymerase I (Klenow fragment)	Amersham
Bst DNA polymerase	Biorad
Mung bean nuclease	Pharmacia
Lysozyme, chicken	Sigma
Proteinase K	Sigma

Ribonuclease A	Sigma
Reverse transcriptase (Mo-MLV)	Bestheda Besearch Laboratories
T4 polynucleotide kinase	Pharmacia, Uppsala, Sweden
T4 DNA-ligase	Boehringer Mannheim
Taq-polymerase	Perkin Elmer Cetus, USA and Biotech International, WA, Australia

All the restriction endonucleases used in this study were obtained from Boehringer Mannheim and New England Biolabs, Beverly, USA.

2.1.3 RADIOCHEMICALS

α-^{32}P-dCTP, 3000Ci/mmole	Radiochemical Centre, Amersham
γ-^{32}P-dATP, 5000Ci/mmole	Radiochemical Centre, Amersham
γ-^{32}P-dATP, 5000Ci/mmole	Bresatec, Adelaide, Australia

2.1.4 BUFFERS

Buffers commonly used in this study are as follows:

5 x Ligation	250 mM Tris-HCl, pH 7.6, 50 mM $MgCl_2$, 25 % (w/v) PEG 6000, 5 mM ATP, 5 mM DTT
PBS	130 mM NaCl, 10 mM Na_2HPO_4, 10 mM NaH_2PO_4, pH 7.2.
10 x PCR	500 mM KCl, Tris-HCl, pH 8.4, 25 mM $MgCl_2$, 1% (v/v) Triton X-100, 0.02% (w/v) gelatin.
20 x SSC	3 M NaCl, 0.3 M tri-sodium citrate.$2H_2O$, pH 7.0.
SM	100 mM NaCl, 10 mM $MgSO_4.3H_2O$, 0.01% (w/v) gelatine, 50 mM Tris-HCl, pH 7.5.
M9 salts	1.05% (w/v) K_2HPO_4, 0.45% (w/v) KH_2PO_4, 0.1% (w/v) $(NH_4)_2SO_4$, 0.05% (w/v) sodium citrate.

TAE	40 mM Tris-acetate, 2 mM EDTA, pH 8.5.
TBE	89 mM Tris-base, 89 mM boric acid, 2.5 mM EDTA, pH 8.3.
TE	10 mM Tris-HCl, pH 7.5, 0.1 mM EDTA.
5 x P	1% (w/v) bovine serum albumin, 1% (w/v) polyvinylpyrrolidone Mr 40,000, 1% (w/v) ficoll Mr 400,000, 250 mM Tris-HCl, pH 7.4, 0.5% (v/v) pyrophophate

2.1.5 BACTERIAL MEDIA

2.1.5a LIQUID MEDIA

All liquid media were prepared in 0.04 μSiemans water and were sterilised by autoclaving for 25 min at a temperature of 120°C and a pressure of 120 KPa. The compositions of the various media were as follows:

L-broth	1% (w/v) Bacto-Tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl.
2 x YT	1.6% (w/v) Bacto-Tryptone, 1% (w/v) yeast extract, 0.5% (w/v) NaCl.
Minimal medium	1 mM $MgSO_4$, 0.1 mM $CaCl_2$, 1 mM thiamine HCl, 0.2% (w/v) glucose, 5 ml 10 x M9 salts, made up to 50 ml.

2.1.5b SOLID MEDIA

Plates were prepared using the appropriate volume of media depending on the size, i.e. 90 mm or 150 mm diameter.

L-Agar	L-broth containing 1% (w/v) Bacto-agar
L-Agarose	L-broth, 1% (w/v) Agarose
L-Amp	L-broth, 1% (w/v) Bacto-agar, ampicillin (120 μg/ml)
Minimal medium plus glucose	Minimal medium, 1% (w/v) Bacto-agar, 0.02% (w/v) $MgSO_4$, 0.2% (w/v) glucose, 0.0005% (w/v) Thiamine-HCl.

2.1.6 ANTIBIOTICS

Ampicillin	Sigma
Chloramphenicol	Sigma
Neomycin (G418)	Gibco, Glen Waverley, Vic, Australia
Kanamycin	Boehringer Mannheim
Streptomycin	Boehringer Mannheim

2.1.7 BACTERIAL STRAINS

The E. coli K12 strains used in this study are described below:

C600	F-thr-1, leuB6, thi-1, lacY1, supE44, rfbD1, fhuA21, mcrA1 (Raleigh *et al.*, 1988).
JM101	F' traD36, $lac^q\Delta$(lacZ), M15, proAB, supE, thi Δ(lac-proAB) (Yanish-Perron *et al.*, 1985).

LE392	F^-, hsdR514 ($r^{k-}m^{k+}$), supE44, supF58, lacY1 or Δ(lacIZY)6, galK2, galT22, metB1, trpR55 (Maniatis *et al.*, 1982).
Y1090	F^-, Δ(lac), U169, Ion-100, araD139, rpsL(str^r), supF, mcrA, trpC22, Tn10(pMC9;Tet^rAmp^r) (Huynh *et al.*, 1985)
ED8799:	$hsdS_k$, met B7, SupE (gln V)44, supF(tyr T)58, Δ(lacZ)M15.

2.1.8 BACTERIOPHAGE STRAINS

The bacteriophage strains used in this study are described below:

λCharon 35	(Loenen and Blattner, 1983)
λgt10	Lam-b527, srl/lam3°, imm^{434}, srl/lam4°, srl/lam5° (Huynh *et al.*, 1984).
λgt11	lam/lac5, srl/lam3°, cl857, srl/lam4°, nin5, srl/lam5°, Sam100 (Young and Davis, 1983)
M13mp18	Yanish-Perron *et al.*, 1985
M13mp19	Yanish-Perron *et al.*, 1985
pBLUESCRIPT SK II	Stratagene

2.1.9 CELL CULTURE

Ham's F-10 nutrient media	Gibco
Glutamine	Flow Laboratories, Sydney, Australia
Phosphate Buffered Saline	Commonwealth Serum Laboratories, Melbourne, Vic., Australia
Fetal Calf Serum	Commonwealth Serum Laboratories

2.1.10 MISCELLANEOUS MATERIALS

Nylon filters (Genescreen Plus)	NEN Research Products, Boston, MA, USA
Colony Plaque Screen filters	NEN Research Products
GF/A glass filter discs	Whatman Ltd., Poole, England
Polaroid film type 667	Polaroid Corp., Cambridge, MA, USA
X-ray film	Fuji Rx X-ray film, Fuji Photo film Co. Ltd., Japan
Sephadex G-50	Pharmacia P-L Biochemicals
Oligo d(T)-cellulose	Pharmacia P-L Biochemicals
DEAE-cellulose paper	Schleicher and Schuell, Bassel, Germany

2.1.11 MISCELLANEOUS FINE CHEMICALS

Caesium chloride (Molecular Biology Grade)	Boehringer Mannheim
Diethlypyrocarbonate	Sigma
Phenol	Wako Chemical Co., Japan
Sarkosyl	Ciba-Geigy, Basle, Switzerland
SDS	Sigma
Osmium tetroxide	Sigma
Hydroxylamine	Sigma
Piperidine	Sigma
Pyridine	Aldrich, USA
Herring Sperm DNA	Sigma

All other chemicals used in this study were analytical reagent grade.

2.2 GENERAL METHODS

2.2.1 PREPARATION OF GLASSWARE AND SOLUTIONS

All solutions were prepared using reverse osmosis water (0.04 μSiemens) and made nuclease-free and microbiologically sterile by autoclaving at 125°C and 120 KPa. Glassware was treated in pyroneg detergent (non-ionic and non-PO_4) and rinsed in deionised water and dried overnight. Spatulas and centrifuge tubes were rinsed in ethanol and air-dried.

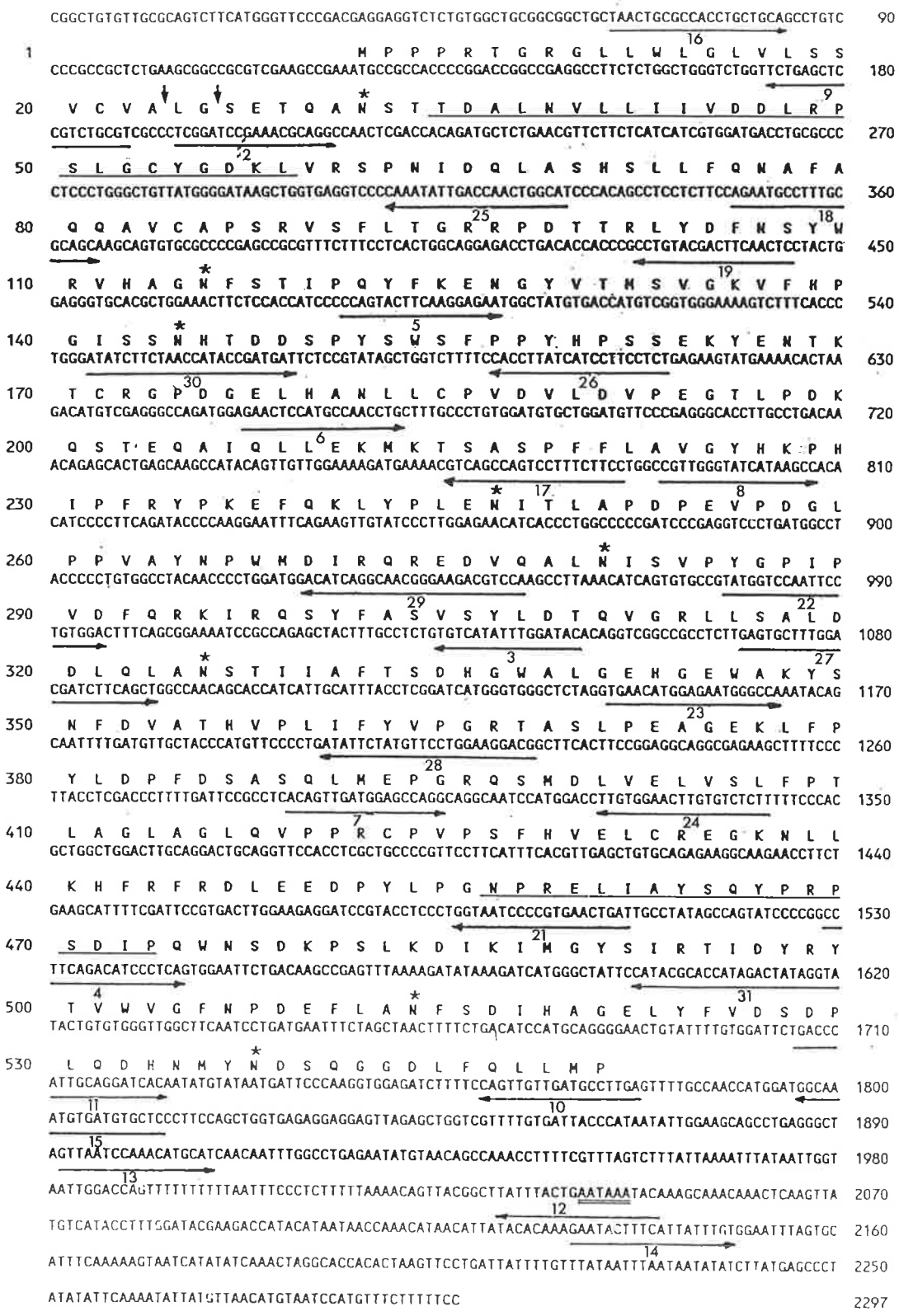
2.2.2 DESIGN AND PREPARATION OF SYNTHETIC DNA OLIGONUCLEOTIDES

The majority of synthetic oligonucleotides used in this study were prepared by Annette Orsborn.

DNA oligonucleotides were designed using an empirical rule, where primers were designed to contain an equal proportion of purine and pyrimidine bases. Long stretches of the same base were avoided if possible. All the primers were synthesised using an Applied Biosystems 391 DNA synthesiser. Primers were deprotected by incubation overnight in 10 M NH_4OH at 55°C. The primers were purified using the method of Sawadogo and Van Dyke (1991). Basically, 100 μl of the deprotected primer solution was added to 1 ml of n-butanol and centrifuged in an Eppendorf model 5414s microcentrifuge (all centrifugations described in this Chapter and elsewhere in this thesis were carried out on this machine unless otherwise stated) at 12,000 rpm for 30 sec. This procedure was repeated once, and the oligonucleotide pellet was washed once in 70% aqueous ethanol, air-dried and resuspended in an appropriate vol of water. The primers used in this study are indicated in Figure 2.1.

Figure 2.1 **SPECIFIC IDS cDNA SEQUENCE PRIMERS**

For design and processing of oligonucleotides see Section 2.2.2.



2.2.3 RADIOLABELLING OF DNA

2.2.3a 5'-TERMINAL RADIOLABELLING

Synthetic oligonucleotides were used at the appropriate concentrations depending on their size and end-labelled according to the manufacturer's specifications (Pharmacia, Sweden) using γ-32-dATP (5000 Ci/mmole).

2.2.3b RANDOM PRIMING AND RADIOLABELLING OF DNA

Labelling double-stranded DNA using the Klenow fragment was carried out according to the manufacturer's specifications (Amersham). Basically, a small quantity of DNA insert (25-50 ng) was denatured and added to a solution containing random hexamers, dNTPs (minus dCTP), α-^{32}P-dCTP, 2 units of Klenow enzyme and buffer. After a sufficient time of incubation (usually 2 hours) the radioactive probe was purified (see below), denatured and used in the hybridisation experiment.

2.2.2c PROBE PURIFICATION

After the probe was labelled it was run on a Sephadex G-50 column to separate the unincorporated nucleotide and to enable the efficiency of incorporation to be determined. This was crudely monitored by analysis of the various fractions with a radiation monitor. Exact specific activity was determined by scintillation counting of a small aliquot of the fractionated probe.

2.2.4 SCREENING OF RECOMBINANT DNA LIBRARIES

The 42-kDa polypeptide N-terminal amino acid sequence was obtained from IDS purified from human liver (Beilicki *et al.*, 1990). The sequence was found to have significant homology with other known sulphatases and was used to design an oligonucleotide probe based on human codon usage. Initially, an EMBL 3 genomic library was screened at 40,000 plaque forming units (pfu) per plate on the host bacteria NM538 with an end-labelled 49-mer oligonucleotide (see Section 2.2.3a). When the plates were grown after approximately 12 hours, they were placed at 4°C to aid the transfer to nylon colony plaque screen filters (NEN, Dupont) according to the manufacturers instructions. When the plaques were transferred the filter was immersed in 0.4 M NaOH for 2 min, then in 1.0 M Tris-HCl, pH 7.5, for 2 min, and the procedure was repeated once. The filters were allowed to air dry and used for the hybridisation experiment according to the manufacturer's specifications as follows.

The filters were placed in 10 ml of pre-hybridisation solution containing 2 ml of 5 x P (1% (w/v) bovine serum albumin, 1% (w/v) polyvinylpyrrolidone Mr 40,000, 1% (w/v) ficoll Mr 400,000, 250 mM Tris-HCl, pH 7.4, 0.5% (w/v) pyrophosphate), 2 ml 5 M NaCl, 1 ml 10% (w/v) SDS, 2 ml 50% (w/v) dextran sulphate, made up to 10 ml with water and 10 mg/ml denatured herring sperm DNA and shaken at 42°C for 30 min.

After the filters were pre-hybridised the radiolabelled probe was added at 5 ng/ml and incubated at 42°C overnight. The filters were washed according to the manufacturer's specifications. The filters were autoradiographed overnight with intensifing screens at -80°C.

When oligolabelled probes were used the conditions were as follows: the hybridisation solution contained 50% (v/v) formamide, 1% (w/v) SDS, 1 M NaCl, 10% (w/v) dextran sulphate and 0.5 ml of denatured herring sperm DNA (10 mg/ml) and was incubated at 42°C for at least 120 min. The denatured radioactive probe was added at a concentration of approximately 10 ng/ml (1-4 x 10^5 dpm/ml) (see Section 2.2.3b for probe preparation). This was incubated overnight at 42°C. The filter was washed in 2 x SSC (0.3 M NaCl; 0.03 M tri-sodium citrate) at room temperature for 5 min with constant agitation. This was repeated with 2 x SSC; 1% (w/v) SDS at 65°C for 30 min. Depending on the radioactive signal obtained, a further wash of 30 min with 0.1 x SSC at room temperature was done. The filter was autoradiographed with intensifying screens at -80°C overnight.

2.2.5 PREPARATION OF LAMBDA DNA FROM PLATE LYSATES

The λ phage were plated-out at a density that would result in confluent lysis, i.e. approximately 1 x 10^5 pfu's per 150 mm plate. The plates were incubated at 37°C for 9 hours, or such a time as to prevent the appearance of resistant bacteria. The plates were covered with 15 ml of 1 x SM solution and gently shaken at 4°C overnight. The SM solution was carefully removed after incubation and placed in 50 ml centrifuge tubes. These tubes as well as the tubes described in this section were centrifuged in a Sorvall J2-21M/E high speed centrifuge (JA-21 roter) at 8,000 rpm for 20 min to ensure the removal of all bacteria. The supernatant was transferred to a fresh tube and 10 μg/ml DNAase I and 100 μg/ml RNAase A was added and the tube was incubated at 37°C for 60 min. An equal vol of 2.5 M NaCl/20% (w/v) PEG 6000 was added and the solution was incubated on ice for 60 min. The tubes were centrifuged at 12,000 rpm for 20 min, the supernatant was discarded and the pellet drained to ensure the complete removal of PEG. The pellet was resuspended in 1.5 to 2.0 ml of SM

buffer, vortexed briefly to ensure complete resuspension, the tube was centrifuged briefly to transfer the solution to the bottom of the tube, and 20 μl of 0.5 M EDTA and 20 μl of 10% (w/v) SDS were added prior to incubation at 55°C for 60 min. The solution was extracted twice with water-saturated phenol, phenol/chloroform and chloroform and then centrifuged at 12,000 rpm for 20 min for each extraction and an equal vol of isopropanol was added before incubation at -20°C overnight to precipitate the DNA. The tube was spun at 12,000 rpm for 20 min and the DNA pellet was washed with 95% (v/v) aqueous ethanol, dried and resuspended in an appropriate vol of water.

2.2.6 PLASMID VECTOR PREPARATION

Plasmid DNA was digested with the appropriate restriction endonuclease depending on the required termini. The digestion was analysed on a 1% (w/v) agarose gel to ensure complete digestion. The DNA was phenol- and chloroform-extracted and ethanol precipitated. The DNA pellet was resuspended at a concentration of 100 ng/μl in water.

2.2.7 LIGATION OF PLASMID VECTORS

The DNA insert and appropriate plasmid vector were combined in a molar ratio of 3:1 in a 20 μl reaction vol containing 50 mM Tris-HCl, pH 7.6, 10 mM $MgCl_2$, 1 mM ATP, 5% (w/v) PEG 6000, 1 mM DTT. For ligations with overhanging ends, 0.5 units of T4 DNA ligase were used, and for ligations with blunt-ends, 1.0 unit was used. Ligations were incubated at 16°C overnight.

2.2.8 TRANSFORMATION OF *E. coli*, ED8799

Essentially as for JM101 (see Section 2.2.19). The transformation procedure was as described in the M13 Biorad handbook.

2.2.9 PLASMID PREPARATION

Recombinant plasmid DNA was prepared as described in the Promega 89/90 catalogue. Recombinant clones were grown overnight in L-broth supplemented with ampicillin (100 μg/ml); 1 ml was transferred to 500 ml of fresh L-broth and grown to stationary phase. The cells were pelleted by centrifugation in a Sorvall J2-21 M/E centrifuge (the other centrifugations described in this section were carried out on this machine) and resuspended in 12 ml of ice-cold, freshly prepared buffer (25 mM Tris-HCl, pH 8.0, 10 mM EDTA, 15% (w/v) sucrose and 2 mg/ml lysozyme). The resuspended cells were kept on ice-water for 10 min. Twenty-four ml of freshly prepared 0.2 M NaOH and 1% (w/v) SDS was added, the solution was carefully mixed by inversion and incubated on ice for 10 min. Fifteen ml of 3 M sodium-acetate, pH 4.6, was added and the solution was again carefully mixed by inversion and incubated in ice-water for 20 min. The solution was centrifuged at 12,000 rpm for 15 min and the supernatant was carefully transferred to a fresh tube. One hundred μl of RNAase A (1 mg/ml) was added to the supernatant and incubated for 20 min at 37°C. The solution was extracted twice with 1 vol of phenol:chloroform (1:1) saturated in TE buffer, and vortexed for 1 min. The solution was centrifuged at 12,000 rpm for 5 min and the aqueous phase was transferred to a fresh tube and 1 vol of chloroform:isoamyl alcohol (24:1) was added. This was also vortexed and centrifuged as described above. The aqueous phase was transferred to a fresh tube and 2 vol of ethanol was added and the solution was left at -20°C for 30 min. The tube was centrifuged at 12,000 rpm for 20

min to pellet the precipitated DNA. The tube was carefully drained, the DNA pellet was dissolved in 3.2 ml of water 1 ml of 4 M NaCl and 4 ml of 13% (w/v) PEG 6000 was added and the solution was mixed and incubated on ice for 60 min. The tube was centrifuged at 12,000 rpm for 10 min, the supernatant was removed and the DNA pellet was washed once with 70% (v/v) aqueous ethanol. The pellet was dried under vacuum and dissolved in an appropriate volume of water or TE buffer. Depending on the requirements of the DNA, i.e. for cell culture expression, the DNA was further purified by ultracentrifugation on a caesium chloride gradient.

2.2.10 PRECIPITATION WITH ETHANOL

In general, all samples were made 300 mM with sodium acetate using a stock solution of 3 M at pH 5.5. Then, 2.5 vol of ethanol were added to the sample, mixed and placed at -20°C for an appropriate time, depending on the concentration of DNA expected. Precipitates were collected by centrifugation at 12,000 rpm for 15 min at 4°C. The supernatant was carefully removed with a drawn-out pasteur pipette. The pellet was rinsed with 500 μl of 70% (v/v) aqueous ethanol, dried briefly under vacuum and resuspended in either water or TE buffer.

2.2.11 DIGESTION OF DNA WITH RESTRICTION ENZYMES

Restriction endonuclease digestion of DNA was carried out using the conditions appropriate for each enzyme as detailed in the Boehringer Mannheim catalogue. Generally, an excess of enzyme was used (typically 10 units for 1 μg of DNA) and the time was also extended two-fold to ensure complete digestion.

2.2.12 AGAROSE GEL ELECTROPHORESIS OF DNA

Electrophoresis of DNA for analytical purposes was carried out in small submarine tanks. The electrophoresis of DNA for transfer to nylon membranes was carried out in a variety of tanks according to the number of samples. Agarose was used at 0.8 to 2.0% (w/v) in TAE buffer. Samples were electrophoresed at 100 V until the bromophenol blue marker dye had migrated a sufficient distance to ensure adequate separation of the DNA fragments. DNA was visualised under UV light after the gel had been briefly stained in 10 μg/ml of ethidium bromide. A photograph was then taken using polaroid film type 667.

2.2.13 POLYACRYLAMIDE GEL ELECTROPHORESIS

2.2.13a NON-DENATURING GELS

For non-denaturing polyacryamide gels, a prepared solution of 20:1 (arylamide:bis-acrylamide) was used at the desired percentage depending on the size separation required (Biorad). The gels were run in TBE buffer, pH 8.3, on a Hoeffer Mighty Tall gel system at 80 V and run for an appropriate time to achieve the desired separation.

2.2.13b DENATURING GELS

For denaturing sequencing polyacrylamide gels a commercially prepared solution was used according to the manufacturer's specifications (usual concentration of 6% (w/v) ; Sequegel, National Diagnostic); these gels were also run in TBE buffer, pH 8.3. The gels were run at 2,000 V, 60 mA and 100 W until the bromophenol dye was 4 cm from the bottom. After electrophoresis, the gel was fixed for 10 min with 10% (v/v) acetic acid, rinsed with 20% (v/v) aqueous ethanol and transferred to a sheet of Whatmann

3MM paper covered with plastic wrap and dried on a Biorad gel dryer. Once dry, the plastic wrap was removed and the gel was autoradiographed overnight at room temperature.

2.2.14 ISOLATION OF DNA RESTRICTION FRAGMENTS

2.2.14a FROM POLYACRYLAMIDE GELS

Restriction fragments were detected under UV light after the polyacrylamide gel was briefly stained with ethidium bromide (10 μg/ml). The required band(s) were excised from the gel and placed in elution buffer (500 mM ammonium acetate; 10 mM $MgCl_2$; 2 mM EDTA; 20 mM Tris-HCl, pH 8.0; 0.2% (w/v) SDS) and incubated overnight at 65°C. Two vol of ethanol was added and the solution was incubated overnight at -20°C. This solution was centrifuged at 12,000 rpm for 15 min at 4°C. The DNA pellet was resuspended in 300 mM sodium acetate, pH 5.5, and ethanol-precipitated again as described above. The DNA pellet was washed with 70% (v/v) aqueous ethanol and either air- or vacuum-dried. The DNA pellet was resuspended in the appropriate solution, depending on the purpose intended.

2.2.14b FROM LOW MELTING POINT AGAROSE GEL

Restriction fragments were detected under UV light after the agarose gel was briefly stained with ethidium bromide (10 μg/ml). The appropriate band was excised from the gel and the DNA was extracted as described by Maniatis *et al.* (1982). Essentially, the gel slice was placed in 5 vol of buffer (50 mM Tris-HCL, pH 8.0, 1 mM EDTA) and incubated at 65°C until the agarose had melted completely. This was phenol-extracted with an equal vol, vortexed, centrifuged for 5 min at 12,000 rpm and the aqueous phase was collected, and this procedure was repeated twice. The DNA was precipitated with 0.1 vol 3 M sodium acetate and 2.5 vol of ethanol.

2.2.14c DNA EXTRACTION WITH GENECLEAN AND DEAE PAPER

For Geneclean, the manufacturer's specifications were followed exactly (Bio 101, California). Essentially, the appropriate band was excised and then placed in a solution of NaI and incubated at 55°C until the agarose had melted. A silica bead solution was then placed in the melted agarose solution and the tube was placed on ice for 5 min. The silica beads were pelleted with centrifugation and subsequently washed and resuspended three times in a wash solution. When this was complete the pellet was resuspended in the elution buffer (water or TE) incubated for 3 min at 55°C and the silica matrix was again pelleted. The aqueous solution was collected and this procedure was repeated once.

When DEAE-cellulose paper was used, the appropriate band was identified after staining in ethidium bromide (10 μg/ml) and a small piece of DEAE-paper was cut and placed alongside the band. The gel was then turned 90°, the current was turned back on and the DNA was collected onto the paper. Once the DNA had been collected the paper was placed in a high salt elution buffer (1 M NaCl; 0.1 mM EDTA; 20 mM Tris-HCl, pH 8.0) and incubated at 65°C for 30 min. Two vol of ethanol were added and the DNA was precipitated as described in Section 2.2.10.

2.2.15 TRANSFER OF DNA TO NYLON MEMBRANES

DNA digested with restriction endonucleases and fractionated on agarose slab gels was transferred using the alkaline method. In this procedure the gel is first treated twice in 0.25 M HCl with shaking for 10 min or until the bromophenol blue dye turns yellow. The gel was then washed twice in 0.4 M NaOH for 20 min with shaking. The nylon membrane was then cut to size and placed in the transfer solution (0.6 M NaCl, 0.4 M NaOH) and allowed to soak until the membrane was saturated. The membrane was

placed on top of the gel and three sheets of 3MM paper and a stack of blotting tissue were placed on top. Transfer was usually done overnight. When transfer was complete the filter was neutralised in 1 M Tris-HCl, pH 7.5, and allowed to air dry and was ready for hybridisation with radiolabelled probes.

2.2.16 HYBRIDISATION OF RADIOACTIVE PROBES TO IMMOBOLIZED DNA

The pre-hybridisation and hybridisation of radioactive probes to DNA immobilised on nylon was carried out according to the manufacturer's instructions (NEN Research Products). Essentially, two procedures were used, one for 5'-end-labelled primers (kinased) and one for random primed insert DNA (oligolabelled). For kinased primers the filters were pre-hybridised in 5 x P (1% (w/v) polyvinylpiridine; 1% (w/v) ficoll; 250 mM Tris-HCl, pH 7.4; 0.5% (w/v) sodium pyrophosphate; 1% (w/v) BSA) for 60 min and the probe was added and the filters were shaken overnight at 42°C. The filters were washed according to the size of the primer used, i.e. for a 49-mer kinased primer, 6 x SSC; 0.1% (w/v) SDS, 42°C for 5 min. The filter was autoradiographed with intensifying screens at -80°C overnight. For oligolabelled probes, the hybridisation solution was: 50% (w/v) formamide, 1% (w/v) SDS, 1 M NaCl and 10% (w/v) dextran sulphate; 0.5 ml of denatured herring sperm DNA (10 mg/ml) and this was pre-hybridised at 42°C for at least 2 hours. After pre-hybridisation the denatured radioactive probe was added at a concentration of approximately 10 ng/ml (1-4 x 10^5 dpm/ml) (see Section 2.2.3b for probe preparation). This was incubated overnight at 42°C. The filter was washed in 2 x SSC (0.3 M NaCl; 0.03 M tri-sodium citrate) at room temperature for 5 min with constant agitation. This was repeated with 2 x SSC; 1% (w/v) SDS at 65°C for 30 min. Depending on the radioactive signal obtained a further wash of 30 min with 0.1 X SSC at room

temperature was done. The filter was autoradiographed with intensifying screens at -80°C overnight.

2.2.17 PREPARATION OF M13 VECTORS

The replicative form (RF) of M13 mp18 and 19 was obtained from Boehringer Mannheim. The RF M13 DNA was digested with the appropriate restriction enzyme(s) to generate the desired termini. The digested DNA was examined on a 1% agarose gel to ensure that there was no intact molecules present. The DNA was phenol- and chloroform-extracted and the DNA was precipitated as described in Section 2.2.10 and resuspended in water to a final concentration of 100 ng/μl. The vector was phosphatased with calf intestinal phosphatase to remove the 5'-terminal phosphate so as to prevent self-ligation of the vector.

2.2.18 LIGATION OF M13 VECTORS

The DNA insert and appropriate M13 vector were combined in a molar ratio of 5:1 in a 20 μl reaction mix containing 50 mM Tris-HCl, pH 7.6, 10 mM $MgCl_2$, 5% (w/v) PEG 6000, 1 mM ATP and 1 mM DTT. For ligations with overhanging termini 0.5 units of T4 DNA ligase were used, and 1.0 unit was used for blunt-end ligations. Both of these reactions were incubated at 16°C overnight.

2.2.19 TRANSFORMATION OF *E. coli*, JM101

E. coli, strain JM101, was made competent according to the Biorad M13 cloning handbook. Essentially, the bacteria were harvested after reaching an A_{600} of between 0.4-0.6. The cells were transferred immediately to an ice-cold solution of 100 mM $MgCl_2$ and incubated on ice for 30 min. The cells were reharvested and resuspended

in 100 mM $CaCl_2$ and left on ice for at least 120 min. The transformation procedure was carried out as described in the Biorad M13 handbook.

2.2.20 HARVESTING (+) STRAND OF RECOMBINANT BACTERIOPHAGE

The white recombinant plaques were harvested from 2 ml cultures as described by the Biorad handbook, except that after the PEG 6000 precipitation stage it was treated according to Eperon (1986). Basically, the PEG pellet was resuspended in 180 μl of water and 20 μl of cracking buffer (100 mM Tris-HCl, pH.7.5; 1 mM EDTA; 5% (w/v) SDS) and incubated at 80°C for 10 min, cooled to room temperature and 80 μl of 7.5 M ammonium-acetate and 560 μl of isopropanol was added. This solution was incubated at room temperature for 15 min and centrifuged at 12,000 rpm for 10 min. The pellet was washed once in 70% (v/v) aqueous ethanol, vacuum-dried and resuspended in 30 μl of water.

2.2.21 DIDEOXY SEQUENCING REACTIONS

2.2.21a HYBRIDIZATION

Four μl of the single-stranded M13 template, 1 μl of primer (2 ng of universal sequencing primer or specific primers), 1 μl of 10 x TM buffer (100 mM Tris-HCl, pH 7.5; 100 mM $MgCl_2$) and 4 μl of water were combined and incubated at 75°C for 5 min and allowed to cool to room temperature. The tubes were centrifuged briefly to bring down condensation.

2.2.21b POLYMERISATION

The chain termination reaction (Sanger *et al.*, 1977) was as described in the Biorad M13 sequencing handbook. The radiolabelled nucleotide is α-^{32}P-dCTP and the reactions were carried out at 45°C with the Klenow fragment. In some cases the enzyme *Bst* polymerase (Biorad) was used and incubation in this case was at 65°C. The final dideoxynucleotide concentrations used in this study varied according to the template to be sequenced. Generally, 44 μM ddATP, 44 μM ddCTP, 53 μM ddGTP and 222 μM ddTTP. An aliquot of each reaction was taken and added to formamide loading dye (95% (v/v) formamide; 20 mM EDTA; 0.03% (w/v) xylene cyanol; 0.03% (w/v) bromophenol blue), heat-denatured and loaded onto a 6% (w/v) sequencing gel as described in Section 2.2.13b.

2.2.22 COMPUTER ANALYSIS OF DNA SEQUENCES

The nucleotide sequence was screened against the GenBank nucleotide sequence data base (Release 71, March 1992) and the encoded protein sequence was screened against the Swissprot (Release 21, March 1992). The compilation of the nucleotide sequence from the multiple subclones was done with a number of programs based on Staden (1980) unless otherwise stated.

RODENT	Used to enter DNA sequence data. (Biotel, North Ryde, Australia)
DBAUTO	Used to construct contig of DNA.
SPCOMP	Used to search for overlapping DNA sequence.

ANALYSEQ Multiple analysis, including search for restriction enzyme sites, translation of DNA sequence into protein sequence plus many other DNA manipulation programs.

MWCALC Calculates the molecular weight of protein sequences.

General sequence analysis and the alignment of protein sequences was done using the programs of Reisner and Bucholtz (1986), Bucholtz and Reisner, (1986) and Lipman *et al.* (1989), respectively.

CHAPTER THREE

ISOLATION AND IDENTIFICATION OF A POTENTIAL IDS GENOMIC CLONE

3.1 INTRODUCTION

The purification of the IDS protein has proved to be extremely difficult due to its very low abundance in human tissues. Previous investigators have reported only partial purification of the IDS enzyme (Cantz *et al.*, 1972; Di Natale and Ronsisvalle, 1981; Archer *et al.*, 1982; Yutaka *et al.*, 1982; Wasteson and Neufeld, 1982; Lissens *et al.*, 1984) (Section 1.4.2). The impetus for commencing the work described in this thesis was the generation of a small amount of peptide sequence data for IDS by our laboratory, discussed in Section 3.3. At this time the latest purification of IDS by other investigators from human placenta reported a 35,000-fold purification to homogeneity (Di Natale and Daniele, 1985), but no IDS amino acid sequence was presented.

More recently, our laboratory reported a 500,000-fold purification of IDS from human liver after a six-column procedure which included a concanavalin A Sepharose Blue A-agarose-coupled step, chromatofocusing, gel filtration on TSK HW 50S-Fractogel, hydrophobic separation on phenyl-Sepharose CL-4B and size separation on TSK G3000SW Ultrapac (Bielicki *et al.*, 1990). The chromatofocusing step led to the separation of two major forms of purified enzyme (A and B) from all the tissue sources studied; both enzyme forms contained a 42-kDa polypeptide and a 14-kDa polypeptide as well as other protein species. When the purified enzyme sample (form A) was run on SDS-PAGE a major polypeptide species of 42-kDa was seen along with lesser amounts of 30-kDa, 20-kDa and 14-kDa polypeptides (Figure 3.1); the 14-kDa species appears to be running at about 18-kDa in Figure 3.1, but under different separation conditions it runs at 14-kDa. When the samples were run under non-reducing conditions the 30-kDa and 20-kDa polypeptides were found to associate to yield a 48-kDa polypeptide. It was therefore thought that these polypeptide species did not interact with the 42-kDa polypeptide. At the time this study was commenced it was thought that the 42-kDa polypeptide was the active IDS enzyme. The basis for

this reasoning was two-fold: firstly, the 42-kDa polypeptide was the most abundant polypeptide species and, secondly, it was consistently present in all the protein purifications of IDS, although the possibility that other polypeptides may interact with the 42-kDa polypeptide to produce active enzyme was not ruled out.

N-terminal amino acid sequencing data was obtained from IDS by transferring the 42-kDa polypeptide to an Immobilon P membrane followed by Edman degradation using a gas phase microsequencer (Bielicki *et al.*, 1990). The amino acid sequence showed strong homology with the other sulphatases that have been isolated in our laboratory (Robertson *et al.*, 1988), as well as the sulphatases that had been published by other groups at the time this work was commenced, i.e. steroid sulphatase (Yen *et al.*, 1987) and arylsulphatase from sea urchin (Sasaki *et al.*, 1988). This strongly supported the hypothesis that the 42-kDa polypeptide was a sulphatase and was potentially derived from IDS.

This chapter describes the use of the 42-kDa polypeptide N-terminal amino acid sequence to design an oligonucleotide probe specifically to isolate the gene encoding the IDS protein. Initially, a cDNA and genomic library were screened with a 49-mer oligonucleotide probe which resulted in the identification of several positive clones. As the location of the IDS locus was known from linkage analysis (Upadhyaya *et al.*, 1988; Roberts *et al.*, 1989) this information was used to eliminate clones with other chromosomal locations. The clones were used to screen somatic cell hybrids containing different regions of the X-chromosome, thereby identifying clones derived from the same region reported for the IDS gene. The 49-mer oligonucleotide-positive clones that were isolated from the cDNA and whole human genomic libraries were shown to be derived from other chromosomal locations demonstrating that they did not encode the IDS protein. Subsequently, an enriched X-chromosome genomic

library was screened and a positive clone mapping to the same position as the IDS gene (Xq27/28) was identified. This chapter describes the isolation and identification of a 1.3-kb *Hin*dIII plasmid subclone from this genomic clone. When this clone was sequenced the 42-kDa N-terminal amino acid sequence was identified within the translated sequence. In addition, a potential initiating methionine and an ER signal peptide sequence were also found within the translation of this sequence.

Some of the work presented in this Chapter was conducted in collaboration with the Department of Cytogenetics and Molecular Genetics at the Adelaide Children's Hospital. The reason for their interest was based on the relatively close proximity of the IDS gene to the fragile X site (FRAXA) and other disease loci on the X-chromosome (Upadhyaya *et al.*, 1988; Roberts *et al.*, 1989). The Department of Cytogenetics had obtained a number of probes that were generated from this region of the X-chromosome (Xq27/q28), however their precise location in relation to the FRAXA site was undetermined. It was postulated that by the use of an IDS clone (reported to be located at the Xq27/28 boundary) a number of these probes could be mapped in relation to the FRAXA site (located at Xq27.3). This approach was successful, and the new order for markers found around the FRAXA site was subsequently reported as centromere-DXS98-DXS369-FRAXA-DXS296-IDS-DXS304-DXS374-telomere (Suthers *et al.*, 1989, 1990, 1991); see Section 5.3.3, Figure 5.2).

CY2 and CY3, somatic hybrid cell lines with various portions of the X-chromosome and chromosome 16, were two of the cell lines used in this study. It was postulated that, with the use of these cell lines, DNA clones isolated with the 49-mer oligonucleotide could be screened prior to determining the DNA sequence. If the isolated DNA clones hybridised to the same region of the X-chromosome as the IDS

gene, then the sequence characterisation of the DNA clones could be done to confirm whether they encoded IDS. Section 3.3.1 has the exact details of the chromosomal constitution of the somatic hybrid cell lines CY2 and CY3.

3.2 SPECIFIC METHODS

3.2.1 *IN SITU* HYBRIDISATION

All the in situ hybridisations carried out in this thesis were performed by Elizabeth Baker in the Department of Cytogenetics and Molecular Genetics, Adelaide Children's Hospital using various probes that are described in this thesis. DNA probes were the 1.3-kb HindIII described in this Chapter and the 1.5-kb EcoRI described in Chapter 4.

The labelled 1.3-kb *Hin*dIII genomic fragment was hybridised *in situ* according to Simmers *et al.* (1986), with denaturation of the chromosomes in 70% deionised formamide, 2 x SSC, at 70°C, pH 7.0, for 2 minutes and hybridisation of 0.4 μg/ml probe at 37°C. The slides were dipped in Kodak NTB-2 nuclear research emulsion diluted 2:1 with water, exposed for 39 days, developed and stained, and the metaphases were scored (Simmers *et al.*, 1988).

3.3 RESULTS AND DISCUSSION

3.3.1 49-MER OLIGONUCLEOTIDE DESIGN AND INITIAL cDNA LIBRARY SCREENING.

The N-terminal amino acid sequence of the 42-kDa polypeptide was used to select a region of low codon degeneracy to design a 49-mer oligonucleotide probe incorporating choices based on human codon usage (Masuyama *et al.*, 1986) (Figure 3.1). The decision to design a larger probe was based upon the experiences of other investigators in the Department of Chemical Pathology. In one case, the use of two overlapping 20-mer oligonucleotides consisting of mixtures to account for every codon choice, was successful in isolating a positive clone coding for glucosamine-6-sulphatase (Robertson *et al.*, 1988), whereas the same approach used for the isolation of the N-acetylgalactosamine-4-sulphatase gene was not as successful (T. Litjens, unpublished observations). It was thought that a larger probe based on codon usage rather than a mixture incorporating every codon choice would generate a more specific probe thereby improving the chances of isolating the IDS gene.

A random primed cDNA library derived from human colon cells was obtained from Clontech. The cDNA library was made using *Eco*RI linkers followed by cloning into the *Eco*RI site of λgt10. The average cDNA insert size was reported to be approximately 1.2-kilobases (kb) and the library was reported to contain 1.5 x 10^6 independent recombinant clones.

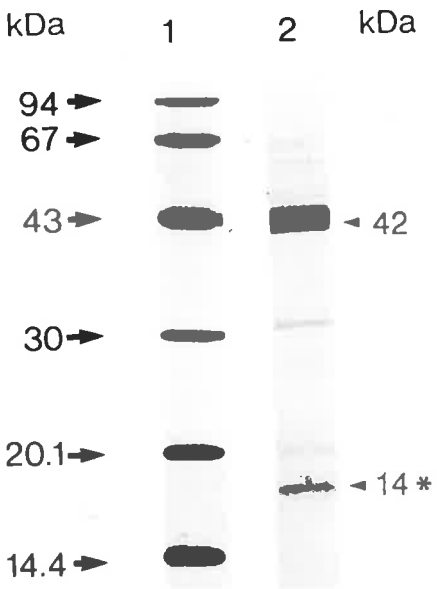
The cDNA library was plated-out as described in Section 2.2.4 at a density of 40,000 pfu per plate and 500,000 recombinant clones were screened. The plaques were

Figure 3.1 **SDS-PAGE OF IDS AND N-TERMINAL AMINO ACID SEQUENCE INDICATING THE POSITION AND DESIGN OF A 49-MER OLIGONUCLEO-TIDE**

The sample obtained from liver was reduced with dithioerythritol. **lane 1,** molecular-mass standards; **lane 2,** liver form A. The stain used was Coomassie Brilliant Blue.

**, Form A from liver was subjected to SDS-PAGE under optimum conditions for polypeptides with molecular masses under 20-kDa (Merle and Kadenbach, 1980). Bielicki et al. (1990) observed that the 14-kDa polypeptide appeared to run at 18-kDa when the conditions for resolving polypeptides under 20-kDa were sub-optimal.*

The N-terminal amino acid sequence of the 42-kDa and 14-kDa polypeptide species are displayed. The position of the 49-mer oligonucleotide is indicated by the amino acids in 3 letter code. The amino acids in bold type reflect errors between the sequenced and the predicted amino acids. The 14-kDa polypeptide species is discussed fully in Chapter 4.



42-kDa N-TERMINAL AMINO ACID SEQUENCE AND 49-MER OLIGONUCLEOTIDE DESIGN

T-**S**-A-L-N-V-L-Lys-Ile-Ile-Val-Asp-Asp-Lys-Arg-Pro-Ser-Lys-

5′-ACG TCG TCG TAG TCG CCC AGG GAG GGC CGC AGG-

-Gly-Cys-Tyr-Gly-Asp-**Tyr-Asp-Asp-Val-Lys-**

TCG TCC ACG ATG ATC A-3′

14-kDa N-TERMINAL AMINO ACID SEQUENCE

Ser-Pro-Arg-Glu-Leu-Ile-Ala-Tyr-Ser-**Gln**-Tyr-Pro-Arg-Pro-

transferred to nylon membranes and hybridised to the end-labelled 49-mer oligonucleotide (Section 2.2.3a). After 16 hours, the filters were removed from the hybridisation solution and washed in 2 x SSC at 42°C (the stringency used in this experiment was relatively low as the probe was degenerate). Two positive clones were identified after autoradiography which were then purified and DNA was made from plate lysates as described in Section 2.2.5. The DNA was digested with *Eco*RI to excise a 1.8-kb (λ2B1) and 0.56-kb (λ2B2)cDNA insert from the λgt10 clones (Figure. 3.2). Large-scale DNA digests were prepared to isolate sufficient quantities of cDNA insert for sub-cloning into the plasmid vector pUC19, and for making oligolabelled probes for Southern analysis of the somatic cell hybrid cell lines obtained from the Department of Cytogenetics and Molecular Genetics.

The location of the IDS gene has been identified by genetic linkage analysis and was reported to lie at the Xq27/28 interface (Mossman *et al.*, 1983; Upadhyaya *et al.*, 1986; Roberts *et al.*, 1989). In order to identify the chromosomal region that the cloned DNA was derived from, it was proposed to use DNA inserts to probe Southern blots with restriction digests of somatic cell hybrids which contained various regions of the X-chromosome. These somatic cell hybrids are generated by a reciprocal translocation between chromosome 16 and the X-chromosome. The somatic cell hybrids used in this study were CY2 (Xpter-Xq26; 16q24-qter) and CY3 (Xq26-qter; 16pter-16q24), i.e. CY3 contains the distal end of the long arm of the X-chromosome and CY2 contains the remainder of the X-chromosome. Other DNA samples used were A9, which contains only mouse DNA and human normal male and female DNA samples (Figure 3.3).

It was anticipated that this approach would enable the positive DNA clones to be screened before commencing DNA sequencing. Thus, if hybridising bands were not

Figure 3.2 **ELECTROPHORESIS OF *Eco*RI DIGESTED cDNA CLONES ISOLATED WITH THE 49-MER OLIGONUCLEOTIDE.**

Clones were prepared, purified and digested as described in Chapter 2 (Sections 2.2.5, 2.2.11 and 2.2.12). DNA was fractionated on a 1.0% agarose gel; DNA fragments were detected with ethidium bromide and visualised with UV. Sizes of cDNA inserts derived from each clone are indicated on the side of the gel

Lane 1: Molecular marker-λ *digested with Hin*dIII

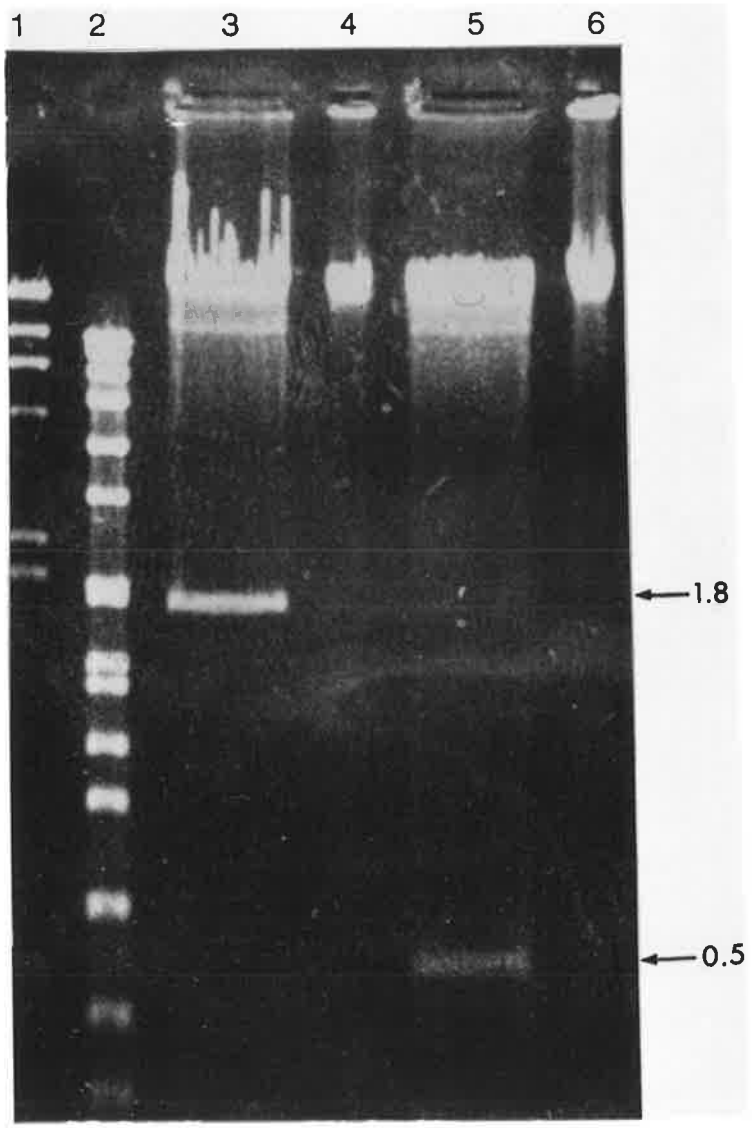
Lane 2: Molecular marker-SP1 phage digested with *Eco*RI.

Lane 3: λ2B1 digested with *Eco*RI.

Lane 4: Undigested λ2B1.

Lane 5: λ2B2 digested with *Eco*RI.

Lane 6: Undigested λ2B2.


1
2
3
4
5
6
1.8
0.5

found in the restriction digests of the DNA derived from CY3, it would be unlikely that the cDNA insert or genomic sub-clone (see Section 3.3.2 and 3.3.3) were derived from the distal portion of the X-chromosome. Therefore these clones would not be from the same region as the IDS gene.

The 1.8-kb cDNA insert of λ2B1 was oligolabelled and used to probe a Southern blot containing the digested somatic cell hybrid DNAs described above. In Figure 3.3, lanes 2 (CY2:Xpter-Xq26), 4 (CY3:Xq26-Xqter), 6 (A9-mouse DNA), 8 (whole female genomic DNA) and 10 (whole male genomic DNA) are digests with the restriction enzyme *Eco*RI. As is evident from the autoradiograph in Figure 3.4, four main bands are seen in lanes 2, 4 and 6 (8.5, 6.7, 6.3 and 5.2-kb respectively) without any corresponding bands in the whole human DNA digests. However, a single band was observed in the whole human DNA digests which was approximately 15-kb. This result indicates that the bands hybridising in the somatic cell hybrids was DNA which was homologous to a region within the mouse genome and the human-specific 15-kb band is not present on either end of the X-chromosome. The 15-kb band seen in the whole human digests is therefore derived from another human chromosome. It therefore appears that λ2B1 was not derived from the X-chromosome and did not require further characterisation. Also evident in this autoradiograph were the hybridising bands in the marker lane (λ digested with *Hin*dIII). Hence, the 1.8-kb cDNA insert is contaminated with pUC19 which contains sequences homologous with λgt10. Restriction digestion with *Sal*I (lanes 3, 5, 7, 9 and 11) was poor, due most likely to the fact that it is a methylation-sensitive enzyme. Therefore, the number of restriction sites available would be decreased because of methylation and the size of the DNA fragments produced would be much larger than for other 6 base recognition sequence enzymes, i.e. greater than 20-kb. These large DNA fragments were difficult to resolve using standard Southern analysis procedures. Figure 3.3 clearly

Figure 3.3 **SOMATIC CELL HYBRIDS DERIVED FROM THE X-CHROMOSOME AND DIGESTED WITH *Eco*RI and *Sal*I**

Somatic cell hybrid DNA was obtained from the Department of Cytogenetics and Molecular Genetics. DNA was digested and electrophoresed on a 1.0% agarose gel as described Chapter 2 (Section 2.2.12). DNA was stained with ethidium bromide and visualised under UV. The schematic at the base of the gel represents the areas covered by the somatic cell hybrids and indicates the position of the IDS gene.

Lane 1: Molecular marker-SP1 phage digested with *Eco*RI.

Lane 2: CY2(Xpter-Xq26; 16q24-qter) digested with *Eco*RI.

Lane 3: CY3(Xq26-qter; 16pter-16q24) digested with *Eco*RI.

Lane 4: A9 (Mouse genomic DNA) digested with *Eco*RI.

Lane 5: Human female genomic DNA digested with *Eco*RI.

Lane 6: Human male genomic DNA digested with *Eco*RI.

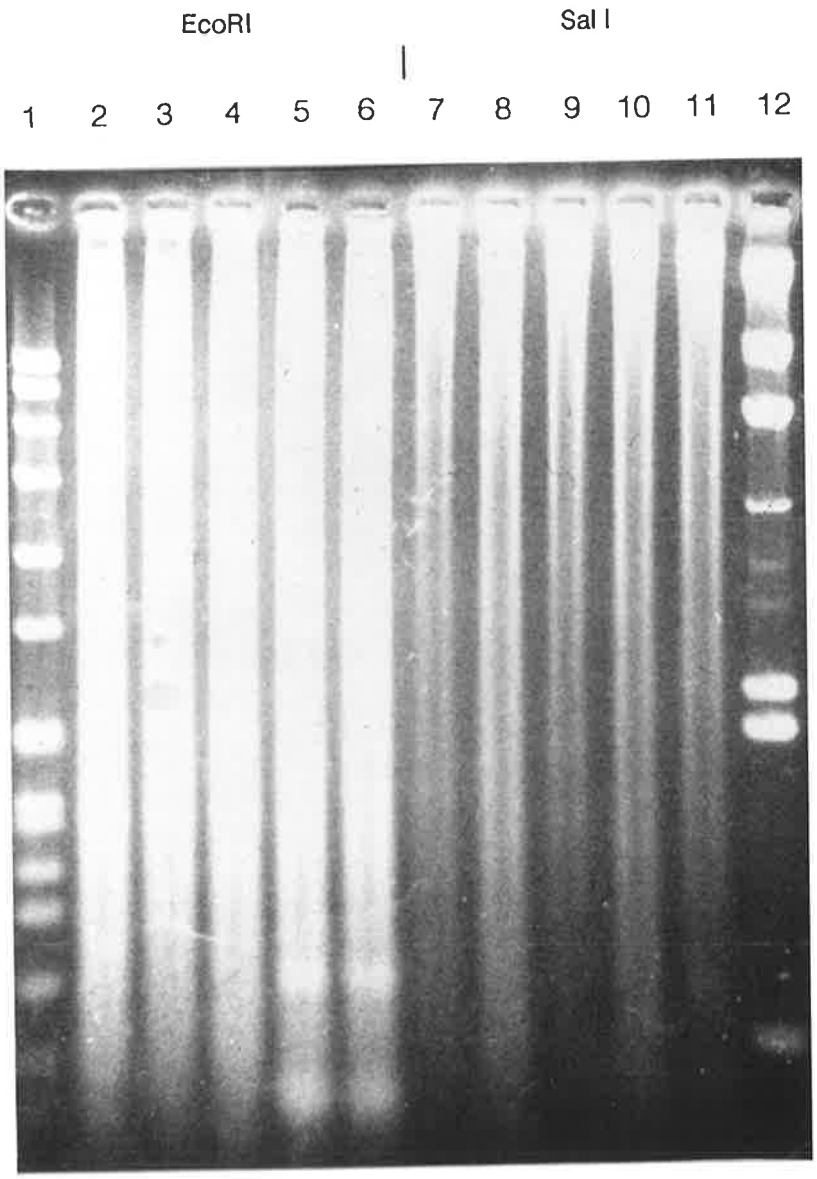
Lane 7: CY2 digested with *Sal*I.

Lane 8: CY3 digested with *Sal*I.

Lane 9: A9 digested with *Sal*I.

Lane 10: Female genomic DNA digested with *Sal*I.

Lane 11: Male genomic DNA digested with *Sal*I.


EcoRI
Sal I
1 2 3 4 5 6 7 8 9 10 11 12

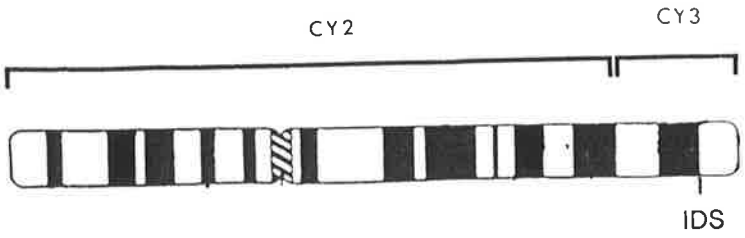

CY2
CY3
IDS

Figure 3.4 **AUTORADIOGRAPH OF SOMATIC CELL HYBRIDS PROBED WITH THE 1.8-kb cDNA INSERT OF λ2B1 ISOLATED WITH THE 49-MER OLIGONUCLEOTIDE**

A gel containing the same DNA samples as shown in Figure 3.3 (in a different order) was transferred to a nylon membrane and probed with the 1.8-kb cDNA insert derived from λ2B1. Each lane was loaded with a similar amount of DNA. However, due to the reduced number of restriction sites (due to methylation sensitivity), the *Sal*I digest appears poorly digested. The probe has also hybridised to some of the molecular weight markers (see text for explanation).

Lane 1: SP1 molecular markers.

Lane 2: CY2 digested with *Eco*RI.

Lane 3: CY2 digested with *Sal*I.

Lane 4: CY3 digested with *Eco*RI.

Lane 5: CY3 digested with *Sal*I.

Lane 6: A9 digested with *Eco*RI.

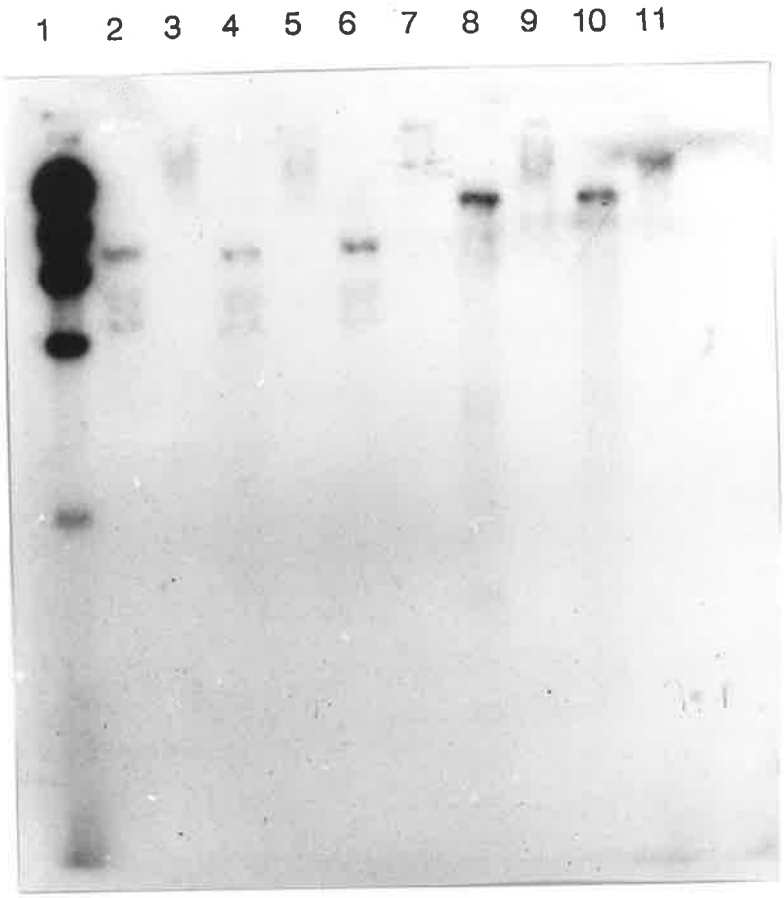
Lane 7: A9 digested with *Sal*I.

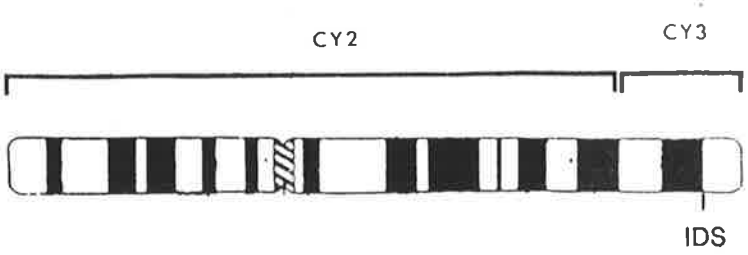
Lane 8: Female genomic DNA digested with *Eco*RI.

Lane 9: Female genomic DNA digested with *Sal*I.

Lane 10: Male genomic DNA digested with *Eco*RI.

Lane 11: Male genomic DNA digested with *Sal* I.


1 2 3 4 5 6 7 8 9 10 11


CY2
CY3
IDS

demonstrates the poor resolution of the larger DNA fragments, making the detection of hybridising fragments difficult (Figure 3.4). In hindsight, this was a poor choice of restriction enzyme and the use of other less methylation-sensitive enzymes would have produced a more satisfactory result, e.g. *Hin*dIII, *Bam*HI, StuI, *Pst*I, etc.

The 0.56-kb insert of λ2B2 was used to probe a Southern blot with *Eco*RI restriction digests of the same somatic cell DNA samples as indicated in Figure 3.4. Figure 3.5 is an autoradiograph of these somatic cell hybrids probed with the 0.56-kb cDNA probe. The bands seen in this gel (lanes 2, 3 and 4) again indicate mouse background bands (3.5, 2.7 and 0.6-kb). However, in the digest containing CY2 (Xpter-Xq26) there was a single band (1.6-kb) that had no corresponding bands in the mouse lane (lane 4) or in the whole human DNA digests (lanes 5 and 6). It is possible that the 0.5-kb cDNA insert is derived from the 1.6-kb fragment which, in turn, is derived from the 6.3-kb fragment seen in the whole human digests. The chromosome translocation disrupts the 6.3-kb fragment, leaving the 0.5-kb positive 1.6-kb fragment in the CY2 hybrid lane. Again, this indicates that the λ2B2 clone is either not derived from the X-chromosome or at least not from Xq27/q28 and is therefore unlikely to encode IDS.

Because of the obvious problems with these cDNA clones, a decision was made at this time to begin screening a whole human genomic DNA library. IDS is a low abundance protein and is therefore likely to have low mRNA levels. A cDNA library is therefore likely to contain IDS clones at a very low frequency thus requiring a larger number of clones to be screened whereas, in a genomic library, IDS will be present at the same level as other single copy genes.

Figure 3.5 **AUTORADIOGRAPH OF SOMATIC CELL HYBRIDS PROBED WITH A 0.56-kb cDNA INSERT ISOLATED WITH THE 49-MER OLIGONUCLEOTIDE.**

The 1.0% agarose gel described in Figure 3.3 was transferred to a nylon membrane and probed with the 0.56-kb cDNA insert derived from λ2B2.

Lane 1: CY2 digested with *Eco*RI.

Lane 2: CY3 digested with *Eco*RI.

Lane 3: A9 digested with *Eco*RI.

Lane 4: Female genomic DNA digested with *Eco*RI.

Lane 5: Male genomic DNA digested with *Eco*RI.

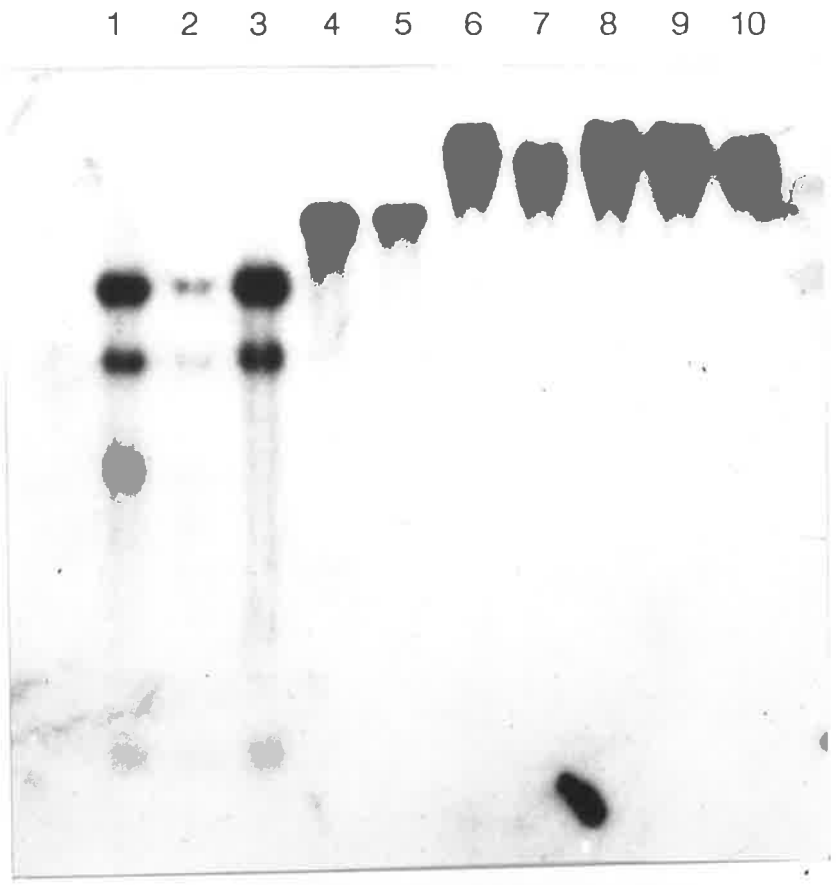
Lane 6: CY2 digested with *Sal*I.

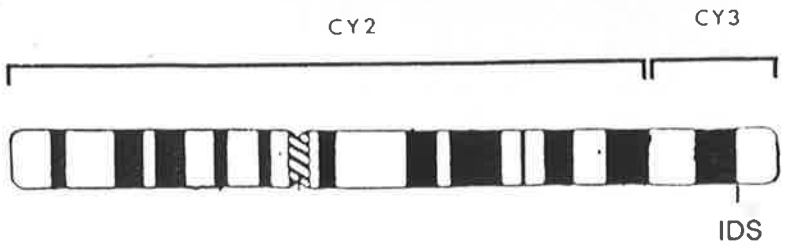
Lane 7: CY3 digested with *Sal*I.

Lane 8: A9 digested with *Sal*I.

Lane 9: Female genomic DNA digested with *Sal*I.

Lane 10: Male genomic DNA digested with *Sal*I.


1 2 3 4 5 6 7 8 9 10


CY2
CY3
IDS

3.3.2 GENOMIC LIBRARY SCREENING

A commercially available library containing human leucocyte genomic DNA which was partially digested with *Sau*3A and cloned into the *Bam*HI restriction site of λEMBL3 and reported to have an average insert size of 15-kb, was obtained from Clontech. The *Bam*HI restriction site is lost during cloning, however the insert can be excised with *Sal*I.

The genomic library was plated-out at a density of 24,000 pfu per plate and a total of 168,000 recombinants were screened with the end-labelled 49-mer oligonucleotide. This number of clones is approximately equivalent to one human genome.

Two positive clones were identified after purification to homogeneity but only one, designated HL-12, remained positive to the 49-mer oligonucleotide. HL-12 DNA was digested with a range of restriction enzymes including *Sal*I, *Eco*RI, *Bam*HI and *Hin*dIII and then analysed on a 0.8% (w/v) agarose gel. Digestion with *Sal*I indicated that an internal restriction site was present in the genomic insert, making isolation of the complete insert difficult. A range of restriction enzymes were used to identify the smallest 49-mer oligonucleotide hybridising fragment to simplify characterisation. In the *Hin*dIII digest of clone HL-12, no 49-mer oligonucleotide hybridising band of 1.3-kb was observed (data not shown). This fragment was later shown to contain the 49-kDa N-terminal amino acid sequence (see the following sections).

3.3.3 SCREENING AN ENRICHED X-CHROMOSOME GENOMIC LIBRARY

It was thought that the problems associated with the cDNA and whole human genomic libraries would be overcome if a library generated almost solely from X-chromosomes

was used, thereby increasing the likelihood of isolating the IDS gene. An enriched X-chromosome genomic library, LAOXNL01, was obtained from the American Tissue Culture Collection. This library was constructed from flow-sorted X-chromosomes and contained inserts which averaged 17-kb in size. The DNA was partially digested with the restriction enzyme *Sau*3A and cloned into the *Bam*HI restriction site of λ charon 35A.

The X-chromosome genomic library was plated-out at a density of 30,000 pfu per plate and a total of 220,000 recombinants (equivalent to approximately 20 X-chromosomes) were screened with the 49-mer oligonucleotide. Fourteen positive clones were identified and picked for further purification. All clones remained positive after the final stage of purification. Two clones were selected arbitrarily and designated λX1 and λX2; DNA was prepared and digested with a range of restriction enzymes including *Sal*I, *Hin*dIII, *Eco*RI, *Xho*I and these enzymes were also combined for double digests. Figure 3.6 shows the digested DNA fractionated on a 1.0% (w/v) agarose gel. The DNA from λX1 was cut to completion, whereas the DNA from λX2 gave an unusual pattern. At first it was thought that the DNA may have been overloaded, however the presence of both weak and intense restriction fragments suggested that this DNA preparation contained DNA from two λ clones, i.e. the clone was not pure. The agarose gel was transferred to a nylon membrane to enable the various restriction fragments to be probed with the 49-mer oligonucleotide in order to identify the smallest DNA fragment that would bind the 49-mer oligonucleotide so that it would be easy to sequence.

Figure 3.7 reveals that the smallest DNA fragment that hybridised to the 49-mer oligonucleotide was a 1.3-kb *Hin*dIII genomic DNA fragment. This fragment was present in all the genomic clones digested with this restriction enzyme. To enable the

Figure 3.6 **ELECTROPHORESIS OF TWO X-CHROMOSOME GENOMIC CLONES ISOLATED WITH THE 49-MER OLIGONUCLEOTIDE DIGESTED WITH A VARIETY OF ENZYMES.**

A 1.0% agarose gel was used to fractionate DNA derived from two separate X-chromosome genomic clones and digested with a range of restriction enzymes. The conditions for electrophoresis and restriction enzyme digestion are described in Chapter 2 (Sections 2.2.11 and 2.2.12).

Lane 1: λDNA digested with *Hin*dIII.

Lane 2: undigested λX1.

Lane 3: *Sal*I digested λX1.

Lane 4: *Hin*dIII digested λX1.

Lane 5: *Sal*I and *Hin*dIII digested λX1.

Lane 6: *Eco*RI digested λX1.

Lane 7: *Xho*I digested λX1.

Lane 8: *Eco*RI and *Xho*I digested λX1.

Lane 9: *Eco*RI and *Sal*I digested λX1.

Lane 10: undigested λX2.

Lane 11: *Sal*I digested λX2.

Lane 12: *Hin*dIII digested λX2.

Lane 13: *Sal*I and *Hin*dIII digested λX2.

Lane 14: *Xho*I digested λX2.

Lane 15: *Eco*RI digested λX2.

Lane 16: *Xho*I and *Eco*RI digested λX2.

Lane 17: *Sal*I and *Eco*RI digested λX2.

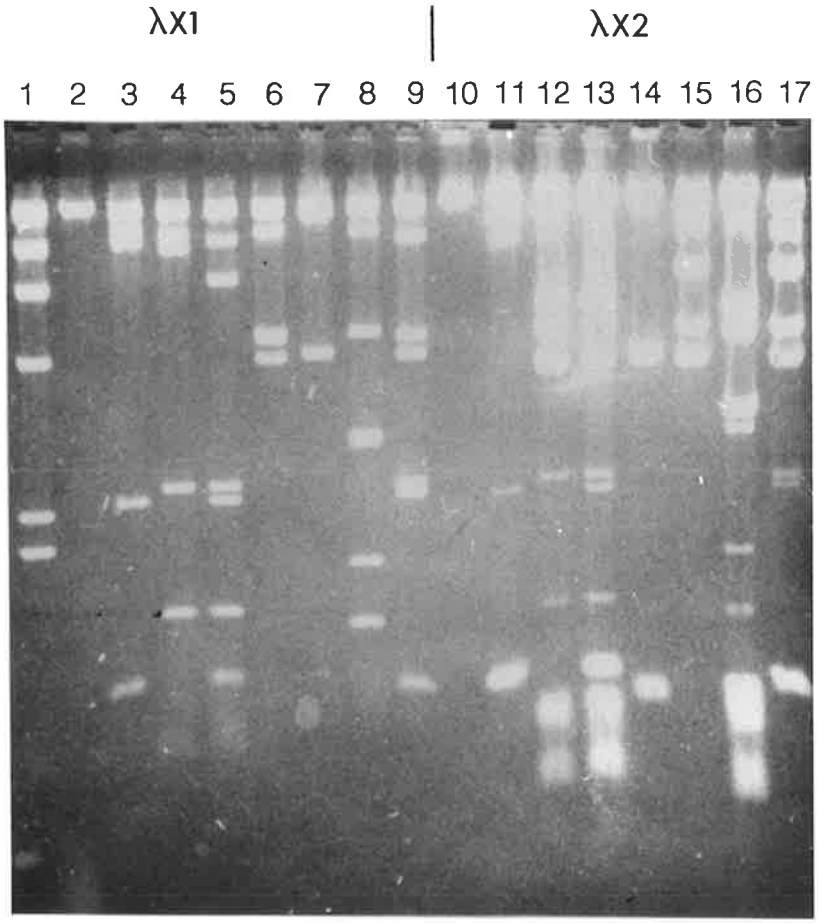

λX1
λX2
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17

Figure 3.7 AUTORADIOGRAPH OF THE DNA SAMPLES FROM FIGURE 3.6.

The gel in Figure 3.6 was transferred to a nylon membrane and probed with an end-labelled 49-mer oligonucleotide as described in Chapter 2 (Section 2.2.3a). Lane one from Figure 3.6 is not represented as it is the molecular weight marker also the DNA (undigested λX2) present in lane 10 of Figure 3.6 has not hybridized to the probe.

Lane 1: Undigested λX1.

Lane 2: *Sal*I digested λX1.

Lane 3: *Hin*dIII digested λX1.

Lane 4: *Sal*I and *Hin*dIII digested λX1.

Lane 5: *Eco*RI digested λX1.

Lane 6: *Xho*I digested λX1.

Lane 7: *Eco*RI and *Xho*I digested λX1.

Lane 8: *Eco*RI and *Sal*I digested λX1.

Lane 9: *Sal*I digested λX2.

Lane 10: *Hin*dIII digested λX2.

Lane 11: *Sal*I and *Hin*dIII digested λX2.

Lane 12: *Xho*I digested λX2.

Lane 13: *Eco*RI digested λX2.

Lane 14: *Xho*I and *Eco*RI digested λX2.

Lane 15: *Sal*I and *Eco*RI digested λX2.

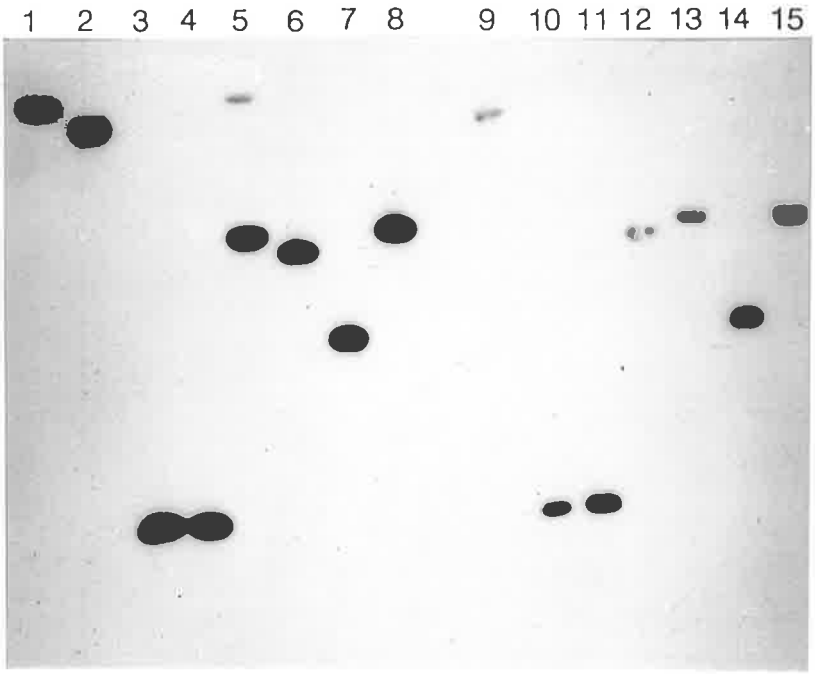

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

Figure 3.8 **AUTORADIOGRAPH OF SOMATIC CELL HYBRIDS PROBED WITH A 1.6-kb Hind III GENOMIC INSERT IDENTIFIED IN FIGURE 3.6.**

The same membrane used in Figure 3.3 was re-used and probed with an oligolabelled genomic fragment, as described in Chapter 2 (Sections 2.2.3a and 2.2.15)

Lane 1: CY2 digested with *Eco*RI.

Lane 2: CY3 digested with *Eco*RI.

Lane 3: A9 digested with *Eco*RI.

Lane 4: female genomic DNA digested with *Eco*RI.

Lane 5: male genomic DNA digested with *Eco*RI.

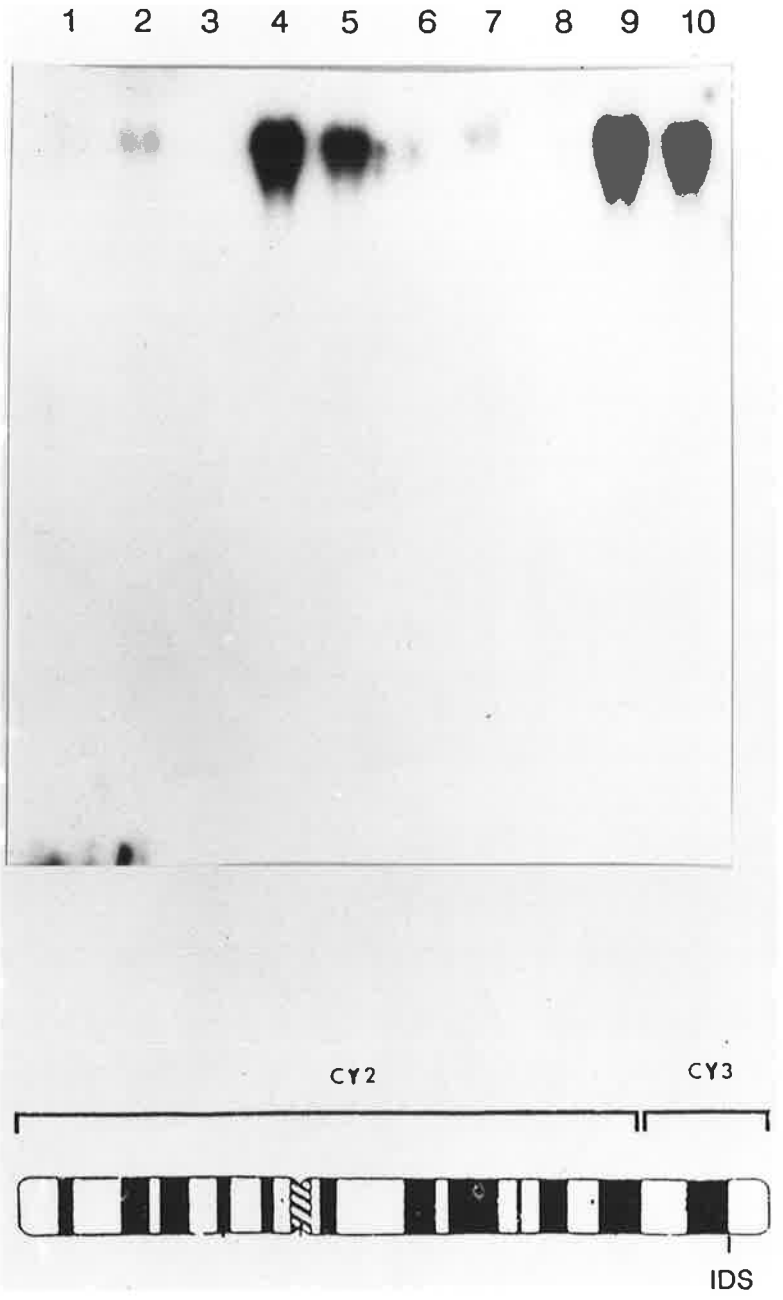
Lane 6: CY2 digested with *Sal*I.

Lane 7: CY3 digested with *Sal*I.

Lane 8: A9 digested with *Sal*I.

Lane 9: female genomic DNA digested with *Sal*I.

Lane 10: male genomic DNA digested with *Sal*I.


1 2 3 4 5 6 7 8 9 10
CY2
CY3
IDS

generation of large quantities of DNA insert, this fragment was purified by agarose gel electrophoresis and sub-cloned into the plasmid vector pUC19. The DNA insert was also sub-cloned into the sequencing vector M13mp19 to enable the nucleotide sequence of the genomic clone to be determined.

The genomic DNA insert was oligolabelled and used to probe a Southern transfer containing digests of DNA from the somatic cell hybrids CY2 and CY3. Figure 3.8 shows that the 1.3-kb *Hin*dIII genomic fragment hybridised to a band in the track containing DNA from CY3 (lane 2). This result indicated that the 1.3-kb *Hin*dIII genomic fragment was derived from the same region (Xq26-Xqter) that is known to contain the IDS gene (Roberts *et al.*, 1989). Hybridising bands were not seen in lane 1 (CY2; Xpter-Xq26) or lane 3 (mouse DNA). Lanes 4 and 5 contained DNA from a human female and male, respectively, and had the same size band as seen in the CY3 DNA. As would be expected, the intensity of the female band was approximately twice that of the male band. This was the first strong evidence that this clone contained the IDS gene. *In situ* hybridisation was performed on human metaphase chromosomes to more accurately localise the 1.3-kb HindIII genomic fragment and the results of this experiment are discussed in the following section.

3.3.4 *In Situ* HYBRIDISATION WITH THE 1.3-kb *Hin*dIII GENOMIC DNA FRAGMENT

The 1.3-kb *Hin*dIII genomic fragment was labelled with ^{3}H-thymidine and used to probe 30 metaphases from normal females (Figure 3.9A). *In situ* hybridisation established that there were no sequences homologous to the 1.3-kb *Hin*dIII genomic clone elsewhere in the human genome. This probe only detected sequences at distal Xq, confirming that only one copy existed in the genome which is consistent with what

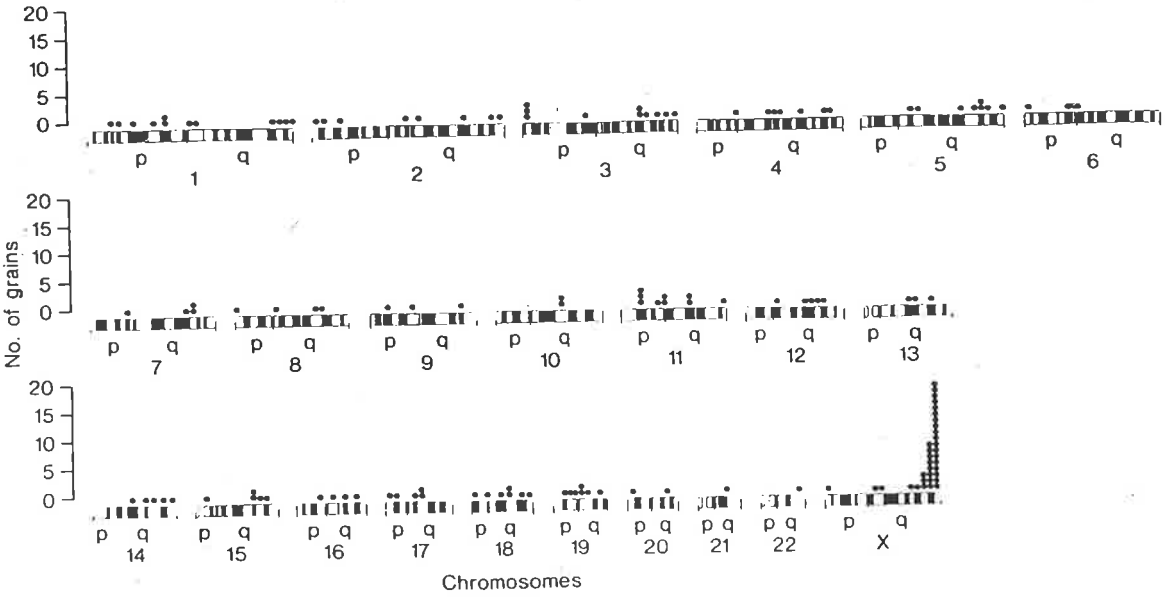
Figure 3.9 ***In Situ*** **HYBRIDISATION OF HUMAN METAPHASE CHROMOSOMES PROBED WITH THE 1.3-kb *Hin*dIII GENOMIC FRAGMENT**

The 1.3-kb *Hin*dIII genomic fragment was labelled and used to probe metaphase chromosomes as described in Section 3.2.1.

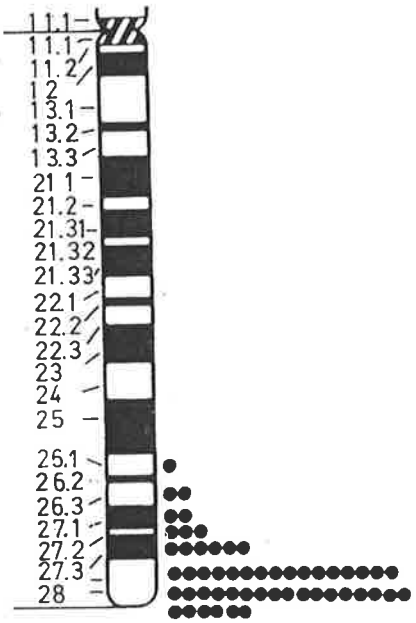
A: Ideogram of G-banded human chromosomes showing the distribution of silver grains after hybridisation of 1.3-kb genomic fragment to individual human metaphase chromosomes.

B: High resolution chromosomal banding of Xq (600-1,000 bands per metaphase).

A


No. of grains
20
15
10
5
0
p q
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
X
Chromosomes

B


11.1
11.1
11.2
12
13.1
13.2
13.3
21.1
21.2
21.31
21.32
21.33
22.1
22.2
22.3
23
24
25
25.1
26.2
26.3
27.1
27.2
27.3
28

is known about the IDS gene. Figure 3.9B represents an additional 40 metaphases with high-resolution chromosome banding of Xq (600-1,000 bands per metaphase) which were scored in this female and confirmed that the genomic clone was located at the boundary between Xq27 and Xq28 (Wilson *et al.*, 1991). This was consistent with the linkage data localising IDS to the boundary of Xq27/28 (Roberts *et al.*, 1989). The fact that the genomic clone hybridises to the same closely defined region of the X-chromosome as the known location for the IDS gene is positive evidence that it encodes IDS.

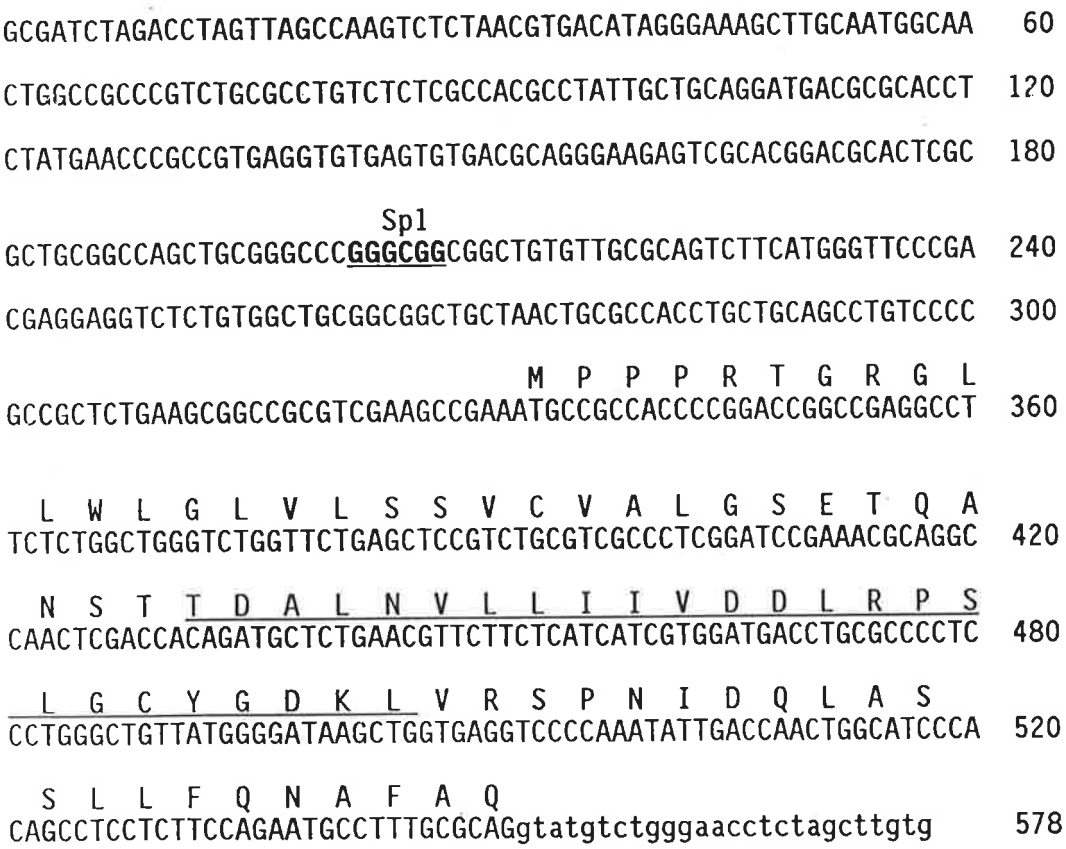
The next stage in the characterisation of the 1.3-kb *Hin*dIII genomic fragment was to obtain the DNA sequence of the clone to determine whether the 49-mer oligonucleotide binding site was present and to establish if this sequence, when translated, was co-linear with the 42-kDa polypeptide N-terminal amino acid sequence (see Figure 3.1). The partial DNA sequence of the 1.3-kb *Hin*dIII genomic fragment is discussed in the following section.

3.3.5 DNA SEQUENCE OF THE 1.3-kb *Hin*dIII GENOMIC FRAGMENT

The 1.3-kb *Hin*dIII genomic fragment was sub-cloned into the sequencing vector M13mp19 for dideoxy sequencing. Both orientations of this fragment were sequenced using the universal sequencing primer and an internal specific primer (designed after some of the sequence had been compiled). In Figure 3.10, the DNA sequence revealed the presence of the 49-mer oligonucleotide binding site and this sequence displayed co-linearity with the 42-kDa polypeptide N-terminal amino acid sequence. The errors found between the direct and predicted amino acid sequence data most likely reflect amino acid sequencing errors resulting from the low signal obtained at

Figure 3.10 **PARTIAL DNA SEQUENCE OF 1.3-kb *Hin*dIII GENOMIC FRAGMENT REVEALS THE 49-MER OLIGONUCLEOTIDE BINDING SITE**

The 1.3-kb *Hin*dIII genomic fragment was sub-cloned into the sequencing vector M13 and subjected to dideoxy sequencing as described in the Materials and Methods. Included in this Figure is the nucleotide sequence (indicated by the numbers on the right-hand side), a potential initiating methionine, the 42-kDa N-terminal amino acid sequence (49-mer oligonucleotide binding domain) indicated by an underline, a potential Sp1 binding site indicated by the bold type and underline; the lower case nucleotide sequence depicts the start of the first intron.


GCGATCTAGACCTAGTTAGCCAAGTCTCTAACGTGACATAGGGAAAGCTTGCAATGGCAA 60
CTGGCCGCCCGTCTGCGCCTGTCTCTCGCCACGCCTATTGCTGCAGGATGACGCGCACCT 120
CTATGAACCCGCCGTGAGGTGTGAGTGTGACGCAGGGAAGAGTCGCACGGACGCACTCGC 180
Sp1
GCTGCGGCCAGCTGCGGGCCCGGGCGGCGGCTGTGTTGCGCAGTCTTCATGGGTTCCCGA 240
CGAGGAGGTCTCTGTGGCTGCGGCGGCTGCTAACTGCGCCACCTGCTGCAGCCTGTCCCC 300
M P P P R T G R G L
GCCGCTCTGAAGCGGCCGCGTCGAAGCCGAAATGCCGCCACCCCGGACCGGCCGAGGCCT 360
L W L G L V L S S V C V A L G S E T Q A
TCTCTGGCTGGGTCTGGTTCTGAGCTCCGTCTGCGTCGCCCTCGGATCCGAAACGCAGGC 420
N S T T D A L N V L L I I V D D L R P S
CAACTCGACCACAGATGCTCTGAACGTTCTTCTCATCATCGTGGATGACCTGCGCCCCTC 480
L G C Y G D K L V R S P N I D Q L A S
CCTGGGCTGTTATGGGGATAAGCTGGTGAGGTCCCCAAATATTGACCAACTGGCATCCCA 520
S L L F Q N A F A Q
CAGCCTCCTCTTCCAGAATGCCTTTGCGCAGgtatgtctgggaacctctagcttgtg 578

the end of the amino acid sequencing run. Also found was a putative ER signal peptide region and a potential initiating methionine (von Heijne, 1986; Kozak, 1987) (see Section 4.3.4 for a more detailed discussion). The 42-kDa N-terminal amino acid sequence was also compared to the N-terminal amino acid sequence of other known sulphatases. Significant homology was found when the 80 amino acids contained in Figure 3.10 were compared to the three known human arylsulphatases, including arylsulphatase A (Stein *et al.*, 1989a), arylsulphatase B (Peters *et al.*, 1990), arylsulphatase C (Yen *et al.*, 1987; Stein *et al.*, 1989b), sea urchin arylsulphatase (Sasaki *et al.*, 1988) and human glucosamine-6-sulphatase (Robertson *et al.*, 1988) (see also Chapter 4, Section 4.3.6 for a complete discussion of homology with other sulphatases). Also indicated in Figure 3.10 is the first 3' intron/exon splice site which follows the consensus sequence for such sites exactly. This splice junction was determined after the complete cDNA sequence of IDS had been determined (Section 4.3.4, Figure 4.3).

To identify possible promoter or controlling elements of the IDS gene (Figure 3.10), a specific primer was designed from the sequence obtained from the 1.3-kb *Hin*dIII genomic fragment to obtain sequence from the 5'-untranslated region of the genomic clone. The promoter and controlling elements searched for include the TATA box (TATAAA) which is present in the majority of protein-encoding genes. Its function appears to be the accurate positioning of RNA polymerase relative to the initiation site (Benoist and Chambon, 1981). A CAP signal was also searched for, however the precise nature of this site has been poorly defined. The CAP signal indicates the start of the mRNA. Its structure has tentatively been described as the nucleotide sequence CA followed by a number of pyrimidines, several potential sites were found but the consensus is very loose. Some possible sites are at positions 221, 228 and 280, however primer extension would be required to determine which of these if any is the true CAP

site (Figure 3.10). No TATA box was identified in the 5'-untranslated sequence of the 1.3-kb genomic clone (Bucher and Trifonov, 1986).

The sequence from the 1.3-kb genomic fragment was also examined for the presence of a CCAAT box or a GC box. These represent a different class of transcription control elements that function at more variable distances and in either orientation; these signals have been termed 'upstream promoter elements' (UPEs) (Maniatis *et al.*, 1987). No CCAAT site was found, however one potential GC-box (GGGCGG) (Bucher, 1990) was found and is underlined at position 202 in Figure 3.10. The GC-box is recognised by transcription factor Sp1 and typically occurs in house-keeping genes (Kadonga *et al.*, 1986, Dynan, 1986). This is consistent with what is known about the IDS gene in that it is apparently expressed at low levels in all cells (see Section 4.3.7b for a more detailed discussion). The 5'-untranslated sequence is also very GC-rich, another feature typical of house-keeping genes.

Although the 49-mer oligonucleotide-binding sequence was identified and the amino acid sequence was found to be co-linear with the 42-kDa polypeptide N-terminal amino acid sequence and had homology with other known sulphatases, it was still not possible to conclusively prove that the genomic clone encoded IDS. In order to establish the authenticity of this genomic clone the complete IDS protein sequence was required. Therefore, a cDNA library was screened with the 1.3-kb genomic fragment and the positive clones were used to probe MPS II patient DNA which revealed the presence of structural alterations to IDS. The cDNA was also used in expression vectors and the product was shown to have specificity towards specific IDS substrates (see the following two Chapters for a full discussion of these results). These results indicated that the cDNA clones encoded IDS and that the 1.3-kb *Hin*dIII genomic clone encoded a portion of the N-terminus of IDS.

CHAPTER FOUR

ISOLATION AND EXPRESSION OF FULL-LENGTH IDS cDNA

4.1 INTRODUCTION

The results presented in Chapter 3 indicated that the genomic clone λX1 was likely to be derived from the IDS gene. The 1.3-kb *Hin*dIII genomic fragment which was a sub-clone of λX1 contained the 42-kDa polypeptide N-terminal amino acid sequence, and probing with various X-chromosome somatic cell hybrids and *in situ* hybridisation localised this fragment to the same region of the X-chromosome as that reported for IDS (Xq27/28). However, before being confident that the IDS gene had been cloned, the complete coding region of IDS would be required to predict the primary amino acid structure of IDS. The amino acid sequence could be used to demonstrate homology with other known sulphatases in order to show the isolated clone encoded a sulphatase gene.

The 1.3-kb *Hin*dIII genomic fragment that was believed to contain a part of the 5'-region of the IDS gene was considered to be an ideal probe for screening a cDNA library to obtain full-length IDS clones. The problems of short probes were highlighted in Chapter 3, where the use of a small degenerate oligonucleotide probe to screen cDNA libraries was found to be unsuccessful in isolating a genuine IDS cDNA clone. With the use of a much larger unique probe the screening specificity would be improved by use of increased washing stringency and the availability of greater probe specific activity.

This Chapter describes the use of the 1.3-kb genomic probe to screen a random primed cDNA library derived from human colon cells. Several positive IDS clones were isolated, the longest of which was sequenced and was found to code for a protein that was much larger than the expected 42-kDa polypeptide. However, this clone was not full-length since it did not contain a stop codon. An endothelial cDNA library was

screened using a restriction fragment from the 3'-end of this clone as a probe and a full-length clone was obtained. It was established that the N-terminal amino acid sequence of the 14-kDa polypeptide species seen on SDS-PAGE (see Figure 3.1) was encoded toward the 3'-end of the cDNA sequence. This accounted for the discrepancy in size of the protein determined by SDS-PAGE and the predicted protein size when the cDNA sequence was translated. This cDNA clone was also used to probe a Northern blot containing poly(A)$^+$ RNA isolated from human placenta, and several mRNA species were identified.

The identity of the cDNA clone as IDS was confirmed in two ways. Firstly, the cDNA was used as a probe to screen MPS II patient DNA and several structural alterations were detected (this is discussed in Chapter 5) which provided conclusive proof that the cDNA encoded IDS. The cDNA was also used to construct an expression vector and the ability of the expressed product to degrade specific IDS substrates indicated that the cDNA was full-length. The predicted amino acid sequence was found to have significant homology with other known sulphatases, primarily in the N-terminal-third of the 42-kDa polypeptide. The cDNA sequence was also found to contain potential N-glycosylation sites and an ER signal peptide sequence. Also, a single potential polyadenylation site was identified in the 3'-untranslated sequence of the full-length cDNA clone. The characterisation of these IDS cDNA clones is discussed in the following sections.

4.2 SPECIFIC METHODS

4.2.1 PCR FROM LAMBDA DNA

Essentially a small quantity of pure high titre λ phage (1-5 μl) was placed in a reaction tube along with PCR buffer (50 mM KCl; 10 mM Tris-HCl, pH 8.4, 2.5 mM $MgCl_2$; 0.02% (w/v) gelatin, 0.1% (w/v) Triton X-100), 2 units of Taq polymerase, 0.2 mM dNTP's, and 0.5 μg of λgt10 or λgt11 forward and reverse primers. Thirty cycles of 95°C for 2 minutes, 50°C for 2 minutes, and 65°C for 3 minutes were performed. The products were electrophoresed and analysed on a 1.0% (w/v) agarose gel.

4.2.2 LINEAR PCR OF IDS M13 SUB-CLONES

One microgram of an M13 sub-clone containing the 1.5-kb IDS cDNA clone in the reverse-orientation was added to 50 mM KCl, 10 mM Tris-HCL, pH 8.4, 2.5 mM $MgCl_2$, 0.02% (w/v) gelatin, 0.1% (v/v) Triton X-100, 2 units of Taq polymerase, 0.5 ug of universal sequencing primer (17-mer) and 0.2 mM dNTP's. Thirty cycles of 94°C for 1 minute, 55°C for 2 minutes, and 70°C for 3 minutes were performed. The single-stranded template was then purified on a 1.0% (w/v) agarose gel according to Section 2.2.14.

4.2.3 LINEAR PCR SEQUENCING

The shotgun cloning and sequencing approach used in this work obtained approximately 95% of the 1.5-kb cDNA sequence in both strands. To complete the DNA sequence in both directions a different approach was used. The single-stranded PCR template generated in Section 4.2.2 was used with specific cDNA primers that were kinased (Section 2.2.3a) and used in normal dideoxy sequencing reactions

(Section 2.2.21). This approach was able to complete the 1.5-kb cDNA sequence in both strands.

4.2.4 NORTHERN RNA ANALYSIS

Northerns were performed basically as described by Ausubel *et al.* (1989). RNA was prepared according to Section 5.2.2. The RNA was fractionated by denaturing agarose gel electrophoresis using formaldehyde as the denaturant. The fractionated RNA was transferred to Genescreen Plus nylon membranes according to the manufacturer's instructions (NEN-Dupont). The size of IDS mRNA species was determined by probing this filter with labelled 1.5-kb cDNA insert. The size standard used in this experiment was the BRL (Bethesda Research Laboratories) 0.24-9.5-kb RNA ladder.

4.2.5 GENOMIC DNA PREPARATION

Genomic DNA was prepared from either periperal blood leucocytes or cultured skin fibroblasts. Essentially, the cells were disrupted by lysis in a Triton X-100 solution (0.32 M sucrose, 10 mM Tris-HCl, pH 7.5, 5 mM $MgCl_2$, 1.0% (w/v) Triton X-100) and the majority of protein was digested with proteinase K. The DNA was precipitated with buffered phenol and washed with chloroform/isoamyl alcohol followed by a 70% (v/v) aqueous ethanol wash. Finally the DNA was dissolved in TE buffer and stored at -80°C. At all stages extreme care was taken to avoid shearing forces to ensure that the DNA remained high molecular weight.

4.2.6 EXPRESSION OF IDS IN CELL CULTURE

4.2.6a ELECTROPORATION OF CHO-DK1 CELLS

CHO-DK1 cells were maintained in Ham's F12 medium supplemented with antibiotics (penicillin, kanamycin and streptomycin) and 10% (v/v) fetal calf serum (FCS) at 37°C in a 5.0% CO_2 atmosphere. Cells were fed with fresh medium every 2-3 days. Cells for electroporation were harvested by trypsinisation, washed once with PBS and then resuspended in PBS at 1.2 x 10^7 viable cells/ml. After equilibration to 0°C, 0.8 ml of cells were electroporated in the presence of 20 μg of plasmid DNA using a BRL Cell-Porator and a pulse of 275 Volts at 330 μF. The cells were grown in non-selective medium for 48 hours and subcultured 1:5 into medium containing 0.375 mg of neomycin (G418)/ml. The selected cells were maintained in medium containing 0.185 mg of G418 /ml.

4.2.6b DETERMINATION OF IDS EXPRESSION

The measurement of IDS enzyme activity in all cases was carried out by Julie Bielicki in the Department of Chemical Pathology.

Cell lysates were prepared by five cycles of freeze/thawing in 0.1% (v/v) Triton X-100 and clarified by centrifugation at 12,000 rpm for 5 min at 4°C. IDS activity was assayed as previously described (Hopwood, 1979) with the substrate L-O-(α-iduronic acid 2-sulphate)-(1>4)-D-O-2,5-anhydro[1-^{3}H]mannitol-6-sulphate (IdoA2S-anM6S) but modified as described below. The total assay volume of 12 μl included 1 μl of 340 μM-IdoA2S-anM6S (specific radioactivity 2,582 cpm/pmol), 3 μl of 0.2 M sodium-

acetate, pH 4.0, containing 2 mg of BSA/ml, 10 mM-$CuCl_2$ and 3 mM NaN_3, and up to 8 μl of the cell lysate or media was diluted in 50 mM sodium acetate, pH 4.0, or dialysed against 50 mM sodium acetate, pH 4.0, containing 10% (v/v) glycerol and 0.1 mM dithioerythritol. Assay tubes were incubated at 37°C for 1-2 hours to give less than 30% breakdown of substrate to product. The reaction was terminated by freezing, and the substrate and product were separated on Whatman 3MM chromatography paper in 0.75 M formic acid, pH 1.7, by using high voltage electrophoresis at 45 V/cm for 30 min on a Shandon Southern model L-24 system (Shandon Southern Products, Runcorn, Cheshire, UK). The chromatography paper was air-dried and scanned on a Packard model 7201 radiochromatogram (Packard, Chicago, Il., USA). Areas of radioactivity (9-10 cm from the origin for the product and 14-15 cm for the substrate) were cut from the strip and placed in 5 ml plastic vials and eluted in 1.5 ml of water. Three ml of Beckman EP ReadySolv scintillant was added and their radioactivities were determined in an LKB Rackbeta II liquid-scintillation counter. Enzyme activity was calculated from the percentage conversion of substrate into product and was expressed as picomoles of product/min/μl for the results obtained see Section 4.3.7.

4.3 RESULTS AND DISCUSSION

4.3.1 SCREENING A COLON cDNA LIBRARY WITH THE 1.3-kb *Hin*dIII GENOMIC FRAGMENT.

The 1.3-kb HindIII genomic fragment which was derived from λX1 was selected as a probe to screen cDNA libraries (see Section 3.3.3). This DNA insert was oligolabelled using α-^{32}P-dCTP according to the procedure outlined in Section 2.2.3b and the efficiency of the labelling was judged by running the sample on a Sephadex G-50 column as described in Section 2.2.3c. The labelled DNA insert was used to screen a human colon cDNA library according to Section 2.2.4. Approximately 5 x 10^5 recombinants were screened at a density of 55,000 pfu per plate using the bacterial host C600. The washing conditions were as described in Section 2.2.16 (Chapter 2). After washing, the filters were autoradiographed. When the film was developed 18 primary positive cDNA clones were identified. The number of positive clones isolated was much greater than would be predicted given the low abundance of IDS. In general low abundance proteins also have low levels of mRNA. Hence, a cDNA library generated using total RNA would be expected to contain low numbers of IDS cDNA clones, assuming that the IDS mRNA level was also low.

Seven cDNA clones were randomly selected and purified to obtain DNA from plate lysates according to Section 2.2.5. Usually, the cDNA insert size would be determined by digesting the DNA with the restriction enzyme *Eco*RI, as the cDNA inserts were cloned using *Eco*RI linkers (see also Section 3.3.1 for a description of this cDNA library). However, at the time that this work was done the polymerase chain reaction (PCR) became routine in the laboratory. It was decided to use λgt10 primers, designed at the borders of the λ vector arms to amplify the cDNA insert (for the conditions used see Section 4.2.4). The amplified cDNA inserts were electrophoresed

on a 1.0% agarose gel (the results of this experiment can be seen in Figure 4.1A). The largest cDNA insert was estimated to be 1.5-kb and was isolated from the clone designated λc2S15. The insert was prepared from the purified λ DNA and sub-cloned into the plasmid pUC19 to enable the insert to be grown in sufficient amounts for further sub-cloning and for use as a probe. This clone was designated as p2Sc15. The cDNA insert was cloned into the sequencing vector M13mp19 to enable the determination of the DNA nucleotide sequence using the dideoxy chain termination method (Sanger *et al.*, 1977). The 1.5-kb cDNA fragment was used to probe somatic cell hybrids described in Section 3.3.1, as well as for *in situ* hybridisation. This work was done in collaboration with the Department of Cytogenetics and Molecular Genetics. The results obtained in both cases were identical to the results found with the 1.3-kb *Hin*dIII genomic fragment (data not shown).

4.3.2 ANALYSIS OF THE DNA SEQUENCE CONTAINED IN THE 1.5-kb *Eco*RI cDNA INSERT.

The analysis of the 1.5-kb cDNA sequence was done in collaboration with Dr. Don Anson and Teresa Occhiodoro in our group.

Several methods can be applied to the analysis of cDNA inserts. One common method is to digest the DNA insert with several four-base specificity restriction enzymes such as *Hae*III, *Alu*I, *Taq*I, etc., to generate a range of different sized DNA fragments. These fragments are then cloned individually into M13 sequencing vectors. The rationale for using a range of restriction enzymes is to obtain clones with overlapping sequence so that the various DNA clones can be linked together to form one continuous DNA sequence. Another way of generating DNA sequence is to apply the same method as described above, but to randomly clone the DNA. This is often

Figure 4.1 **ELECTROPHORSIS OF PCR-AMPLIFIED DNA DERIVED FROM INDEPENDENT cDNA CLONES.**

Purified DNA from clones isolated from two different cDNA libraries was subjected to PCR amplification using the forward and reverse primers specific for λgt10 and λgt11 cDNA libraries as described in Section 4.2.1.

A: PCR amplification from clones isolated from a human colon λgt10 cDNA library. The largest identified cDNA insert was 1.5-kb.

Lane 1: Molecular markers-SP1 digested with *Eco*RI.

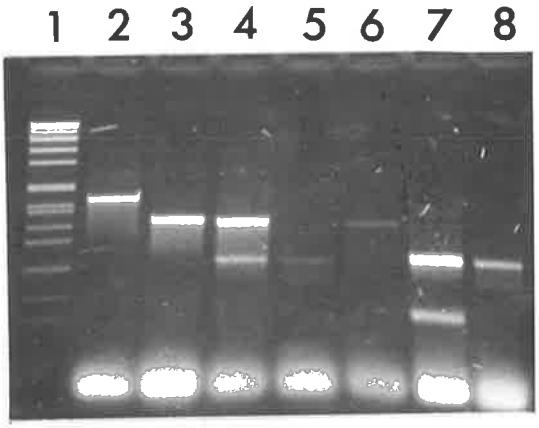
Lane 2 to 7: PCR amplified λgt10 cDNA clones

B: PCR amplification of clones isolated from a human endothelial λgt11 cDNA library. The largest cDNA insert shown here was 2.3-kb.

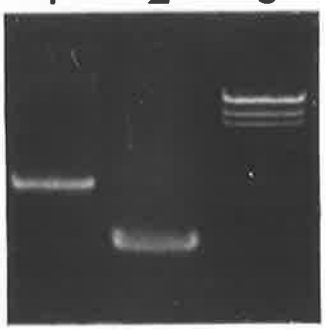
Lanes 1 and 2: PCR-amplified λgt11 cDNA clones

Lane 3: molecular marker-SP1 digested with EcoRI.

A



B



referred to as 'shotgun' cloning of restriction fragments. There are advantages and disadvantages to both of these methods. One disadvantage of the later method is that, although both methods enable the rapid generation of DNA sequence data, not all the DNA sequence will be covered by overlapping clones. This makes the remaining DNA sequence difficult to analyse. Another possible way of generating DNA sequence is to ligate the DNA insert to make a long polynucleotide chain of repeat units which can then be sheared with sonication. This method generates a pool of DNA fragments with different sizes and different regions of overlap. These fragments can then be randomly cloned into an M13 sequencing vector. It was decided to follow this approach as the generation of randomly overlapping clones was thought to have many advantages compared to the sub-cloning of restriction enzyme fragments.

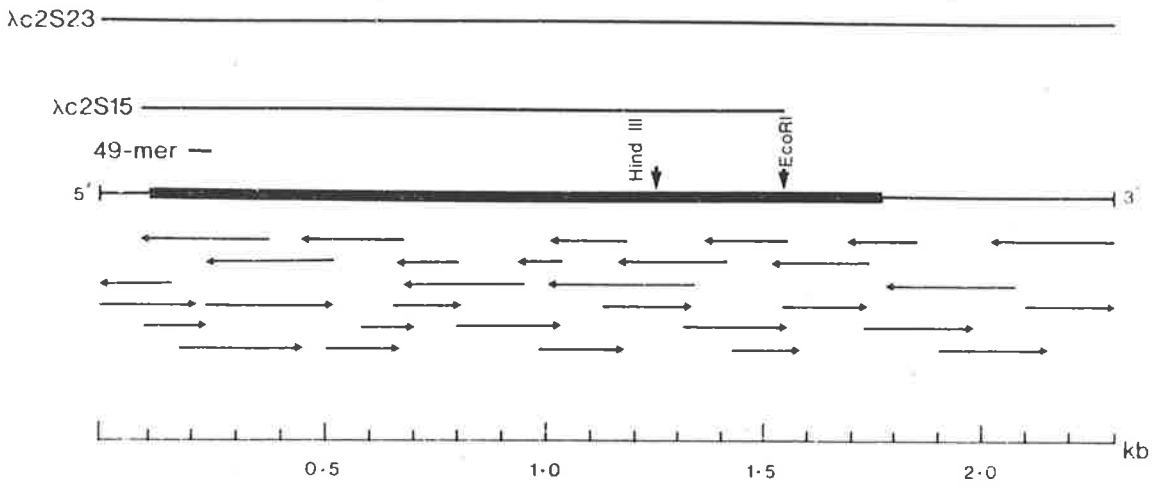
Sonicated DNA fragments from the 1.5-kb cDNA insert were sub-cloned into M13mp19 for nucleotide sequence analysis by the dideoxy-nucleotide chain-termination method using the Klenow fragment of DNA polymerase I (Sanger *et al.*, 1977) (for a detailed description of the method used see Section 2.2.21). The majority of the 1.5-kb cDNA sequence was compiled in this manner using the DNA sequence analysis program DBAUTO which joins overlapping sequences from both strands of DNA (Section 2.2.22). Using this program, two non-overlapping sequence contigs were generated. Specific primers were designed at the 3'-end of one of these DNA contigs and used to sequence an M13mp19 clone containing the complete 1.5-kb cDNA insert in the reverse-orientation. The cDNA-specific primer was able to join the two contigs, however there remained some areas of sequence that were incomplete in both strands. Sequence from both DNA strands is required to ensure the sequence obtained is completely accurate. This is especially true of areas that are G+C rich and are therefore especially prone to sequencing artefacts. To complete the sequence of the 1.5-kb cDNA the universal sequencing primer was annealed to an M13 clone

containing the entire 1.5-kb cDNA insert. Single-sided PCR was performed to generate single-stranded 1.5-kb cDNA insert (for the method used see Section 4.2.2). End-labelled specific cDNA sequence primers were used to anneal to the single-stranded PCR template and normal dideoxy sequencing reactions were performed. Using this approach the sequence of the 1.5-kb cDNA insert was completed in both strands. Figure 4.2 indicates the sequencing strategy employed for the 1.5-kb cDNA sequence and also includes the sequence derived from the 2.3-kb endothelial cDNA insert; this clone is discussed in Section 4.3.3.

The sequence contained the initiating methionine and a continuous open reading frame but did not contain a termination codon. The translated 1.5-kb cDNA sequence encoded a protein with a much larger molecular weight than the 42 kDa polypeptide seen on SDS-PAGE. A number of possibilities could explain this discrepancy, i.e. the molecular weight estimation from SDS-PAGE may be in error, the 42-kDa polypeptide could have been proteolytically processed from a precursor protein, or an additional polypeptide species may be required to complete the IDS enzyme. In Chapter 3, the SDS-PAGE gel of purified IDS (Figure 3.1) revealed multiple polypeptide bands, apart from the 42-kDa polypeptide which was consistently present as the major and most abundant polypeptide. The next most abundant and consistently observed species was a 14-kDa polypeptide species which was also seen in both Form A and B of the protein purifications from several tissue sources (liver, kidney, lung, placenta). The 14-kDa polypeptide was isolated and subjected to N-terminal amino acid sequencing (Bielicki *et al.*, 1990). The nucleotide sequence of the 1.5-kb cDNA was examined and the N-terminal amino acid sequence of the 14-kDa polypeptide was found towards the 3'-end of the 1.5-kb cDNA insert. This indicated that the cDNA clone contained the 42-kDa polypeptide positioned N-terminal to the 14-kDa polypeptide.

Figure 4.2 **SEQUENCING STRATEGY OF IDS cDNA CLONES DEPICTING THE RELATIONSHIP OF THE TWO CLONES.**

The majority of the sequence from the 1.5-kb cDNA insert was generated from sonicated sub-clones and sequenced with the dideoxy chain termination method with the universal sequencing primer as described in Section 2.2.21. The remainder was sequenced from linear PCR templates with specific primers (see Section 4.2.3). The sequence of the 2.3-kb cDNA insert (the last 800-bps is not found in the 1.5-kb cDNA insert) was derived entirely from sub-clones in M13 using and a combination of the universal sequencing primer and specific primers. Shown in the Figure is the relationship of the 1.5-kb and 2.3-kb cDNA insert. Also indicated is the position of the 49-mer oligonucleotide and the 300-bp *Hin*dIII/*Eco*RI restriction fragment derived from the 1.5-kb cDNA insert and used to isolate the 2.3-kb cDNA insert. The black boxed area represents the coding region and the arrows represent the direction and extent of sequencing reactions.


λc2S23
λc2S15
49-mer
Hind III
EcoRI
5'
3'
0·5
1·0
1·5
2·0
kb

The lack of a termination codon required the screening of another cDNA library to isolate a full-length cDNA clone. In order to obtain the remaining 3'-sequence it was decided that the best approach would be to use a restriction fragment derived from the 3'-end of the 1.5-kb cDNA insert as a probe. A 300-bp *Hin*dIII-*Eco*RI restriction fragment was identified at the 3'-end of the 1.5-kb cDNA clone and this was oligolabelled and used to screen a human endothelial cDNA library. This resulted in the isolation of a full-length IDS cDNA clone.

4.3.3 ISOLATION OF A 2.3-kb cDNA CLONE FROM AN ENDOTHELIAL cDNA LIBRARY

To obtain a full-length IDS cDNA clone, a 300-bp *Hin*dIII-*Eco*RI restriction fragment was oligolabelled with α-^{32}P-dCTP and used to screen a commercially available λgt11 endothelial cDNA library (Clontech). The endothelial library was selected based on the success in obtaining full-length cDNA clones encoding other low abundance lysosomal enzymes of interest to the laboratory when this library was used. This library was generated by cloning endothelial cell cDNA into the *Eco*RI site of λgt11. The average insert size of this library was 0.9-kb, and the library was reported to contain at least 2.1 x 10^6 independent recombinant clones.

Approximately 5 x 10^5 recombinants were screened at a density of 55,000 pfu per plate using the bacterial host NM538 (for the method used see Section 2.2.4). Twenty-seven clones were identified, 5 of which were also positive to the 49-mer oligonucleotide probe, indicating that these clones were likely to contain additional sequence when compared with the 1.5-kb cDNA clone. The cDNA inserts contained in these 5

endothelial clones were generated using the same PCR procedure employed with the 1.5-kb insert. This experiment indicated that the largest cDNA insert was 2.3-kb in size. This clone was designated λc2S23 (Figure 4.1B). Due to the many reports concerning the lack of fidelity of Taq polymerase, the purified λc2S23 DNA was digested with *Eco*RI to obtain a DNA insert suitable for sub-cloning into a plasmid vector. Two fragments were generated indicating either that the clone was not sufficiently pure or that there was an internal *Eco*RI site. The purification strategy was checked and appeared to be satisfactory. This left the second possibility as the likely explanation for the presence of two DNA fragments. Figure 4.2 indicates that the 3'-end of the 1.5-kb cDNA sequence ends in an *Eco*RI restriction site. At first this was thought to be part of the linker used to generate the cDNA library, but sequencing across the boundary between the 1.5-kb and 2.3-kb cDNA clones now shows that this restriction site is part of the IDS cDNA sequence. This suggests that the λgt10 colon cDNA library used to isolate the 1.5-kb cDNA clone was only partially methylated and may explain why a full-length cDNA clone was not isolated. The purified endothelial cDNA clone, λc2S23, was used as a template in a PCR reaction using the λgt11 forward and reverse primers (see Section 4.2.1). A single product of 2.3-kb in size was observed, indicating that the clone was both pure and that the *Eco*RI restriction site was part of the IDS cDNA sequence (Figure 4.1B).

The sequencing strategy employed in the nucleotide sequence characterisation of both the 2.3-kb cDNA insert and the 1.5-kb cDNA insert is indicated in Figure 4.2. Also shown in this diagram is the relationship of the two cDNA inserts and the position of the 49-mer oligonucleotide. A number of methods were used to generate the complete IDS cDNA sequence. Firstly, as was indicated in the previous section, the majority of the 1.5-kb cDNA sequence was compiled by random-sheared and 'shotgun'-cloned DNA. Secondly, areas not covered in both directions were completed by using linear

PCR-derived template and specific primers. Thirdly, the majority of the remaining 3' sequence derived from the 2.3-kb endothelial cDNA (λc2S23) was generated using restriction fragment M13 sub-clones and sequenced with the universal sequencing primer. The detailed description of the complete cDNA sequence will be discussed in the following section.

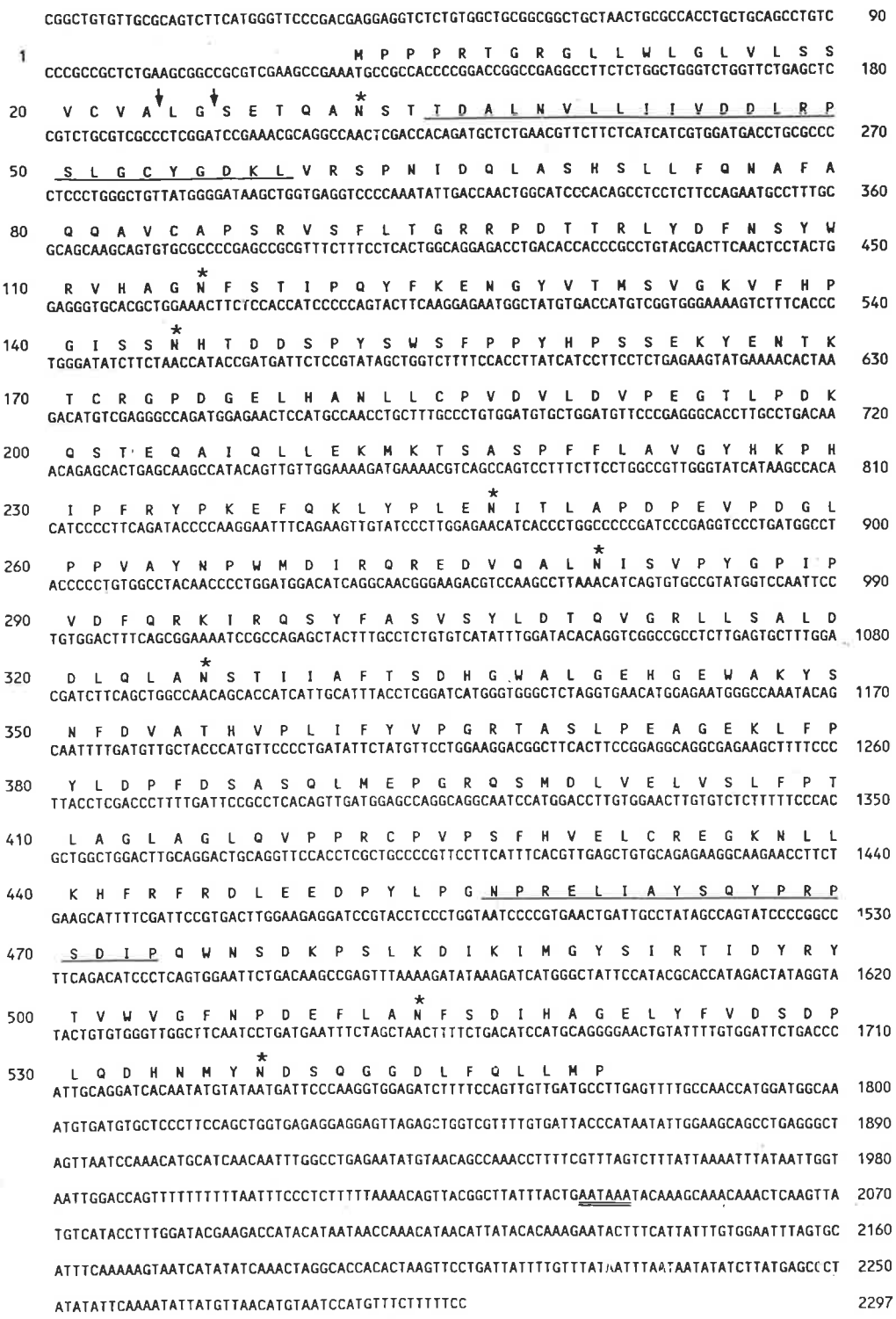
4.3.4 THE COMPLETE IDS cDNA NUCLEOTIDE SEQUENCE

Figure 4.3 shows the complete nucleotide sequence of the 2297-bp cDNA insert derived from both the 1.5-kb cDNA (λc2S15) and 2.3-kb cDNA (λc2S23) clones. Except for a few differences, the deduced amino acid sequence was colinear with the determined N-terminal amino acid sequence of both the 42-kDa and 14-kDa polypeptides. The discrepancies seen (residues 35, 53, 55 and 57) between the direct and predicted amino acid sequence most likely reflect errors due to the low signal obtained at the end of the amino acid sequencing run. The 2.3-kb cDNA sequence contains an open reading frame from the initiation codon (ATG) starting at position 125, to the termination codon (TGA) starting at position 1775. The 1650-bp coding region encodes a polypeptide of 550 amino acids as shown in Figure 4.3.

The sequence flanking the ATG codon at position 125 correlates well with the consensus sequence for initiator codons. A survey has reported that at position -3 from the initiating ATG, a purine (usually an A residue) is used 97% of the time and this conservation is seen across the animal kingdom (Kozak, 1987). The IDS cDNA has an A residue at this position. IDS also contains a G residue at position -6 and -9 and this is in agreement with the survey (Kozak, 1987). It is thought that the periodicity of A

Figure 4.3 **COMPILED NUCLEOTIDE SEQUENCE OF THE IDS cDNA CLONES AND THE DEDUCED AMINO ACID SEQUENCE OF THE ENCODED PROTEIN**

Amino acid sequence is shown in the one-letter code above the nucleotide sequence. Nucleotide numbers are depicted on the right-hand margin and amino acid numbers are depicted on the left-hand margin. Possible sites for peptidase cleavage of the signal peptide are indicated with arrows. Underlined amino acids are co-linear with amino-terminal sequences. Potential N-glycosylation sites are starred. A potential polyadenylation signal is doubly underlined.



and G residues at these positions helps the ribosome stay in frame during translation (Trifonov, 1987).

Lysosomal enzymes are trafficked through the ER and the presence of an ER signal peptide sequence is required to ensure that they are transported across the ER membrane into the lumen of the ER to undergo further processing (Section 1.4.1 and Figure 1.2). For lysosomal enzymes this processing includes the attachment of a M6P group to high mannose oligosaccharide chains. The function of the non-phosphorylated oligosaccharide chains is at present unknown but it is possible that they aid in the correct folding of the protein, or may help to protect the amino acid sequence from degradation within the acidic and proteolytic environment of the lysosome. The attachment of a M6P residue is critical for the sequestering of lysosomal enzymes to the pre-lysosome body (see Chapter 1, Section 1.4.1 and Figure 1.2).

ER signal peptide sequences have a basic N-terminal region, a central hydrophobic region and a more polar C-terminal region (von Heijne, 1985). The first 25 amino acids at the amino terminus of the deduced protein were examined using a weight matrix score based on the signal sequences derived from 161 eukaryotic signal sequences. The first 25 amino acids of the deduced amino acid sequence from the 2.3-kb cDNA were found to have features characteristic of an ER signal peptide (von Heijne, 1986). Two putative sites for cleavage of the ER signal peptide sequence from the mature protein were predicted from the weight matrix scores and these are indicated by arrows in Figure 4.3 (between amino acid residues 23-24 and 25-26). The matrix scores predicted that the best possible ER signal sequence cleavage site was between residues 25-26, but this leaves a further eight amino acids prior to the sequence found at the N-terminus of the 42-kDa polypeptide. These eight amino acids

may be removed from the IDS precursor either during transport or when the protein reaches the lysosome (Hasilik and von Figura, 1984). It is generally thought that this additional proteolytic processing in the case of the lysosome is a consequence of the proteolytic environment and is not generally required for the acquisition of enzyme activity (von Figura and Hasilik, 1986; Kornfeld and Mellman, 1989; Neufeld, 1991).

The 14-kDa polypeptide N-terminal amino acid sequence commenced at amino acid residue Asn_{456}, giving a total of 95 amino acids to the carboxyl terminus. When the full amino acid sequence of IDS was examined eight potential N-glycosylation sites were observed (Asn-Xaa-Ser/Thr; marked in Figure 4.3 with an asterisk). The molecular weight of the deduced polypeptide for the 14-kDa component was calculated as 11,093 Da. This suggests that the two potential N-glycosylation sites may in fact be glycosylated with a complex-type oligosaccharide of approximately 2-kDa to give a total calculated molecular weight of approximately 15,000 Da. However, this will require further detailed examination to determine if this is the case. One possible way of determining the number of glycosylation sites utilised is by the use of endoglucosaminidase H in conjunction with a range of truncated cDNA's that have different numbers of glycosylation sites that are then translated *in vitro* in the presence of microsomes (Stein *et al.*, 1989b). SDS-PAGE can then be used to determine the change in molecular weight of the various truncated proteins and the N-glycosylation sites used can be calculated. This approach was used successfully to determine the number of N-glycosylation sites utilised in steroid sulphatase (Stein *et al.*, 1989b).

The number of N-glycosylation sites used in the 42-kDa polypeptide is unknown. The first N-glycosylation site (residue 31) is not contained within Form A of the protein preparation since this asparagine residue is thought to be removed during amino-terminal processing. The molecular weight of the deduced polypeptide for the 42-kDa

component was calculated as 47,404 Da (mature 42-kDa starts at Thr_{34}). This suggests that the value obtained by SDS-PAGE may be in error or that only some of the N-glycosylation sites are used, or that further C-terminal proteolytic processing occurs to produce the observed 42-kDa polypeptide. The 14-kDa polypeptide does not contain any cysteine residues required for disulphide bonds. This means that the 42-kDa and 14-kDa polypeptides are likely to be associated by either hydrophobic or hydrophilic interaction. This proposal is supported by the result obtained from SDS-PAGE run under reducing and non-reducing conditions (Bielicki *et al.*, 1990), where the 42-kDa and 14-kDa polypeptide remain unchanged.

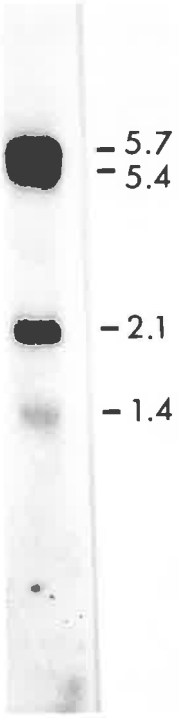
4.3.5 NORTHERN ANALYSIS OF IDS

This work was undertaken in collaboration with Teresa Occhiodoro in our group.

Total RNA was isolated from human placental tissue according to Section 5.2.1. Poly(A)$^+$ RNA was prepared using an oligo(dT)-cellulose column, and the RNA sample was loaded onto a denaturing agarose gel (see Section 4.2.4 for description of the procedure used). The 1.5-kb cDNA insert derived from λc2S15 was oligolabelled and used to probe the Northern filter to determine the size and number of IDS mRNA species present in human placenta. The results obtained are shown in Figure 4.4. Three major species of 5.7, 5.4 and 2.1-kb were observed and there was also one minor species of 1.4-kb. It is likely that IDS, like other known lysosomal sulphatases (e.g. arylsulphatases A, B and C (Stein *et al.*, 1989a,b; Peters *et al.*, 1990; Yen *et al.*, 1987) has mRNA species that differ in length at their 3'-ends due to differential polyadenylation. Arylsulphatase C has two major RNA transcripts that result from the use of different polyadenylation sites (2.7 and 5.2-kb major mRNA species and minor

Figure 4.4 **NORTHERN ANALYSIS OF IDS**

Total RNA was isolated from placental tissue as described in Chapter 5 (Section 5.2.2) and poly(A)$^+$ was obtained by oligo(dT)-cellulose chromatography. RNA was electrophoresed under denaturing conditions and transferred to a nylon membrane as described in Section 4.2.4. The Northern blot was then probed with oligolabelled 1.5-kb cDNA insert as described in Materials and Methods. The size (kb) of each RNA species is shown in the right margin


5.7
5.4
2.1
1.4

species at 7.2, 4.3 and 3.4-kb) (Yen *et al.*, 1987). They also found that there was heterogeneity in mRNA length from different tissue sources. It is possible that some developmental or tissue-specific regulation may be conferred at the level of mRNA production. Differential polyadenylation can account for the three major IDS mRNA species seen in Figure 4.4, but it cannot explain the 1.4-kb minor species. An mRNA of this size would not encode enough information for the 550 amino acids of the IDS precursor protein.

There are a number of possibilities to explain this observed mRNA species. The 1.4-kb mRNA could either be a degradation product, a cross-reacting species, or it may represent an alternatively spliced mRNA that may produce a different functional protein. Differential splicing has also been observed in other lysosomal enzymes including β-glucuronidase (Oshima *et al.*, 1987) and β-galactosidase (Morreau *et al.*, 1989). However, this possibility was ruled out when PCR was performed on IDS mRNA and only the product of the expected size was observed (data not shown).

Examination of the 520-bp of 3'-untranslated sequence identified a potential polyadenylation signal (**AATAAA**) at position 2041 (see Figure 4.3; indicated by a double underline) which may direct the position of polyadenylation for the observed 2.1-kb mRNA species (Figure 4.4). If this is correct, the 124-bp of 5'-untranslated sequence in the 2.3-kb cDNA clone is sufficient to account for most, if not all, of the 5'-untranslated region expected for the 2.1-kb mRNA species (assuming a typical 50-100 residues of poly(A) tail). The exact function of the poly(A) tail has often been open to much speculation. Recently, fusion constructs between the bacterial neomycin resistance gene and either a histone 3' processing or a β-globin cleavage/polyadenylation signal show that these 3'-end formation signals contain not only the necessary information for 3'-end generation but also for efficient nucleo-

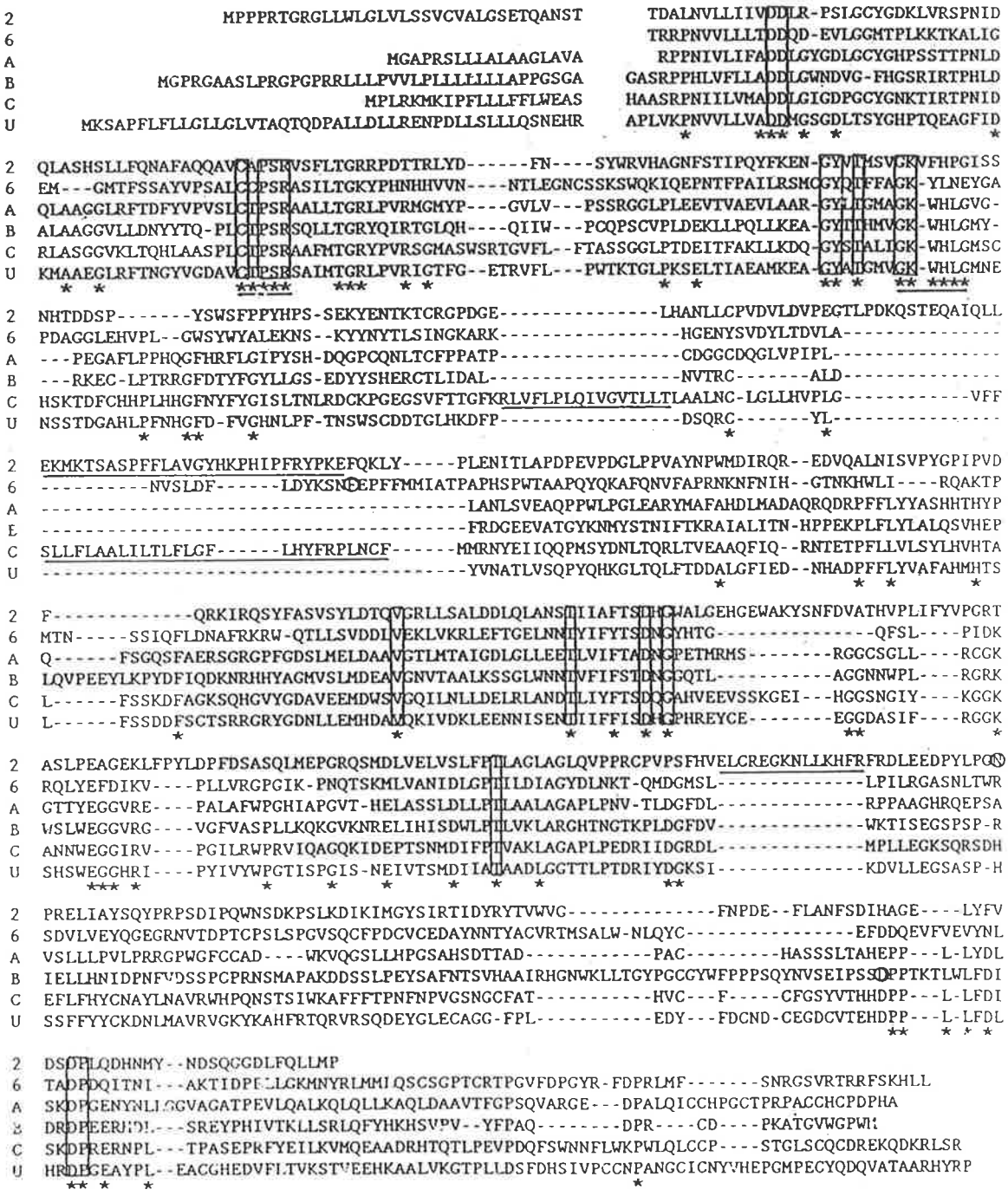
cytoplasmic mRNA transport (Eckner *et al.*, 1991). If the 3'-end is sufficient for transport from the nucleus to the cytoplasm then the function of the poly (A) tail remains unknown. It is possible that it may be required for efficient translation or greater mRNA stability.

4.3.6 COMPARISON OF THE IDS AMINO ACID SEQUENCE WITH OTHER KNOWN SULPHATASES

Robertson *et al.* (1988) were the first to show that the known sulphatases show sequence homology. It was thought this suggested a common origin for these enzymes and when the deduced amino acid sequence of IDS became available, the homology with other known sulphatase sequences was determined. Figure 4.5 reveals many areas of identical and conserved amino acid matches within the arylsulphatase group (galactose-3-sulphatase, ASA; N-acetylgalactosamine-4-sulphatase, ASB; steroid sulphatase, ASC) and the non-arylsulphatases including iduronate-2-sulphatase (IDS) and glucosamine-6-sulphatase (G6S), (comparison was also made without G6S) (Wilson *et al.*, 1990). Sea urchin (*hemicentrotus pulcherrimus*) arylsulphatase is also aligned and was found to have sequence homology with the human sulphatases. Sea urchins are believed to have evolutionarily separated from humans more than 500 million years ago. When conservative amino acid substitutions are taken into account the high degree of evolutionary conservation suggests that the genes for all of these sulphatases (aryl and non-aryl) arose from a common ancestral gene and therefore can be considered members of a sulphatase gene family (Robertson *et al.*, 1988; Peters *et al.*, 1990; Wilson *et al.*, 1990; Robertson *et al.*, 1992). The aryl- and non-arylsulphatases are defined according to their ability to degrade the fluorogenic substrate 4-methylumbelliferyl sulphate. The arylsulphatases degrade this substrate

Figure 4.5 **HOMOLOGY ALIGNMENT OF THE KNOWN SULPHATASES**

Alignment of amino acid sequences of human IDS, human glucosamine-6-sulphatase (Robertson *et al.*, 1988), human galactose-3-sulphatase (ASA) (Stein *et al.*, 1989a), human N-acetylgalactosamine-4-sulphatase (ASB) (Peters *et al.*, 1990), human steroid sulphatase (ASC) (Yen *et al.*, 1987; Stein *et al.*, 1989b) and sea urchin arylsulphatase (Sasaki *et al.*, 1988) shown in lines 2, 6, A, B, C, and U, respectively. Amino acids identical in all sulphatases are boxed. Amino acids identical in the arylsulphatases (lines A, B, C, and U) are starred on the bottom line. The ringed residues in lines 2, 6 and B indicate the first amino-terminal amino acid in polypeptides produced by internal proteolysis. Underlined sequences are unique to each particular sulphatase sequence and underlined and starred sequences are blocks of conserved residues (see Section 4.3.6 for detailed discussion).



efficiently whereas the non-arylsulphatases such as G6S and IDS, even though they are able to bind to this substrate, are unable to degrade it efficiently (Robertson *et al.*, 1992; J. Bielicki, unpublished data).

The multiple alignment of the amino acid sequences has the highest level of homology in the amino-terminal third of each sulphatase, however only a small number of these conserved regions will be discussed (Figure 4.5). The human arylsulphatase group has conserved blocks of up to six identical amino acid residues, for example Leu-Cys-Thr-Pro-Ser-Arg (LCTPSR), Gly-Lys-Trp-His-Leu-Gly (GKWHLG) and Thr-Gly-Arg (TGR) (indicated by blocks in Figure 4.5). Peters *et al.* (1990) speculated that the histidine residue in the highly conserved GKWHLG region of the arylsulphatases, the conserved arginine residues in and close to the LCTPSR region and the TGR region may contain the catalytic residues required for sulphate ester hydrolysis. This hypothesis was supported from studies reporting that one histidine residue (Lee and van Etten, 1975) and two or more arginine residues (James, 1979) are essential for the catalytic activity of ASA. However, when these regions are examined in IDS and G6S, the conservation is not as rigid and casts doubt on the catalytic importance of these residues in these enzymes. For example, in IDS, this sequence is VCAPSR and GK-V-FHPG and for G6S it is LCCPSR and GKYLNE. Apart from the insertion of a valine residue in IDS the majority of the amino acid substitutions are conservative. Also, the histidine residue is present in IDS but absent in G6S. IDS displays better homology than G6S in the LCTPSR and TGR region with the exception of a conservative substitution of valine for leucine at residue 83 and alanine for threonine at residue 85. Recently, the cDNA encoding N-acetylgalactosamine-6-sulphate sulphatase (NG6S) was isolated and the translated amino acid sequence was used to determine the homology of this enzyme with other published sulphatases. In the case of NG6S the LCTPSR region of the human arylsulphatases was LCSPSR, and the

GKWHLG and TGR regions were unaltered (Tomatsu *et al.*, 1991). This strengthens the proposal that all of these sulphatases arose from a common ancestral gene. From the alignment it can be speculated that IDS and NG6S are perhaps closer in evolution to the arylsulphatases than is G6S.

The alignment of the six sulphatases in Figure 4.5 suggests that IDS has several peptide inserts (Wilson *et al.*, 1990). The microsomal membrane-bound steroid sulphatase (ASC) was shown to contain two membrane-spanning regions. IDS was also shown to contain an amino acid insert in the same region as the second membrane-spanning region of ASC, but it is probably not membrane-spanning because it is highly hydrophylic. A second peptide insert in IDS is present just before the amino-terminal sequence of the 14-kDa polypeptide. Wilson *et al.* (1990) speculated that the position of this insert may direct internal proteolysis. However, such alignments can only serve as a useful guide and it is unwise to attach importance to particular regions based solely on this information. The predicted inserts in IDS may therefore be artefactual. When the alignment was done in the absence of G6S no predicted peptide insert regions in IDS were observed (data not shown). However, both of these alignments agree that the major conservation occurs in the amino-terminal third of the sulphatases. With the reduction of these areas of absolute homology it should enable the detailed study of the remaining conserved areas across all the sulphatases (aryl and non-aryl).

In order to gain insights as to where the regions of functional importance in IDS lie, a multi-disciplinary approach will be required and would include determination of the tertiary structure of the IDS protein to enable the prediction of regions likely to be critical for IDS enzyme function, and the study of patient mutations which should also help our predictions. This would involve examination of both the severe and mild

cases which may make it possible to locate regions of critical importance for enzyme function. In addition, site-directed mutagenesis of certain conserved residues proposed to be important for specific sulphatase functions (ie. GKWHLG, LCTPSR, and TGRXXX) may help to identify the catalytic site of IDS and regions critical for the binding of IDS substrate. Also by comparing regions that differ between each sulphatase it may be possible to identify sequences in functions that are unique to each sulphatase or sulphatase group. For example, regions important for defining substrate specificity may be identified by comparing a non-arylsulphatase such as IDS to an arylsulphatase such as ASB. In this case the regions responsible for arylsulphatase specificity toward 4-methylumbelliferyl should be strongly conserved among the arylsulphatases but not in the non-arylsulphatases. Once the appropriate conserved residues are selected for mutagenesis, the nature of the mutated residue must be considered. Generally, a conservative substitution should be made to ensure that the structural integrity of IDS is maintained although there is scope for substituting the amino acid with a small neutral amino acid like alanine. The principal step in studying transiently expressed mutant IDS is the measurement of crude enzyme activity and quantification of levels of mutant protein using monoclonal antibodies in order to determine the specific activity of the mutant enzyme (Brooks *et al.*, 1991). A more detailed analysis can then be made which would include the measurement of the mutant enzymes K_m, V_{max} and substrate structure/function studies, as well as the analysis of enzyme maturation in order to see if the mutant enzyme follows the same processing pathway as the normal enzyme.

4.3.7 EXPRESSION OF THE cDNA CLONE IN CHO CELLS CONFIRMS THAT THE cDNA ENCODES FUNCTIONAL IDS PROTEIN

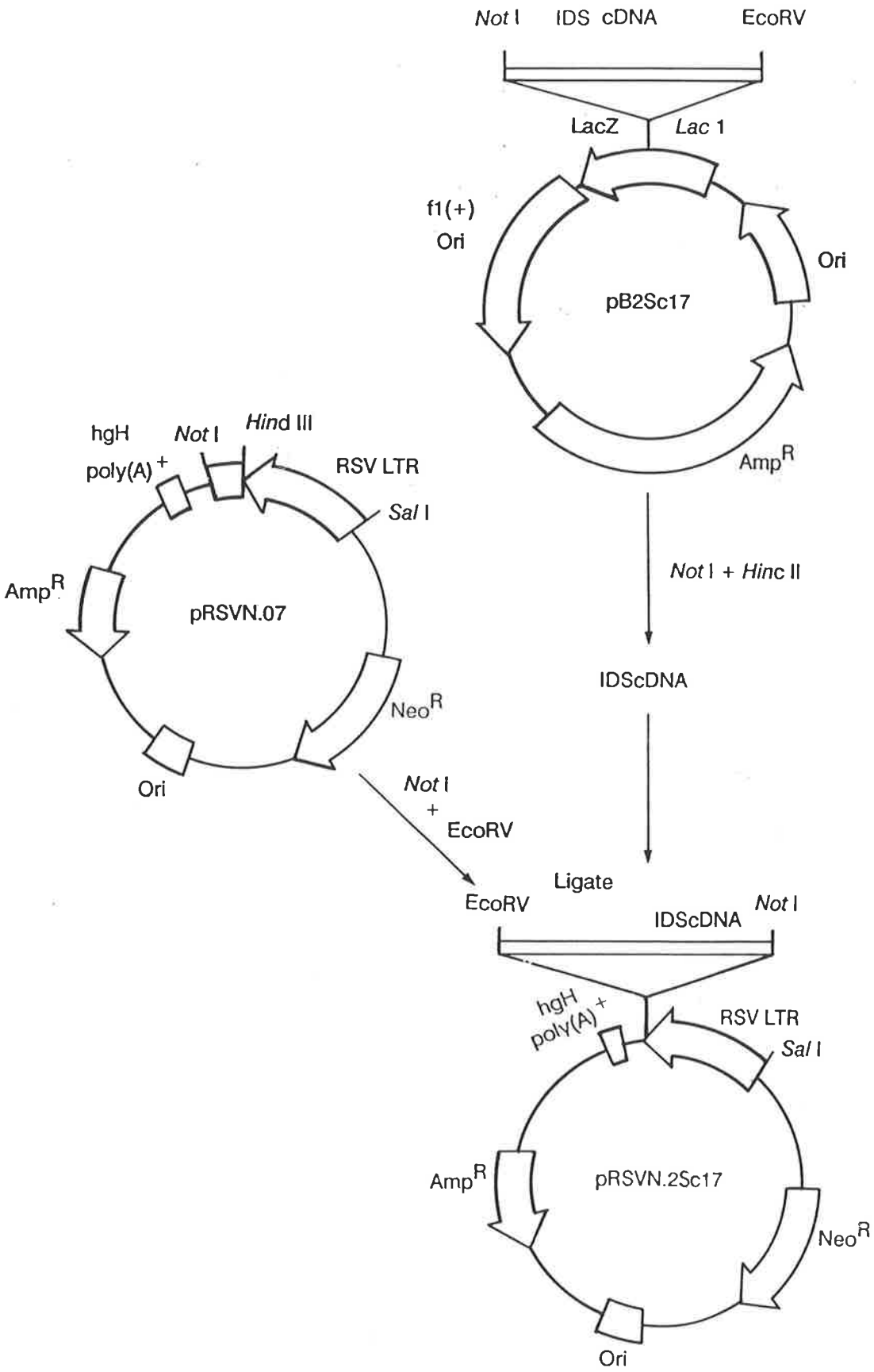
4.3.7a CONSTRUCTION OF AN IDS EXPRESSION VECTOR

A cDNA fragment encompassing the entire coding region of IDS was generated from λc2S23 by complete *Not*I digestion and partial digestion with *Bst*XI (position 107-1866; Figure 1 in Wilson *et al.* (1990); Figure 4.3 in this thesis). Firstly, the cDNA was partially digested with *Bst*XI and the *Bst*XI overhang was made blunt-ended by digestion with mung bean nuclease. The cDNA was then completely digested with *Not*I. The resulting *Not*I/*Bst*XI fragment was cloned into the *Not*I/*Eco*RV cut and phosphatased plasmid vector pBLUE SK II, creating pB2Sc17. This cDNA insert was then excised by *Not*I/*Hin*cII digestion, gel purified and cloned into the *Not*I/*Eco*RV digested and phosphatased pRSVN.07, creating the expression construct pRSVN.2Sc17 (Figure 4.6). This places the IDS cDNA under the transcriptional control of the Raus Sarcoma Virus Long Terminal Repeat (RSV LTR). The clones were checked for the correct orientation by restriction enzyme analysis. A large amount of one clone, designated pRSVN.2Sc17, was then grown and purified by caesium chloride gradient ultracentrifugation.

Two further expression constructs were subsequently generated (Bielicki *et al.*, 1992). In the first of these, the normal 5'-non-coding sequence of 2Sc17 (from pB2Sc17) was removed and replaced with 45-bp of the rat preproinsulin 5'-non-coding sequence. This construct was designated pRSVN/2SNC1. The last expression construct was made by replacing the RSV LTR in pRSVN.2SNC1 with the human polypeptide elongation factor 1-α (EF-1α) promoter from pEFBOS (Mizushima and Nagata, 1990).

Figure 4.6 **SCHEMATIC DIAGRAM DEPICTING THE ENGINEERING OF THE EXPRESSION VECTOR, pRSVN.2S17**

See Section 4.3.7a for a detailed description of the process involved in the construction of this expression vector, and Table 4.1 for the level of IDS expressed in CHO-DK1 cells by this vector.


Not I
IDS cDNA
EcoRV
LacZ
Lac 1
f1(+)
Ori
Ori
pB2Sc17
Amp^R
Not I + Hinc II
IDScDNA
hgH
poly(A)+
Not I
Hind III
RSV LTR
Sal I
Amp^R
pRSVN.07
Neo^R
Ori
Not I
+
EcoRV
Ligate
EcoRV
IDScDNA
Not I
hgH
poly(A)+
RSV LTR
Sal I
Amp^R
pRSVN.2Sc17
Neo^R
Ori

4.2.7b IDS EXPRESSION

Julie Bielicki was responsible for the electroporation and maintenance of the second and third expression vector constructs described in this section. Julie also performed all of the IDS enzyme assays. Dr. Don Anson constructed the second and third expression vectors.

pRSVN.2Sc17 was introduced into CHO-K1 cells as described in Section 4.2.6a, and $G418^R$ clones were isolated and expanded. Conditioned medium from 24 individual $G418^R$ clones containing the pRSVN.2Sc17 expression constructs were assayed for IDS enzyme activity as described in Section 4.2.6b. Table 4.1 shows that the best of these clones (clone 11i) produced 20 ng/ml of IDS in the medium, whereas untransfected CHO-K1 cells produced 0.42 ng/ml of IDS in medium. The level of IDS expression was low when compared to that obtained for N-acetylgalactosamine-4-sulphatase (G4S) using the same expression vector, where up to 6 μg/ml was produced (Anson *et al.*, 1992). One possible reason for this difference could be the truncated 5'-untranslated sequence used in the IDS expression construct which includes only 18 bp out of the 124 bp sequenced in the IDS cDNA prior to the initiation methionine. It is possible that this sequence does not contain all the translational signals contained in the IDS mRNA.

Further expression constructs were generated by Dr. Don Anson and these are included here for comparative purposes. In the first of these constructs, pRSVN.2SNC1, the normal IDS 5'-non-coding sequence was replaced with 45 bp of the 5'-non-coding sequence from the rat preproinsulin II gene. This sequence has been reported to greatly enhance the efficiency of interlukin-2 (IL-2) mRNA translation (Cullen, 1988). Several factors have been reported to affect mRNA translational efficiency (Kozak, 1984; 1986a,b; 1991). This includes the sequence

Table 4.1 **TABLE SHOWING THE RELATIVE LEVELS OF IDS EXPRESSED IN CHO-DK1 CELLS BY THREE DIFFERENT EXPRESSION VECTORS**

See Section 4.2.6a for description of how vectors were introduced into the CHO-DK1 cells and also how the IDS activity was determined.

Vector	IDS produced in culture medium (μg/ml)	IDS produced in cell lysates (μg/ml)
1. pRSVN.2Sc17	0.01	0.62
2. pRSVN.2SNC1	4.7	3.15
3. pEFN.2SNC1	8.5	nd
4. No Vector	0.0002	0.00025

nd, not done

1. All amounts of IDS are calculated from specific activity of IDS (20.8 x 10^6 pmoles $min^{-1}.mg^{-1}$) (Bielicki *et al.*, 1992).

2. Cells were grown in 10 mm dishes until confluent. The medium was changed and the cells were incubated for a further 48 hours. Medium was collected for IDS enzyme assays (described in Section 4.2.6).

surrounding the initiation codon, however this sequence in the IDS cDNA was found to be very close to the ideal rules proposed by Kozak (1987). An alternative possibility could be the presence of secondary structures in the mRNA, however the short sequence from the IDS cDNA does not appear to contain any significant regions susceptible to secondary folding. IDS is a house-keeping gene and the 5'-non-coding sequence may have evolved to poduce low levels of protein. The IDS 5'-non-coding sequence is also extremely G+C-rich, a feature characteristic of other known house-keeping genes. It has been proposed that IDS is under the control of an Sp1 promoter element (described in Section 3.3.5) which is also typical of house-keeping genes (Kadonga *et al.,* 1986; Bucher, 1990). Low level expression is also a feature typical of other known lysosomal enzymes. It would appear that the presence of the preproinsulin 5'-non-coding sequence in the second IDS expression construct greatly enhances expression, presumably by improving translational efficiency or IDS mRNA stability (Table 4.1). Expressed IDS in the second vector construct represents a 900-fold increase in the levels of IDS produced when compared with the levels expressed by the first vector construct (Table 4.1). While the exact reason for this is unknown, the result demonstrates that a 5'-non-coding sequence derived from an efficiently translated mRNA may also enhance the expression of other coding regions when present in *cis* (Jobling and Gehrke, 1987; Cullen, 1988).

The third expression vector construct, pEFN.2SNC1, contained the powerful promotor EF-1α (Mizushima and Nagata, 1990). Table 4.1 indicates that this vector construct gave almost a 2-fold increase over the second vector construct which containred RSV LTR as promoter (Bielicki *et al.,* 1992).

The size of IDS, purified from the medium produced by the third expression vector construct on an SDS-PAGE gel was calculated as 90-kDa, which is considerably larger

than the value obtained for the IDS protein isolated from lysosomes, namely a 42-kDa and 14-kDa species (Bielicki *et al.*, 1992). Incubation of the CHO expression cells in NH_4Cl induces expressed M6P IDS to be secreted into the culture medium. This occurs because the M6P IDS and the M6P receptor require an acidic pH in order to dissociate. This dissociation occurs in an acidic endosome compartment (see Section 1.3.1 and Figure 1.2- compartment C). NH_4Cl increases the pH of the endosome and prevents this dissociation. This causes saturation of the available M6P receptors and, as a consequence, there are no M6P receptors to transport lysosomal proteins. Instead, these proteins follow the path of secretory proteins and are secreted into the culture medium.

The calculated molecular weight of the IDS mature protein is 58.5-kDa without N-glycosylation. It is therefore likely that the expressed IDS with molecular size of 90-kDa is heavily N-glycosylated with complex oligosaccharide chains accounting for the increased molecular weight (there are 8 potential N-glycosylation sites; see Section 4.3.4). This proposal was supported when the expressed protein was digested with endoglycosidase F, reducing the protein from a diffuse 90-kDa to a sharp 60-kDa which is consistent with the value calculated for unprocessed IDS (Bielicki *et al.*, 1992).

The expressed products generated from all three vectors efficiently and specifically degraded IDS disaccharide substrates, demonstrating that the cDNA used was full-length and encoded IDS. The availability of sufficient quantities of IDS should enable the production of monoclonal antibodies which have been singularly difficult to produce using both crude protein preparations and artificial peptides designed from the IDS cDNA sequence. It is thought that this difficulty is partly due to the relatively poor antigenicity of IDS as judged by the computer program that identifies immunogenic regions within a polypeptide sequence (Krchnak *et al.*, 1987), making the

design of synthetic peptides difficult (Dr. Doug Brooks, personal communication). The availability of antibodies would enable an accurate comparison of IDS levels between MPS II patients and unaffected controls. The high level expression of IDS also makes enzyme-replacement therapy a possibility (Bielicki *et al.*, 1992).

The 1.5-kb cDNA was also used to probe MPS II patient DNA and revealed a wide variety of gene alterations. This provides further evidence that the cDNA isolated in this Chapter encodes the IDS. In addition, using a full-length IDS cDNA clone as a probe, chemical cleavage was used to study small MPS II mutations. These results are discussed fully in Chapter 5.

CHAPTER FIVE

MOLECULAR ANALYSIS OF MPS II PATIENTS

5.1 INTRODUCTION

The results presented in Chapter 4 proved conclusively that the identity of the cDNA clone isolated with the 1.3-kb *Hin*dIII genomic fragment was full-length and encoded IDS. The availability of a full-length IDS cDNA clone therefore enabled the analysis of the mutations responsible for MPS II. Further confirmation of the identity of the cDNA clone was obtained when it was demonstrated that in some MPS II patient DNA samples, a variety of gene rearrangements were observed by Southern analysis.

The diverse variation in clinical phenotype seen in MPS II patients prompted studies of the molecular nature of the gene defect in a range of clinically different MPS II patients. This chapter describes the identification of patients who revealed gene alterations, including some patients who had a deletion of the entire IDS gene. A collaborative survey of 220 MPS II patients has revealed that approximately 18% of the patients probed showed abnormal Southern patterns, and 4% of these patients probed had complete IDS gene deletions (see Table 5.1). The extent of the IDS gene deletion in two MPS II patients extended well beyond the IDS locus. DNA from an MPS II female was also studied and it was shown that the translocation of her X-chromosome disrupted the IDS gene. The remaining 82% of MPS II patients are thought to have mutations that are beyond the limits of Southern analysis detection, including point mutations, small deletions or small insertions.

The last section of this Chapter deals with the detection of mutations in MPS II patients with normal Southern blot patterns. It is possible that mutations at the nucleotide level may reveal amino acid regions critical for IDS function. In addition it was thought that it may be possible to generate a correlation between the patient genotype and the patient clinical phenotype. The isolation and sequencing of IDS

provides an excellent opportunity to search for specific mutations that result in the various clinical phenotypes seen in MPS II. Early methods for mutation detection involved the laborious and tedious process of re-cloning and sequencing each individual patient gene (DiLella *et al.*, 1986, 1987). Once the mutation was characterised in these early studies, it was possible to screen other patients for the same mutation, however many patients remained in which these mutations were not detected.

Methods were developed that were less labour-intensive and detected the majority of mutations. The techniques and methods available for mutation research has expanded at a rapid rate in recent years and this progress is reviewed by Cotton, 1989. Briefly, the early methods involved either denaturing gradient gel electrophoresis (Myers *et al.*, 1985b) or RNAase cleavage methods (Myers *et al.*, 1985a). With the development of PCR technology (Saiki *et al.*, 1985) regions of interest were studied with greater efficiency and speed (Sheffield *et al.*, 1989).

Since that time, one reliable method of screening has been developed which detects mismatches in heteroduplexes formed between normal and patient DNA or RNA that has been used to search for mutations in many genetic disorders (Cotton *et al.*, 1988; Dahl *et al.*, 1989; Grompe *et al.*, 1989; Montandon *et al.*, 1989; Dahl *et al.*, 1990; Howells *et al.*, 1990; Forrest *et al.*, 1991; Dianzani *et al.*, 1991; Litjens *et al.*, 1992; Scott *et al.*, 1992; Wilson *et al.*, 1992). At the position of the mutation, a mismatch occurs which is more reactive with hydroxylamine or osmium tetroxide than the surrounding bases. These susceptible bases are cleaved with piperidine and the position of the mutation can be determined by fractionation of the sample by denaturing gel electrophoresis. This technique has the potential to detect all mutations, however it has been reported that certain classes of T-G mismatches appear to be more resistant

to cleavage (R.G.H. Cotton, unpublished data). Some of these T-G mismatches can be detected by using a normal probe of the opposite sense, with indirect reactivity of nearby matched bases indicating the presence of a mismatch (Cotton and Campbell, 1989; see also Section 5.4.1). For the remaining T-G mismatches a different approach is necessary where both the patient and normal DNA is labelled thus enabling 100% mutation detection (Pilz *et al.*, 1990; Forrest *et al.*, 1991). .

PCR was used to amplify the coding region of the IDS gene from reverse transcribed messenger RNA isolated from MPS II patient cultured skin fibroblasts. Using chemical cleavage of heteroduplexes between patient and normal IDS and direct PCR sequencing, a number of mutations and polymorphisms were detected which were all unique in the small pool examined. Their association with MPS II will be the subject of discussion in Section 5.4.1.

5.2 SPECIFIC METHODS

5.2.1 PREPARATION OF RNA FROM FIBROBLASTS

The MPS II patient skin fibroblasts described in this thesis were revived and maintained by Cathy Nelson and Greta Richardson in the Department of Chemical Pathology.

The extraction of total RNA from fibroblasts was carried out essentially as described by Chomczynski and Sacchi (1987). Primarily, monolayers of confluent cultured fibroblast cells were scraped from the culture flasks using a rubber policeman into phosphate buffered saline (PBS). The cells were pelleted by centrifugation at 1,500 rpm for 5 min (Heraeus Sepatech benchtop centrifuge). The pellet was resuspended in an appropriate volume of the guanidium isothiocyanate solution (solution D; 4 M guanidinium isothiocyanate, 25 mM tri-sodium citrate, 0.5% (w/v) sarkosyl, 0.1 M β-mercaptoethanol: this solution was made fresh) (Chomczynski and Sacchi, 1987), judged by the weight of the fibroblast pellet, for example 20 mg of cultured fibroblast cells were resuspended in 200 μl of solution D.

The pellet was resuspended by vortexing, with 1/10 vol of 2 M sodium acetate, pH 4.0. The mixture was vortexed before adding an equal volume of phenol saturated with water, then 1/5 vol of isoamyl alcohol/chloroform was added and the mixture was left on ice for 15 min. The suspension was centrifuged at 12,000 rpm for 10 min at 4°C to separate the solvent and water soluble phases, and the supernatant was transferred to a fresh tube to which an equal vol of isopropanol was added then this solution was placed at -20°C overnight to aid in the precipitation of RNA. The mixture was centrifuged at 12,000 rpm for 15 min at 4°C, the supernantant was drained, and the RNA pellet was resuspended in 20 μl of Solution D, 1/10 vol of 3 M sodium acetate,

pH 5.2, 2.5 vol of ethanol was added and the mixture was incubated at -20°C for 120 min to aid in RNA precipitation. The mixture was centrifuged as described above and the RNA pellet was washed once with 70% aqueous ethanol, vacuum-dried and dissolved in 20 μl of water. A 2 μl sample was then used to determine the concentration of RNA by absorbance spectroscopy at a wavelength of 260 nm.

5.2.2 cDNA SYNTHESIS FROM PATIENT RNA

cDNA was prepared from total cellular RNA by reverse transcription. Aliquots of RNA dissolved in water (6 μg) that was extracted from patient cultured skin fibroblasts (see above) were added to reaction mixtures containing 1 x Moloney murine leukemia (M-MLV) reverse transcriptase buffer (50 mM Tris-HCl, pH 8.3, 40 mM KCl, 6 mM $MgCl_2$, 1 mM DTT), 80 units of RNAsin (Promega), 0.5 mM dNTP's (Boehringer Mannheim), 5 μg of random octamers, and 1,200 units of M-MLV reverse transcriptase (BRL) to a final volume of 100 μl. This mixture was incubated at 37°C for 60 min, followed by the addition of 10 μl 3 M NaOH and incubation at 37°C for 45 min to achieve the alkaline hydrolysis of RNA. The reaction mixture was neutralised by the addition of 2.5 μl 32% (v/v) HCl, and the cDNA was precipitated with 1/10 vol 3 M sodium acetate, pH 5.2, and 2.5 vol of ethanol. After incubation overnight at -20°C the solution was centrifuged at 12,000 rpm for 15 min at 4°C to pellet the precipitated cDNA. The cDNA pellet was washed once with 70% aqueous ethanol, and dried and resuspended in 100 μl of water. Aliquots of 5 μl were used in typical PCR reactions.

5.2.3 POLYMERASE CHAIN REACTION

PCR conditions were essentially the same as described by Saiki *et al.* (1988) with the exception of the final deoxynucleotide concentration and the need for dimethyl sulphoxide (DMSO) in some reactions. There were two main PCR reactions performed, covering the entire coding region of IDS. The 5'-PCR reaction was performed using both a normal cDNA clone λc2S23 (described in Section 4.3.3) and patient cDNA transcribed from mRNA (see Section 5.2.2). Essentially, the reaction was performed in 10 mM Tris-HCl, pH 8.4, 50 mM KCl, 2.5 mM $MgCl_2$, 0.1% (v/v) Triton X-100, 0.02% (w/w) gelatin, 0.4 mM dNTP's, 2 units Taq polymerase (Cetus and/or Biotech International, WA) and 10% (w/v) DMSO. The addition of DMSO was found to be necessary to reduce the problems with secondary structure thought to be the reason for non-specific products generated in the 5'-PCR product (the IDS sequence is G+C-rich at the 5'-end). Thirty cycles of 94°C for 45 sec, 58°C for 45 sec, and 72°C for 2 min were performed. The reactions for the 3'-PCR product were essentially the same except for the omission of DMSO and the annealing temperature was lowered to 56°C instead of 58°C. The 5'-PCR primer set was: **primer 16**, 5'-TAACTGCGCCACCTGCTGCA-3' (position 64-83) and **primer 17**, 5'-GGAAGAAAGGACTGGCTGAC-3' (position 766-785). The 3'-PCR primer set was: **primer 6**, 5'-GAACTCCATGCCAACCTG-3' (position 653-672) and **primer 15**, 5'-GAGCACATCACATTTGCC-3' (position 1796-1813). The 5'-PCR primer set generates a 722-base pair (bp) fragment and the 3'-PCR primer set generates a 1161-bp fragment respectively. There is a 133-bp overlap between the two DNA fragments (Figure 4.3; also see Figure 2.1 for the nucleotide positions of the primers and Figure 5.4 for relationship of PCR products A and B).

5.2.4 CHEMICAL CLEAVAGE OF MISMATCH BASES

This technique provides a powerful tool to analyse mutations at the nucleotide level. Essentially, the methods described by Cotton *et al.* (1988) and Howells *et al.* (1990) were followed. The reaction involves forming a heteroduplex between a normal radiolabelled DNA and a mutant DNA template and then treating the DNA template with base modifying chemicals. The two chemicals used in this study were osmium tetroxide and hydroxylamine hydrochloride. Osmium tetroxide (OsO_4) reacts specifically with thymidine residues to form 5,6-dihydroxy thymidine, a change that in itself does not result in chain cleavage, but rather leads to labilization of the polynucleotide chain at the position of modified T residues (Friedmann and Brown, 1978). Subsequent treatment with piperidine removes the thymidine base which leads to the cleavage of the phospho-diester bond. In the case of hydroxylamine hydrochloride ($NH_2OH.HCl$) at a slightly acidic pH reacts specifically with C residues. It has been proposed that NH_2OH adds across the 5-6 double bond in C residues, which then labilizes the ring to internal rearrangement and subsequent cleavage with piperidine (Rubin and Schmid, 1980). When a mutation is present, the strands do not exactly match and they fail to anneal correctly. This makes the mismatched bases more accessible to the modifying reagents and these bases are cleaved first with piperidine. Analysis of the chemical reactions on a denaturing polyacrylamide gel then reveals the approximate position of the most susceptible site in the DNA heteroduplex.

The normal templates were generated from the cDNA clone λc2S23 (Section 4.3.3) as described in the previous section. The only difference being that the primers were radioactively end-labelled, as described in Section 2.2.3a the DNA probes (722-bp and 1161-bp PCR products) were run on a 1.0% agarose gel, and the DNA probes were purified by the DEAE-cellulose paper method described in Section 2.2.14c. The

patient DNA fragments were generated using cDNA synthesised as described in Section 5.2.2.

5.2.4a HETERODUPLEX FORMATION

Labelled normal PCR product (10-20 ng) was added to the patient PCR product (100-200 ng) in a reaction tube containing 0.3 M NaCl, 3.5 mM $MgCl_2$, 3 mM Tris-HCl, pH 7.7, and made up to 100 μl with water. The reactions were performed in 0.5 ml PCR tubes, placed in a Perkin-Elmer thermal cycler and subjected to 100°C for 5 min followed by 100°C to 65°C and rapid cooling from a temperature ramp which was 65°C to 42°C over a period of 60 min. The DNA was precipitated with 10 μl 3 M sodium acetate, pH 5.2, and 250 μl ethanol, centrifuged at 12,000 rpm for 10 min at 4°C, the supernatant was drained, and the DNA pellet was vacuum-dried and resuspended in 12 μl of water.

5.2.4b HYDROXYLAMINE MODIFICATION

The hydroxylamine solution was made fresh by dissolving 1.39 g hydroxylamine in 1.6 ml of water and the pH was adjusted with approximately 1 ml of diethylamine to pH 6. To each tube was added 1 μl of herring sperm DNA (10 mg/ml), 6 μl of the heteroduplex described above and 20 μl of the hydroxylamine solution. The solution was mixed and the reaction was incubated at 37°C for 45 min. After the incubation, 200 μl of stop solution was added (0.3 M sodium acetate, pH 5.2, 0.1 mM EDTA, pH 8 and 25 μg/ml tRNA) and 500 μl of 95% aqueous ethanol to precipitate the DNA. After centrifugation at 12,000 rpm to bring the DNA pellet down the supernatant was drained and the DNA pellet was dissolved in 50 μl water, and the DNA was re-precipitated with 5 μl 3 M sodium-acetate, pH 5.2, and 125 μl 95% aqueous ethanol.

This was centrifuged at 12,000 rpm for 10 min, the supernatant was discarded, and the DNA pellet was carefully washed with 70% aqueous ethanol and vacuum-dried ready for piperidine cleavage.

5.2.4c OSMIUM TETROXIDE MODIFICATION

To each of the tubes was added 10 mM Tris-HCl, pH 7.7, 1 mM EDTA, 1.5% (v/v) pyridine, 1 μl of herring sperm DNA (10 mg/ml), 6 μl of heteroduplex (described in Section 5.2.4a) and 15 μl of a 4% (w/v) osmium tetroxide solution in water. The tubes were mixed and incubated at 37°C for 5 min. To each of the tubes, 200 μl of stop solution (0.3 M sodium acetate, pH 5.2, 0.1 mM EDTA and 25 μg/ml tRNA) was added and then 500 μl of 95% aqueous ethanol was added to precipitate the DNA. The tubes were mixed and centrifuged at 12,000 rpm for 10 min at 4°C. The supernatant was discarded and the DNA pellet was dissolved in 50 μl of water, and then re-precipitated with 5 μl 3 M sodium acetate, pH 5.2 and 125 μl 95% aqueous ethanol. The tubes were centrifuged as described above and the DNA pellet was rinsed with 70% aqueous ethanol and vacuum-dried, ready for piperidine cleavage. All of these reactions were carried out in a fume hood due to the volatile and toxic nature of osmium tetroxide.

5.2.4d PIPERIDINE CLEAVAGE

To both the hydroxylamine- and osmium tetroxide-treated samples, was added 50 μl of freshly prepared 10% (v/v) piperidine and these reactions were incubated at 90°C for 30 min. All of these reactions were processed in a fume hood. After incubation the DNA in both sets of reactions was precipitated with 5 μl 3 M sodium acetate, pH 5.2 and 125 μl 95% aqueous ethanol. The hydroxylamine-treated tubes were centrifuged at 12,000 rpm for 10 min at 4°C to pellet the DNA. The osmium tetroxide tubes were

incubated in a dry-ice/ethanol bath for 15 min, and centrifuged as described for the hydroxylamine-treated samples. For both sets of samples the supernatant was discarded and the DNA pellets were carefully rinsed with 70% aqueous ethanol and vacuum-dried, 5 μl of sequencing dye was added to the dried DNA pellets, the tubes were vortexed and heat-denatured at 100°C for 3 min. One to 2 μl of this sample was then loaded onto a 6% sequencing gel and processed as described in Section 2.2.13b.

5.2.5 DIRECT PCR SEQUENCING

Essentially, this procedure is based on Murray (1989), except, the buffer used was standard PCR buffer (see Section 5.2.3). The deoxynucleotide to dideoxynucleotide ratios were also modified to suite IDS DNA templates. Basically, the double-stranded PCR template was added at a concentration of 200 ng along with, 7.5 μM dNTP's, 125 μM ddATP, 25 μM ddCTP, 25 μM ddGTP, 188 μM ddTTP, 2.5 mM $MgCl_2$, 2.2 units Taq polymerase (Cetus and/or Biotech International). Ten cycles of 94°C for 30 sec, 58°C for 30 sec, and 72°C for 1 min were performed. Five μl of the PCR sample was added to 4 μl of formamide loading dye, heat-denatured and loaded onto a 6% sequencing gel and treated thereafter as a standard sequencing gel (Section 2.2.13b).

5.3 RESULTS AND DISCUSSION

5.3.1 SOUTHERN ANALYSIS OF MPS II PATIENTS.

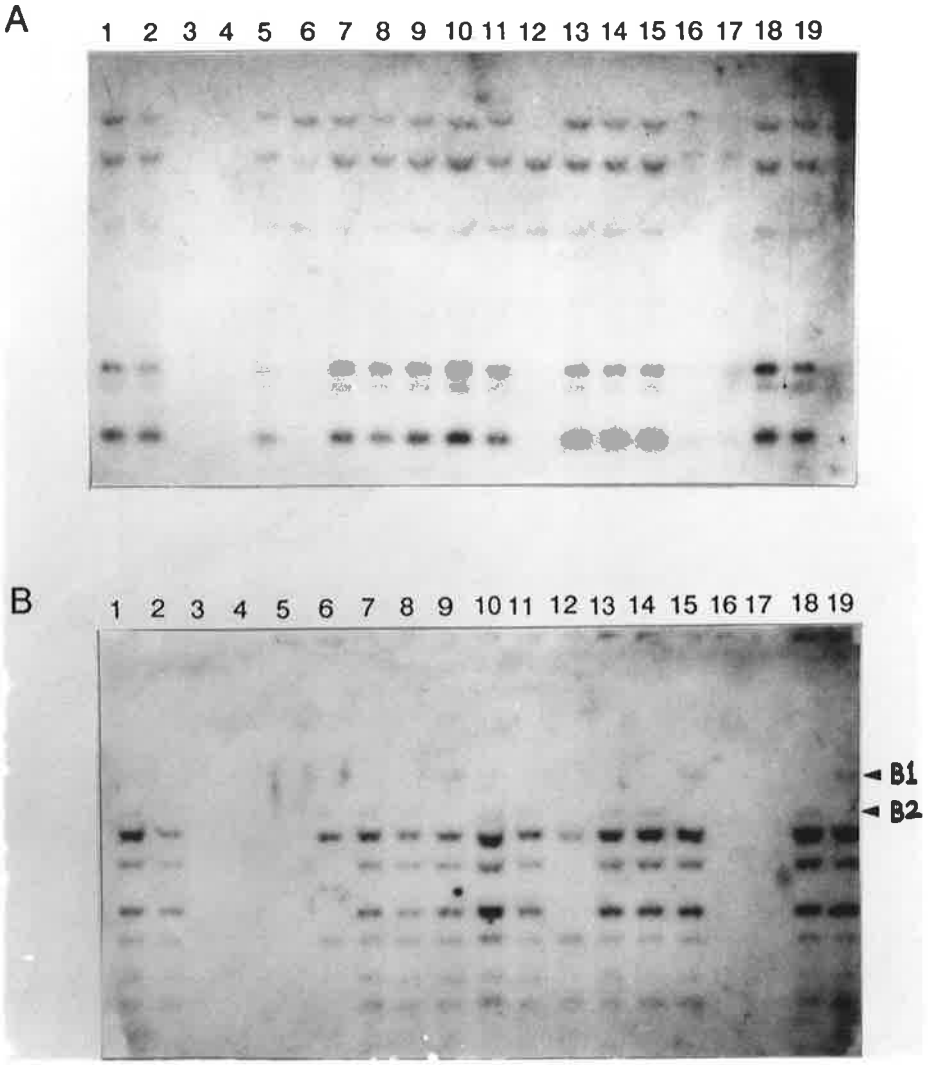
The Southern analyses performed in this section were done in collaboration with Paul Nelson in the Department of Chemical Pathology.

The 1.5-kb IDS cDNA fragment derived from λc2S15 (Section 4.3.1) was labelled and used to probe Southern blots containing MPS II patient genomic DNA. The majority of the MPS II patients studied were males unless otherwise stated and in most cases these patients had severe clinical phenotypes. In Figure 5.1A-D, six of the MPS II patient DNA samples showed either deleted bands and novel bands (lanes 3, 4, 6 and 12) or the appearance of novel bands (lanes 7, C and lane 15, D) when compared with the DNA from unaffected controls (lanes 18 and 19). There were no hybridising bands seen in the MPS II patient DNA samples in lanes 3 and 4 (Figure 5.1A-D), indicating that the entire IDS gene was deleted. It was observed that both of these MPS II patients had extremely severe clinical phenotypes (see Section 5.3.2 for detailed discussion of these two MPS II patients). The DNA samples from 2 MPS II patients (lanes 6 and 12; Figure 5.1A-D) revealed the loss of multiple bands when compared to the unaffected controls. The MPS II DNA sample (lane 6, Figure 5.1B) also revealed the appearance of a novel band at 2.5-kb. Two other MPS II patient DNA samples revealed changes to their DNA patterns, these were the MPS II DNA samples probed in lane 7 (Figure 5.1C) and lane 15 (Figure 5.1D). The MPS II patient DNA sample from lane 7 (Figure 5.1C) revealed the appearance of 2 novel bands at 11.4-kb and 2.3-kb as well as the disappearance of 2 bands at 10-kb and 7.4-kb compared to the unaffected controls. One possible explanation for this result could be that either the 10-kb or 7.4-kb DNA fragment loses a *Hin*dIII restriction site from one end and a larger *Hin*dIII restriction fragment is generated the size of this fragment is dependent

Figure 5.1 **SOUTHERN ANALYSIS OF MPS II PATIENT DNA**

DNA was digested with *Pst*I (A), *Taq*I (B), *Hin*dIII (C) and *Stu*I (D), electrophoresed and transferred to nylon membrane and probed with IDS cDNA as described in Section 2.2.11, 2.2.12 and 2.2.15 (Chapter 2).

Lane 1-15: MPS II males
Lane 16: SF2849 MPS II female.
Lane 17: SF2256 MPS II female.
Lane 18: unaffected female control.
Lane 19: unaffected male control.


A
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
B
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
B1
B2

on the position of the next flanking *Hin*dIII restriction site i.e. the 11.4-kb DNA fragment. To explain the appearance of the novel band seen at 2.3-kb it is possible that an additional *Hin*dIII restriction site is generated within either the 10-kb or 7.4-kb DNA fragment depending on which restriction site is involved in the first change. It is unknown whether either one of these changes is the mutation leading to MPS II in this patient (Figure 5.1C, lane 7) while at least one of the changes present must be a polymorphism. In addition a further possiblity to explain the appearance of 2 DNA changes could be that an inversion occurs between 2 *Hin*dIII restriction sites flanking a third internal *Hin*dIII restriction site, causing both a larger and a smaller DNA fragment to be generated ie. the 11.4-kb and 2.3-kb DNA fragments. However, the chance of this type of genetic change occurring without being detected by the other restriction enzymes used in this study are small (Figure 5.1A-C, lane 7). Also these 2 DNA fragments should add up to the 2 missing DNA fragments and clearly they do not. The MPS II patient DNA sample probed in Figure 5.1D, lane 15 revealed the loss of a single band at 2.6-kb and the generation of 2 new bands at 1.4-kb and 1.2-kb. The same MPS II patient DNA sample showed no change when compared to unaffected controls when different restriction enzymes were used. A possible explanation for this phenomenon could be that a *Stu*I restriction site has been generated within the 2.6-kb band and as a result upon digestion resolves into 2 bands of 1.4-kb and 1.2-kb respectively. It is unknown whether this change is a mutation or a polymorphism. The Southern blots described in this study (Figure 5.1A-D) were probed with other gene probes and revealed normal DNA patterns indicating that the alterations described above were localised to DNA around the IDS gene and that equivalent amounts of DNA were used for each lane (data not shown).

The nine MPS II patient DNA samples shown in (Figure 5.1A-D, lanes 1, 2, 5, 8, 9, 10, 11, 13, 14) revealed normal Southern patterns when compared to the unaffected

controls which suggests that their mutations were localised to small sequence alterations, such as point mutations, small deletions or insertions. The DNA samples from the two MPS II females (Figure 5.1A-D, lanes 16 and 17) had insufficient DNA to make any positive conclusions, although the weak pattern appeared to be identical to the pattern seen in the unaffected controls.

The 1.5-kb IDS cDNA detected two-allele RFLPs at the IDS locus in DNA digested with *Stu*I (alleles A1 and A2) or *Taq*I (alleles B1 and B2) (Figures 5.2C and D). The polymorphic fragments in the *Stu*I digest were 17.8-kb (A1 allele) and 15 and 2.8-kb (A2 allele). Constant fragments of the following sizes were also detected: 2.6, 3.2, 3.4, 3.6 and 4.7-kb. The polymorphic fragments in the *Taq*I digest were 5.1-kb (B1 allelle) and 3.8-kb (B2 allele). Constant fragments of the following sizes were also detected: 1.1, 1.3, 1.7, 2.0, 2.8 and 3.3-kb. The calculated heterozygote frequencies for the two RFLPs were 0.50 and 0.08, respectively. Among 27 X-chromosomes from unrelated males and females, the haplotype frequencies were A1B1 (0.60), A1B2 (0.04), A2B1 (0.36) and A2B2 (0.0). RFLPs were not detected with the following restriction enzymes: *Bam*HI, *Bcl*I, *Bgl*II, *Eco*RI, *Hinc*II, *Hin*dIII, *Msp*I, *Pst*I, *Pvu*II, and *Sau*3A (Suthers *et al.*, 1990).

Table 5.1 shows a survey of 206 MPS II patients that have been probed with the 1.5-kb IDS cDNA. This data represents a collaborative survey using the 1.5-kb cDNA clone described in Section 4.3.2. The data from eight countries has approximately 18% of the MPS II patients with abnormal Southern blot patterns with 4% showing complete IDS gene deletions. This indicates that the majority of IDS gene mutations are small, i.e. point mutations, small deletions or insertions, and as a result are unable to be detected by Southern analysis. The detection and characterisation of these mutations in some MPS II patients will be described in Section 5.4.1.

Table 5.1 **SOUTHERN ANALYSIS OF 220 MPS II PATIENTS**

The IDS cDNA probe was distributed to other investigators and the data collected in this table is a result of this collaboration.

SOUTHERN ANALYSIS OF IDS IN MPS II PATIENTS

Total Patient Numbers	Number of Patients with Partial Deletion	Complete Deletion			Source
49	12	2	(i)	DXS295	Australia & UK Wilson *et al.* 1991 & P. Nelson
46	7	2	(i) (ii)	DXS466/295 DXS466	Sweden Steen-Bondesen *et al.* 1992
25	4	-			Italia Palmieri *et al.* 1992
20	3	-			France E. Le Guern pers. commun.
11	1	2	(i) (ii)	DXS304 DXS296	Germany A. Gal pers. commun.
7	1	-			Israel G. Bach, pers. commun.
14	2	1		DXS466	Germany F. Herrmann pers. commun.
49	8	-			Japan T. Orii pers. commun.
220	**39 (18%)**	**8 (4%)**			

5.3.2 COMPLETE IDS GENE DELETIONS IN 2 MPS II PATIENTS.

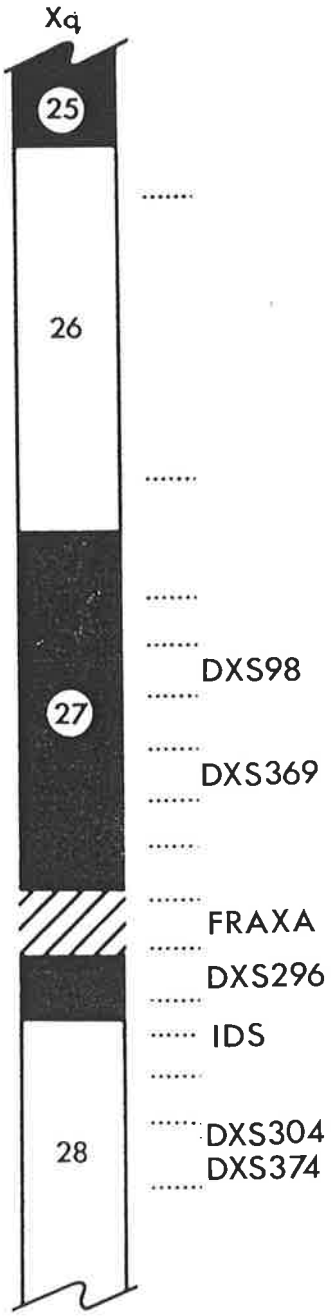
The work described in this section was done in collaboration with Dr. Graeme Suthers and Dr. Sui Yu in the Department of Cytogenetics and Molecular Genetics, Adelaide Children's Hospital.

The filters shown in Figures 5.1A, B, C and D were re-probed to determine whether markers near the IDS gene were deleted in the MPS II patient DNA samples that had no hybridizing bands (lanes 3 and 4). The X-chromosome marker DXS296 described by Suthers *et al.* (1989) was absent in the MPS II patient analysed in lane 4 (Figure 5.1A-D) indicating that the deletion extended proximally from the IDS gene. The deletion in the MPS II patient analysed in lane 3 (Figure 5.1A-D) did not extend to DXS296 and neither deletion included the other X-chromosome markers DXS98 (Kidd *et al.*, 1989), DXS369 (Kidd *et al.*, 1989; Oostra *et al.*, 1990), DXS304 (Vincent *et al.*, 1989; Dahl *et al.*, 1989) or DXS374 (Kidd *et al.*, 1989) (see Figure 5.2 to see the relationship of IDS to these markers). Pulsed field gel electrophoresis studies have demonstrated that DXS296 is approximately 800-kb proximal to the IDS gene (Dr. S. Yu, unpublished data).

Both MPS II patients (lane 3) and (lane 4) had extremely severe clinical phenotypes. They presented in their first year of life with developmental delay and, by their second year, had developed hernias, curvature of their spines, enlarged livers and spleens and developmental regression. In contrast with other patients, neither attained speech and both were troubled by epileptic seizure from an early age. MPS II patient (lane 3) also had a congenital abnormality of the eyelids causing excessive drooping (ptosis). Neither had any other congenital malformations, although the severity of their MPS II

Figure 5.2 **RELATIONSHIP OF VARIOUS X-CHROMOSOME MARKERS TO IDS**

Schematic representation of the distal portion of the long arm of the X-chromosome (Xq26-part of Xq28). The locations of the markers in relation to the chromosome banding are approximate. The fragile site is indicated by the hatched region at distal Xq27.


Xq
25
26
27
28
DXS98
DXS369
FRAXA
DXS296
IDS
DXS304
DXS374

clinical phenotype makes it difficult to exclude minor abnormalities attributable to their deletions extending beyond the IDS gene (Wraith *et al.*, 1991).

The observation of one patient with an extremely severe clinical phenotype and a deletion extending from the IDS gene to an adjacent conserved sequence (DXS296), suggests that some MPS II patients may have additional symptoms due to deletions of genes other than IDS (Schmickel, 1986; Wraith *et al.*, 1991). For example, it was shown that the probe which detects DXS296 (VK21; Suthers *et al.*, 1989) also detects a single conserved *Hin*dIII fragment in both mouse (1.9-kb) and hamster DNA (1.8-kb) (data not shown) therefore it is possible that this evolutionarily conserved sequence when deleted could lead to seizures and ptosis seen in the MPS II patient were the deletion extends beyond this marker. Another example, is the FRAXA site which is adjacent to the IDS gene, the FRAXA site is involved with the fragile X syndrome which has as a dominant feature mental retardation and it is possible that deletions extending to this site could lead to extreme mental retardation in MPS II patients. DXS296 shows no recombination with FRAXA (Suthers *et al.*, 1989), and deletions extending from the IDS gene to DXS296 could conceivably include FRAXA. It is also possibile that deletions extending distally from the IDS gene could encompass genes in Xq28. With the development of a large-scale restriction map of this region it may be possible to correlate the extent of deletions in MPS II patients with the clinical phenotype.

5.3.3 A TRANSLOCATION DISRUPTS IDS IN AN MPS II FEMALE

The work described in this section was done in collaboration with Dr. Graeme Suthers in the Department of Cytogenetics and Molecular Genetics.

An MPS II female patient who has a t(X;5) (q28;q33) reciprocal translocation which results in preferential inactivation of her normal X-chromosome, and is deficient in IDS enzyme activity, has HS and DS-uria, and shows a typical MPS II clinical phenotype (Mossman *et al.*, 1983; Roberts *et al.*, 1989) was shown to have an X-chromosomal breakpoint at the Xq27/28 boundary (Roberts *et al.*, 1989). The translocation was postulated to disrupt the function of IDS, although it was not clear that the translocation breakpoint was actually within IDS. DNA was analysed from a somatic cell hybrid, CY34, which contains the derived X-chromosome from this patient and lacked the Xq28-qter fragment (Suthers *et al.*, 1989). The breakpoint was mapped in relation to DNA markers at Xq28 and the chromosomal breakpoint was found to be between the markers DXS296 and DXS304, the locus order at Xq28 was centromere-DXS98-DSX369-FRAXA-DXS296-IDS-DXS304-DXS374-telomere (see Figure 5.2; Suthers *et al.*, 1989; Mandel *et al.*, 1989).

Figure 5.3 shows that in normal DNA the 1.5-kb cDNA fragment detected eight fragments of 7.5, 5.5, 4.1, 2.5, 2.3, 1.3 and 0.76-kb (not visible) (see lane 1 and 2) whereas, in DNA from the somatic hybrids CY34 (Xpter-Xq28) and CY34A (a subclone of CY34 containing only Xq24-Xq28) (lanes 3 and 4) four fragments were missing (5.5, 2.5, 1.3 and 0.76-kb) and a new 1.0-kb fragment was present but not easily seen on this exposure. In addition, a very faint 14-kb fragment was visible in the mouse background (lane 5). The IDS cDNA fragment did not detect polymorphic *Hin*dIII fragments in DNA samples from 16 normal X-chromosomes. This clearly

Figure 5.3 **SOUTHERN ANALYSIS OF A FEMALE MPS II PATIENT**

Southern blot of *Hin*dIII digested DNA samples. The somatic cell hybrid line CY34 contained Xpter-q28, while CY34A was a sub-clone of CY34 having just Xq24-q28; both cell lines had a mouse (A9) background (Suthers *et al.*, 1989). Approximate DNA fragment size indicators are shown on the left.

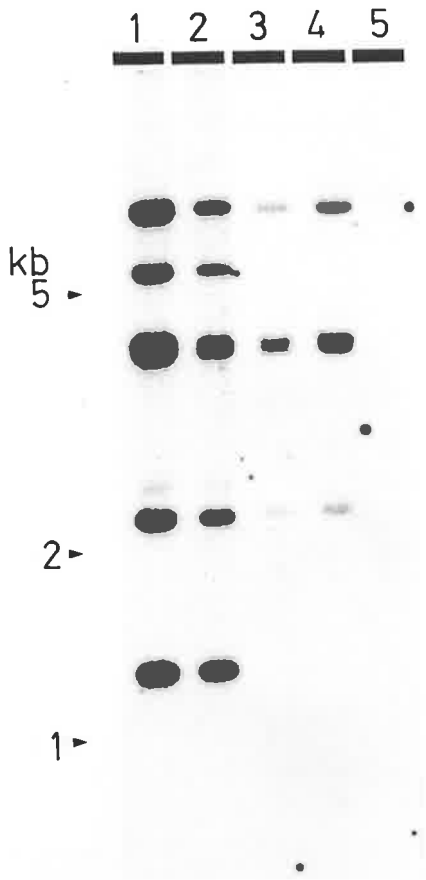
Lane 1: normal female DNA.

Lane 2: normal male DNA.

Lane 3: human/mouse cell line CY34.

Lane 4: human/mouse cell line CY34A.

Lane 5: mouse cell line A9.


1 2 3 4 5
kb
5
2
1

demonstrates that the translocation disrupts IDS and that this gene lies between the markers DXS296 and DXS304 (Wilson *et al.*, 1991). Using the 1.3-kb *Hin*dIII genomic fragment (5'-IDS) (see Chapter 3) to probe DNA from a normal female, a single 1.3-kb fragment was revealed, however when this same probe was used to screen DNA from CY34, no fragment was observed. This indicates that IDS is orientated with the 5'-end on the telomeric side of the CY34 breakpoint ie. the gene is transcribed from the telomere toward the centromere (Wilson *et al.*, 1991).

5.4 MOLECULAR ANALYSIS OF SMALL MUTATIONS IN MPS II PATIENTS

5.4.1 CHEMICAL CLEAVAGE OF IDS cDNA FROM MPS II PATIENTS

The majority of MPS II patients showed normal Southern blot patterns when probed with the 1.5-kb IDS cDNA. It therefore became a goal of this study to determine the nature of the small mutations in some of these MPS II patients. Identification of the mutation may reveal protein structures critical for IDS function. In addition, it was anticipated that this information could be used to establish a correlation between the mutation and the MPS II patient clinical phenotype. With a significant number of MPS II patients screened it may be possible to determine if a specific mutation results in destabilized IDS with activity or whether a specific mutation results in relatively stable IDS with minimal activity. To gain this level of information it is necessary to study the maturation process of IDS and to determine the kinetics of IDS from each MPS II patient.

A group of unrelated MPS II patients were selected on the basis of their clinical phenotypes and their biochemical phenotype, i.e. residual IDS enzyme activity (see

Table 5.2 **MPS II PATIENTS SELECTED FOR SCREENING.**

Table shows the patient number, clinical phenotype and the relative IDS enzyme activity. The source of the patient fibroblasts is also indicated.

Patient	Clinical Phenotype	IDS Enzyme Activity ($pmoles.min^{-1}.mg^{-1}$)	Source
SF 811	S	0.09	NSW
SF 532	M	0.332	GM614 New Jersey Cell Line Repository
SF3362	M	1.04	USA
SF1779	S	0.126	NSW
SF2069	S	nd	NSW
SF 635	M	nd	VIC
SF 970	S	0.288	NSW
SF3055*	M	0.779	UK
SF3254	M	0.666	DM
SF 181	S	nd	VIC
SF 193	S	nd	NSW
SF3116	S	0.019	UK
Normal (n=3)	N	206; 456; 859	

*nd, not done; *, cells were incubated in fetal calf serum. The remaining patient cells were incubated in heat-inactivated fetal calf serum.*

S : Severe

M : Mild

N : Normal

n : number of individual samples

Table 5.2). The reasoning behind this selection was based on the possiblility of identifying protein structures that are crucial to both IDS substrate binding and substrate catalysis. If both mild and severe MPS II patients were screened, these critical domains may be identified. The mild MPS II patients are most likely to have mutations that do not totally disable IDS activity enabling the level of stored GAG to be relatively less. However, the MPS II patients with a severe clinical phenotype are most likely to have mutations that greatly reduce IDS activity. It is therefore possible that identifying the mutations from a broad range of MPS II clinical phenotypes regions involved in catalysis and substrate binding may be located. Total RNA was isolated from MPS II patient fibroblasts and cDNA was generated from transcribed mRNA. Using the PCR-specific IDS primers, two PCR products were generated that covered the entire coding region of IDS. The two PCR products which overlap by 133-bp, are shown in Figure 5.4: primers 16-17 generated the 722-bp PCR product A, and primers 6-15 generated the 1161-bp PCR product B (Figure 2.1 for primer sequence and position). The two PCR products were also generated from the normal IDS sequence in λc2S23 and used to form heteroduplexes with MPS II patient PCR products. The following section describes seven MPS II patients where chemical cleavage and direct DNA sequencing have identified a base change, the possible involvement of these changes in relation to MPS II will be discussed at the the end of this section (see Table 5.3).

The chemical cleavage pattern obtained for the DNA from patient SF811 is shown in Figure 5.5. No abnormal cleavage sites were observed in product A. However, the cleavage pattern from product B was observed to have three reactive sites in the hydroxylamine lane. Both of the primers that generate this PCR product B were end-labelled making it impossible to determine which end of product B the cleavage sites are derived from. The size of the cleavage product was determined from a logarithmic

Figure 5.4 **IDS cDNA AMPLIFICATION PRODUCTS**

IDS mRNA showing the open reading frame (open rectangle) and the position of primers used to generate two PCR products (A and B) that encompass the entire coding region of IDS. The ER signal peptide sequence is represented by the hatched area and the positions of the 42 and 14-kDa polypeptides are indicated.

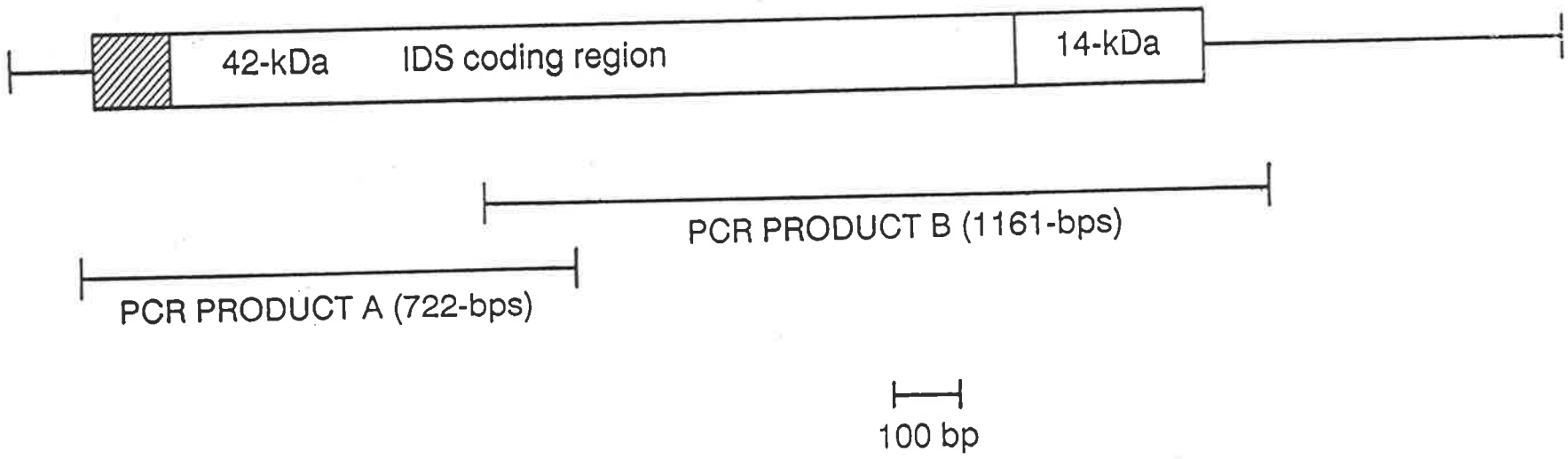

42-kDa
IDS coding region
14-kDa
PCR PRODUCT B (1161-bps)
PCR PRODUCT A (722-bps)
100 bp

Figure 5.5 **CHEMICAL CLEAVAGE OF PCR PRODUCTS FROM PATIENT SF811**

Autoradiograph of chemical cleavage method analysis of heteroduplexes formed between normal IDS and patient PCR templates (B, see Figure 5.4). The heteroduplexes were reacted with hydroxylamine (H) for 40 min and with osmium tetroxide (OT) for 5 min. The samples were electrophoresed on a 6% sequencing gel (Section 5.2.4). Indicated on the Figure are the homoduplex (NORMAL-normal/normal) and the heteroduplex (SF811-normal/patient). The arrows indicate the uncleaved product B (1161-bp) and the size of the patient mutation product (595-bp).

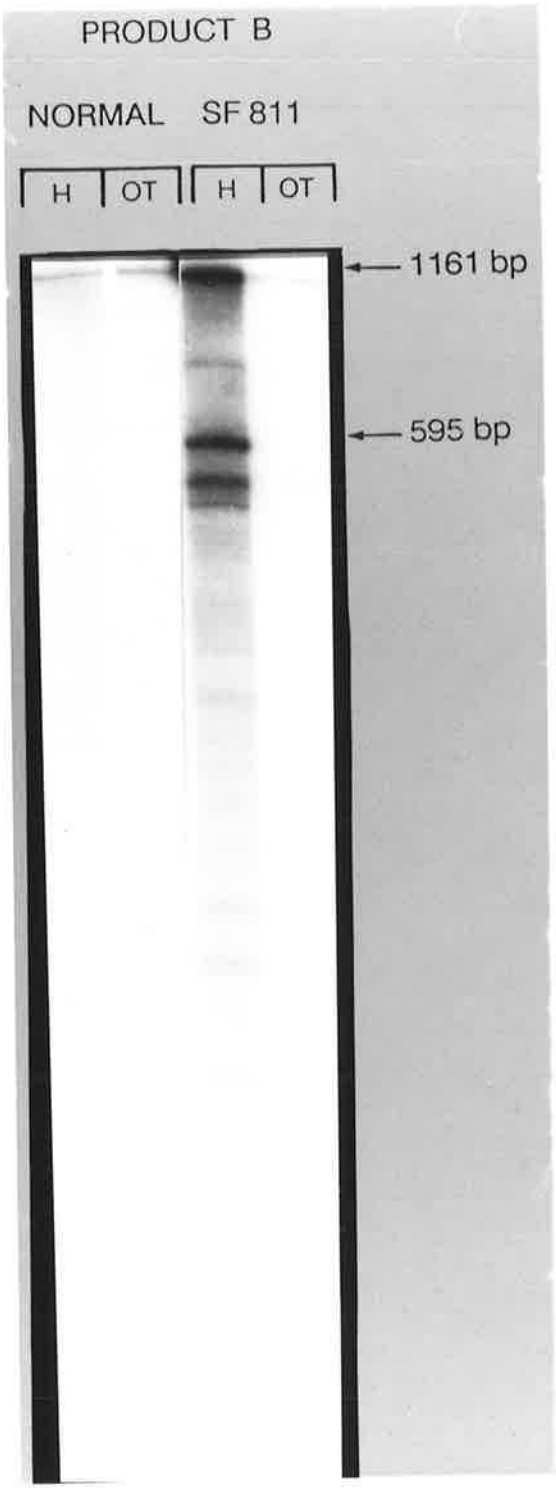
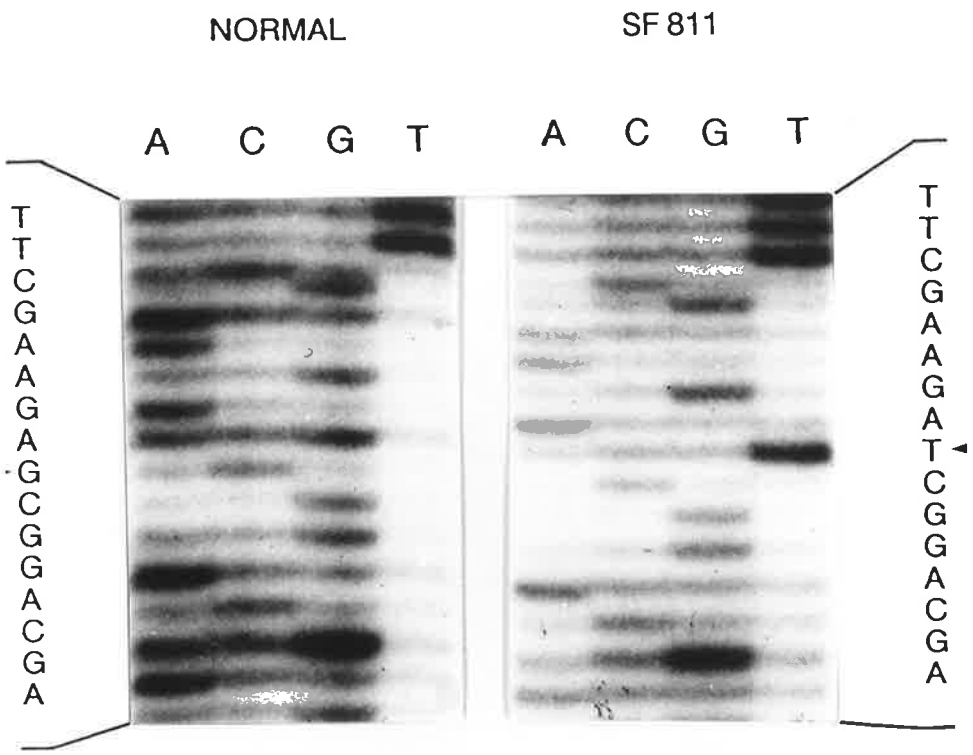

PRODUCT B
NORMAL SF 811
H
OT
H
OT
1161 bp
595 bp

Figure 5.6 **DIRECT PCR SEQUENCING OF THE MUTATION IN PATIENT SF811**

Sequence analysis of the normal and patient SF811 IDS sequence using primer 23. The position of the mutation is indicated with an arrow. The change is a G to a T transition at nucleotide position 1247.


NORMAL
SF 811
A C G T
A C G T
TTCGAAGAGCGGACGA
TTCGAAGATCGGACGA

plot with products of known size, placing the cleavage site in patient SF811 at approximately 700-bp from either end of product B. Primers 23 and 24 were end-labelled and used in direct PCR sequencing reaction to cross the sites of the predicted mutation. Figure 5.6 shows the IDS sequence from patient SF811 in this area compared with normal IDS sequence (Figure 4.3), using primer 23 (see Figure 2.1 for primer position in IDS). The sequence shows a G to T substitution at position 1247, which results in the mutation of a glutamic acid codon at amino acid 375 (GAG) to a termination codon (TAG). Using standard mutation nomencalature this mutation has been named $E_{375}X$. There were no other IDS sequence differences detected in patient SF811 that would explain the presence of the other reactive bands observed in the cleavage reaction. It is possible that the single base pair change (a G to T substitution at nucleotide position 1247) may act to destabilise the surrounding sequence making them more susceptible to chemical modification.

The pattern obtained from chemical cleavage of DNA from patient SF3362 is shown in Figure 5.7. Reactive sites were not observed in product A, however, a reactive site was observed in the osmium tetroxide lane of product B. The position of this cleavage site was calculated to be approximately 500-bp from either end of product B. Primer 22 was end-labelled and used to sequence across the predicted region containing the mutation. Figure 5.8 shows the patient IDS sequence at the position of the cleavage compared with normal IDS DNA sequence. The sequence reveals a 7-bp insertion (GGAACTA) at position 1130 which results in a reading frameshift with a stop codon at nucleotide position 1144, this causes IDS to be produced missing 208 amino acids from the carboxyl end.

The chemical cleavage pattern of DNA from patient SF532 is shown in Figure 5.9. No reactive sites were observed in product B. However, the cleavage pattern from

Figure 5.7 **CHEMICAL CLEAVAGE OF PCR PRODUCTS FROM PATIENT SF3362**

Autoradiograph of chemical cleavage method analysis of heteroduplexes formed between normal IDS and patient PCR templates (B, see Figure 5.4). The heteroduplexes were reacted with hydroxylamine (H) for 40 min and with osmium tetroxide (OT) for 5 min. The samples were electrophoresed on a 6% sequencing gel (Section 5.2.4). Indicated on the figure are the homoduplex (NORMAL-normal/normal) and the heteroduplex (SF3362-normal/patient). The arrows indicate the uncleaved product B (1161-bp) and the size of the patient mutation product (674-bp).

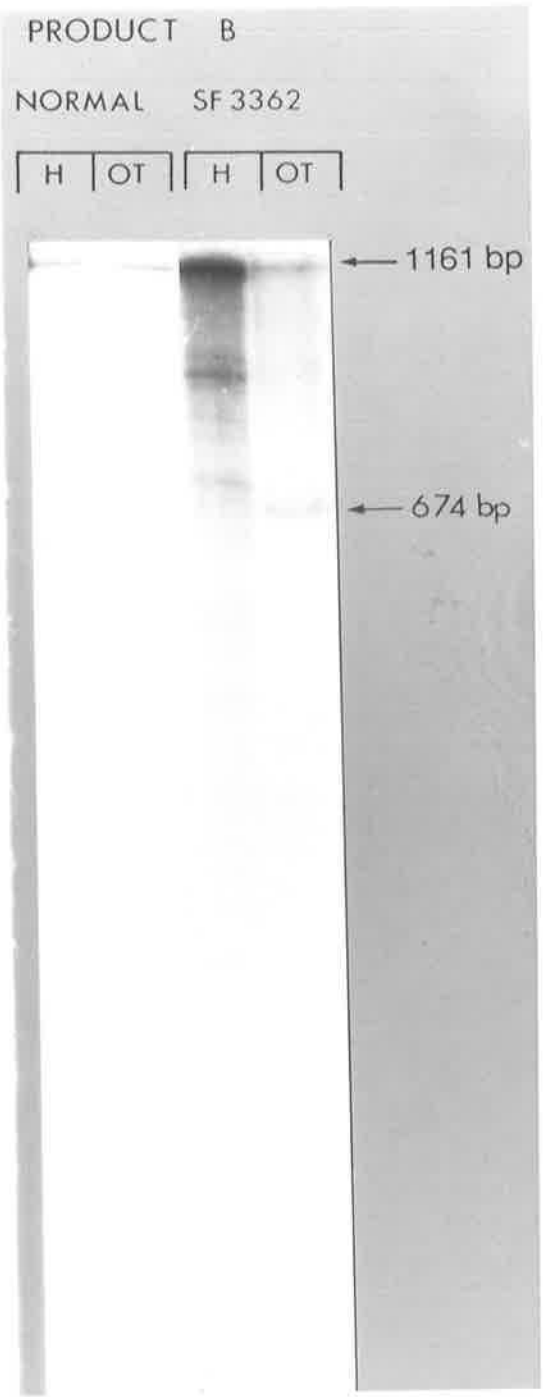

PRODUCT B
NORMAL SF 3362
H
OT
H
OT
1161 bp
674 bp

Figure 5.8 **DIRECT PCR SEQUENCING OF THE MUTATION IN PATIENT SF3362**

Sequence analysis of the normal and patient SF3362 IDS using primer 22. The position of the mutation is indicated with an arrow. The change detected was a seven base pair insertion (GGAACTA) (indicated by the bar) at nucleotide position 1130.

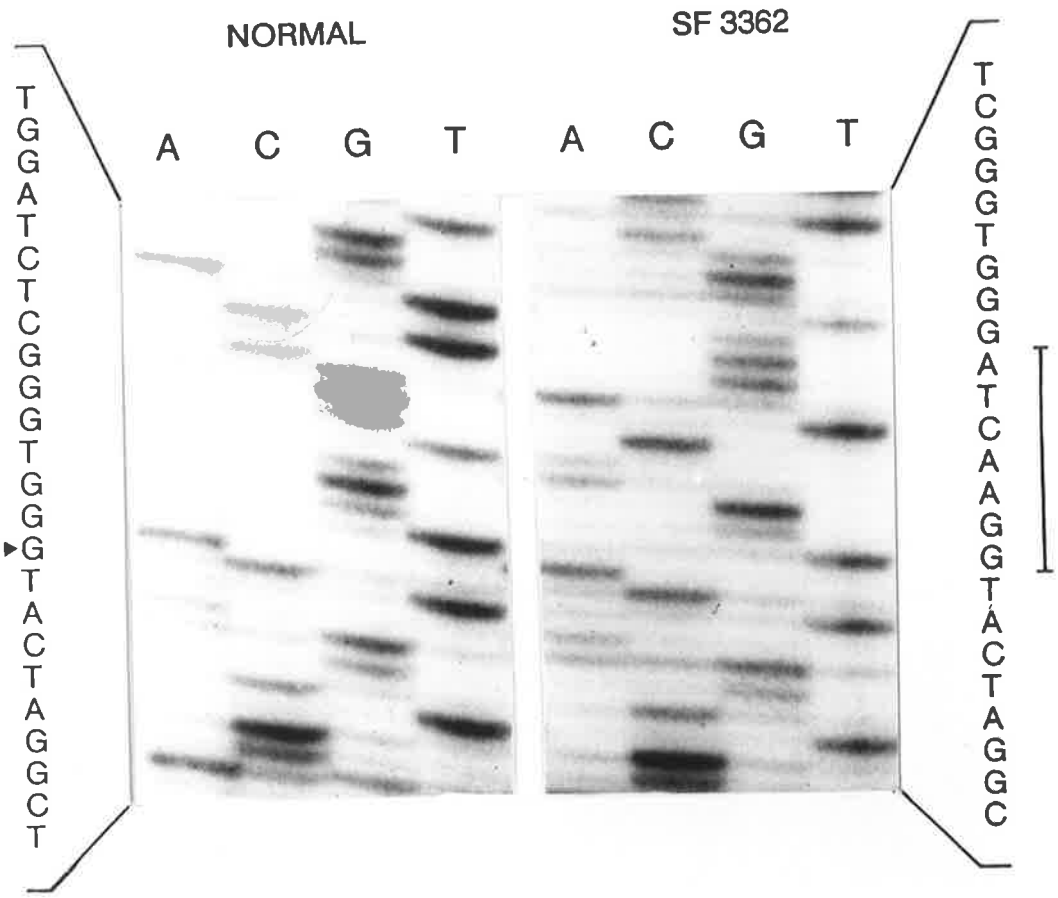

NORMAL
SF 3362
A C G T A C G T
TGGATCTCGGGTGGGTACTAGGCT
TCGGGTGGGATCAAGGTACTAGGC

Figure 5.9 **CHEMICAL CLEAVAGE OF PCR PRODUCTS FROM PATIENT SF532**

Autoradiograph of chemical cleavage method (CCM) analysis of heteroduplexes formed between normal IDS and patient PCR templates (A, see Figure 5.4). The heteroduplexes were reacted with hydroxylamine (H) for 40 min and with osmium tetroxide (OT) for 5 min. The samples were electrophoresed on a 6% sequencing gel (Section 5.2.4). Indicated on the Figure are the homoduplex (NORMAL-normal/normal) and the heteroduplex (SF532-normal/patient). The arrows indicate the size of the patient mutation product (167-bp).

PRODUCT A

NORMAL SF 532

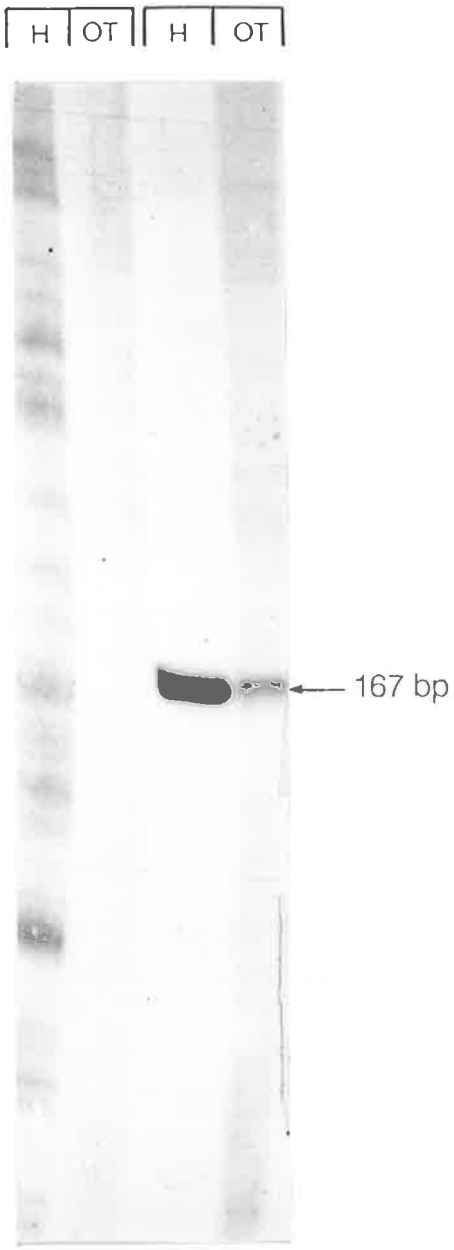
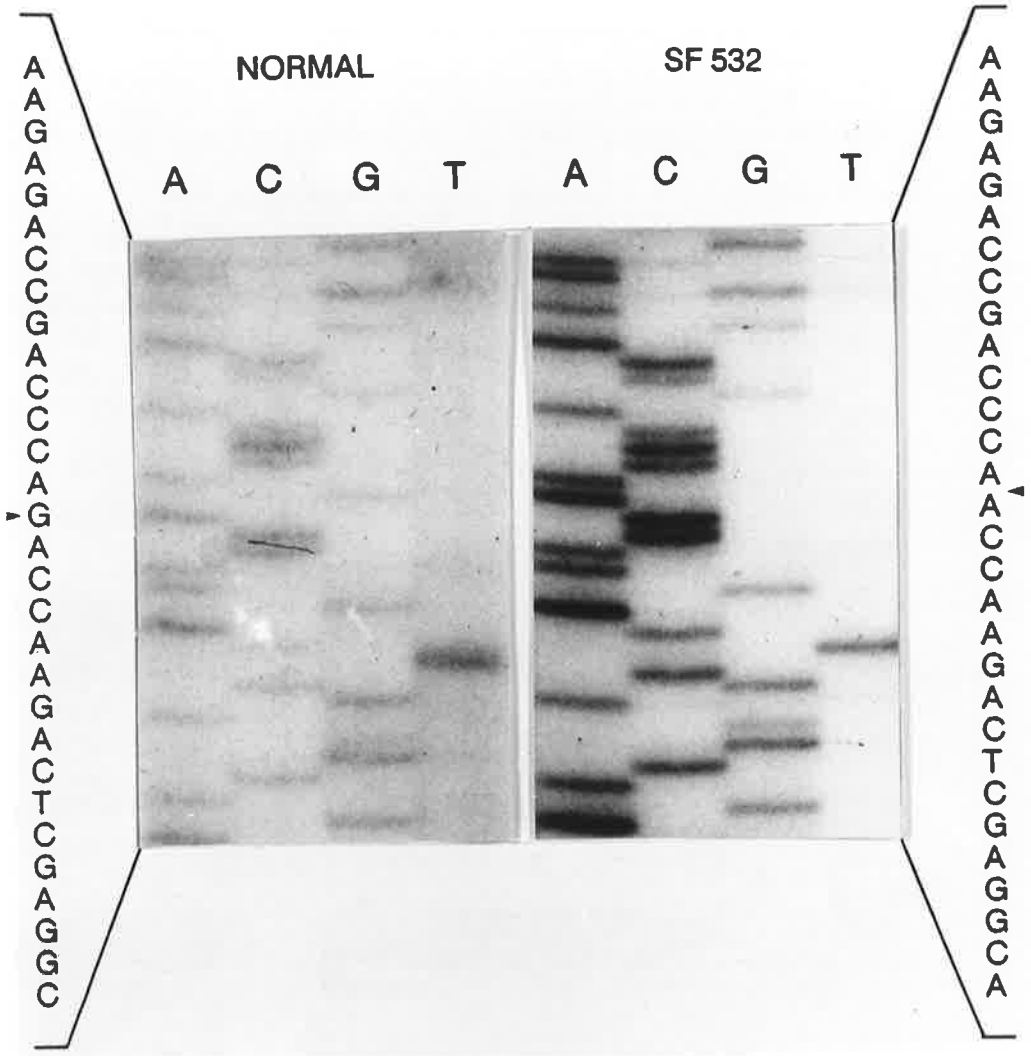


Figure 5.10 **DIRECT PCR SEQUENCING OF THE MUTATION IN PATIENT SF532**

Sequence analysis of the normal and patient SF532 IDS using IDS primer 25. The position of the mutation is indicated with an arrow. The change detected was a single base pair deletion at nucleotide position 167.


NORMAL
SF 532
A C G T A C G T
AAGAGACCGACCCAGACCAAGACTCGAGGC
AAGAGACCGACCCAACCAAGACTCGAGGCA

product A revealed a strongly reactive site in the hydroxylamine lane, indicating a mismatched cytosine or guanine (depending on which strand was cleaved). It was calculated that the cleavage site was approximately 150-bp from either end of product A. Figure 5.10 shows the IDS DNA sequence of patient SF532 with primer 25 which reveals a 1-bp deletion (C) at position 167 when compared with normal IDS sequence. This change was the only sequence difference detected in PCR product A. This mutation predicts a reading frameshift which results in a termination codon at nucleotide position 173.

The chemical cleavage pattern of DNA from patient SF635 is shown in Figure 5.11 . No reactive sites were observed in product A. However, 2 strong reactive sites were seen in both the hydroxylamine and osmium tetroxide lanes of PCR product B. The nature of the cleavage sites seen in product B suggest that the mutation affects a large area on both strands of DNA. The position of the mutation was predicted to lie approximately 600-bp from either end of product B. Figure 5.12 shows that when the patient IDS was sequenced with primer 23, there was found to be 60-bp missing when compared to the normal IDS sequence. It is uncertain that this is a deletion. The reason for the uncertainty is the sequence surrounding the site of the possible deletion is composed of a small repeat (**AGGCAGGC**) that contains splicing consensus sites (AG/intron sequence). It is therefore possible that this mutation is an mRNA splicing defect. It is therefore possible that the 60-bp of IDS sequence missing in this patient could represent a single exon. To determine which is the case the intron/exon structure of IDS is required so that the intron/exon boundary of patient SF635 at the area of this mutation can be spanned and the integrity of the splice junction in SF635 determined.

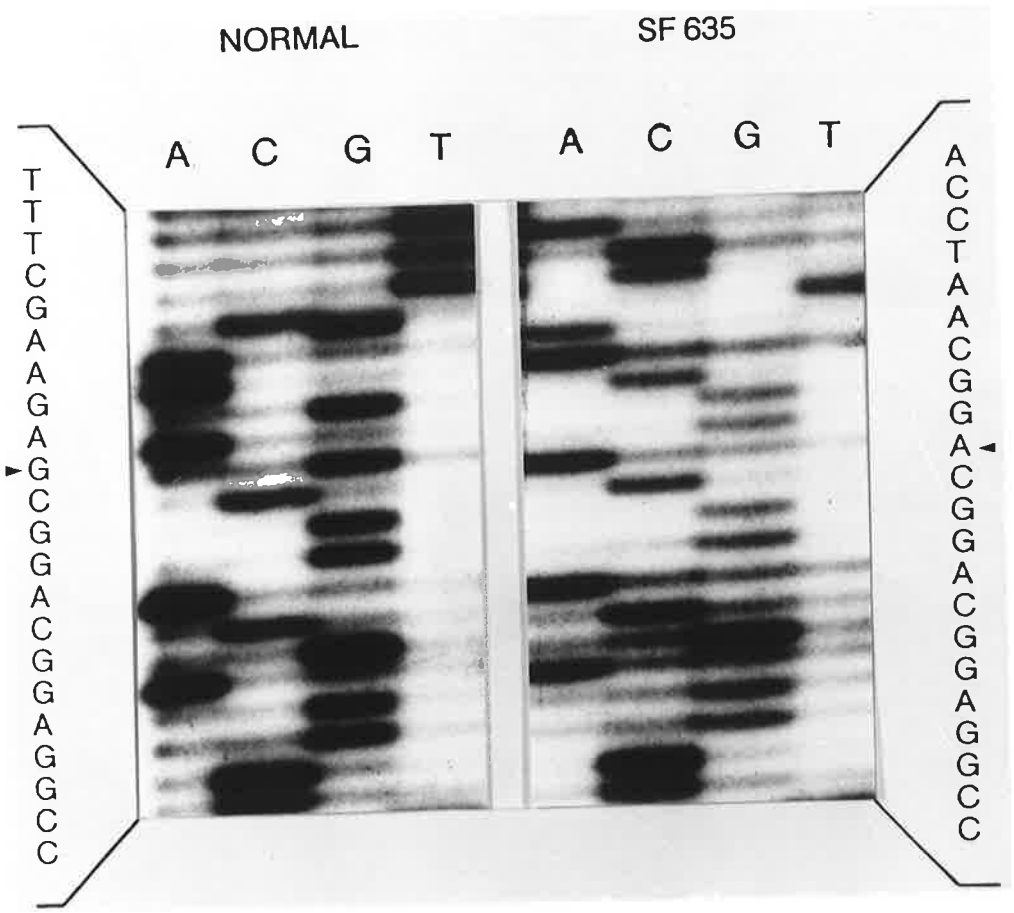
Figure 5.11 **CHEMICAL CLEAVAGE OF PCR PRODUCTS FROM PATIENT SF635**

Autoradiograph of chemical cleavage method analysis of heteroduplexes formed between normal IDS and patient PCR templates (B, see Figure 5.4). The heteroduplexes were reacted with hydroxylamine (H) for 40 min and with osmium tetroxide (OT) for 5 min. The samples were electrophoresed on a 6% sequencing gel (Section 5.2.4). Indicated on the Figure are the homoduplex (NORMAL-normal/normal) and the heteroduplex (SF635-normal/patient). The arrows indicate the uncleaved product B (1161-bp) and the size of the two patient mutation products (595-bp and 506-bp).


PRODUCT B
NORMAL SF 635
H
OT
H
OT
1161 bp
595 bp
506 bp

Figure 5.12 **DIRECT PCR SEQUENCING OF THE MUTATION IN PATIENT SF635**

Sequence analysis of the normal and patient SF635 IDS using primer 23. The position of the mutation is indicated with an arrow. The change detected was a sixty base pair deletion (see text for explanation) at nucleotide position 1247.


NORMAL
SF 635
A C G T
A C G T
TTTCGAAGAGCGGACGGAGGCC
ACCTAACGGACGGACGGAGGCC

The chemical cleavage pattern of DNA from patient SF1779 is shown in Figure 5.13. No reactive sites were observed in product A. However, the cleavage pattern from product B revealed a strong reactive band in the hydroxylamine lane. The position of this cleavage site was difficult to calculate due to its location at the very top of the gel. An estimation was made putting the cleavage site at approximately 200-bp from either end of product B. Appropriate primers were selected and the patient product was sequenced. Figure 5.14 shows that when the patient IDS was sequenced with primer 31, a C to G substitution was found at position 1526 when compared with the normal IDS sequence. This results in the mutation of an arginine at position 469 (CGG) to a glycine (GGG) ($R_{469}G$).

The chemical cleavage pattern of DNA from patient SF2069 is shown in Figure 5.15 . No cleavage sites were observed in PCR product A. However, a strong cleavage site was observed in the osmium tetroxide lane of product B. In this case it was much easier to calculate the position of the cleavage site as the position was much lower on the gel. Calculations placed the cleavage site at around 150-bp from either end of PCR product B. Primers were selected and, in Figure 5.16 using primer 6, an A to G substitution was detected when the patient IDS sequence was compared with normal IDS sequence. This results in the mutation of a lysine (AAG) codon at position 228 to a glutamine codon (CAG) ($K_{228}Q$).

The chemical cleavage pattern of DNA from patient SF3055 is shown in Figure 5.17. No cleavage sites were observed in product A. However, a strong cleavage site was observed in the hydroxylamine lane of product B. The position of the cleavage site was calculated to lie approximately 400-bp from either end of product B. Using primer 22 a C to T substitution was identified at position 1081 when compared to the normal

Figure 5.13 **CHEMICAL CLEAVAGE OF PCR PRODUCTS FROM PATIENT SF1779**

Autoradiograph of chemical cleavage method analysis of heteroduplexes formed between normal IDS and patient PCR templates (B, see Figure 5.4). The heteroduplexes were reacted with hydroxylamine (H) for 40 min and with osmium tetroxide (OT) for 5 min. The samples were electrophoresed on a 6% sequencing gel (Section 5.2.4). Indicated on the Figure are the homoduplex (NORMAL-normal/normal) and the heteroduplex (SF1779-normal/patient). The arrows indicate the uncleaved product B (1161-bp) and the size of the patient mutation product (874-bp).

PRODUCT B

NORMAL SF 1779

Figure 5.14 **DIRECT PCR SEQUENCING OF THE MUTATION IN PATIENT SF1779**

Sequence analysis of the normal and patient SF1779 IDS using primer 31. The position of the mutation is indicated with an arrow. The change detected was a C to a G transition at nucleotide position 1526.

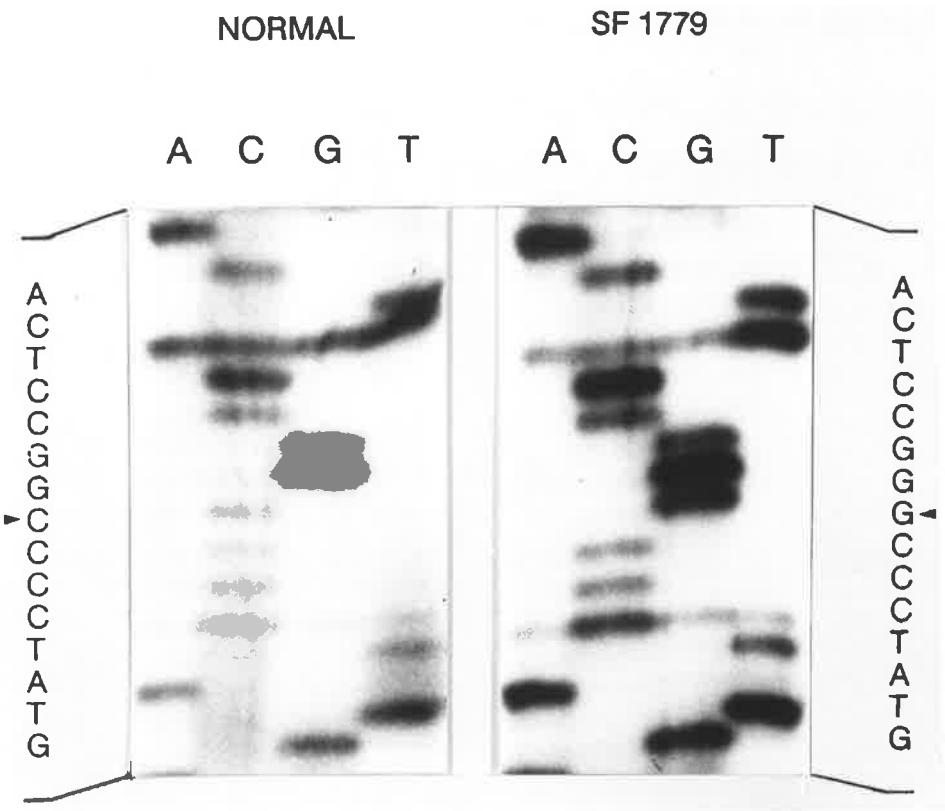

NORMAL
SF 1779
A C G T
A C G T
ACTCCGGGCCCCTATG
ACTCCGGGCCCCTATG

Figure 5.15 **CHEMICAL CLEAVAGE OF PCR PRODUCTS FROM PATIENT SF2069**

Autoradiograph of chemical cleavage method analysis of heteroduplexes formed between normal IDS and patient PCR templates (B, see Figure 5.4). The heteroduplexes were reacted with hydroxylamine (H) for 40 min and with osmium tetroxide (OT) for 5 min. The samples were electrophoresed on a 6% sequencing gel (Section 5.2.4). Indicated on the Figure are the homoduplex (NORMAL-normal/normal) and the heteroduplex (SF2069-normal/patient). The arrows indicate the uncleaved product B (1161-bp) and the size of the patient mutation product (152-bp).

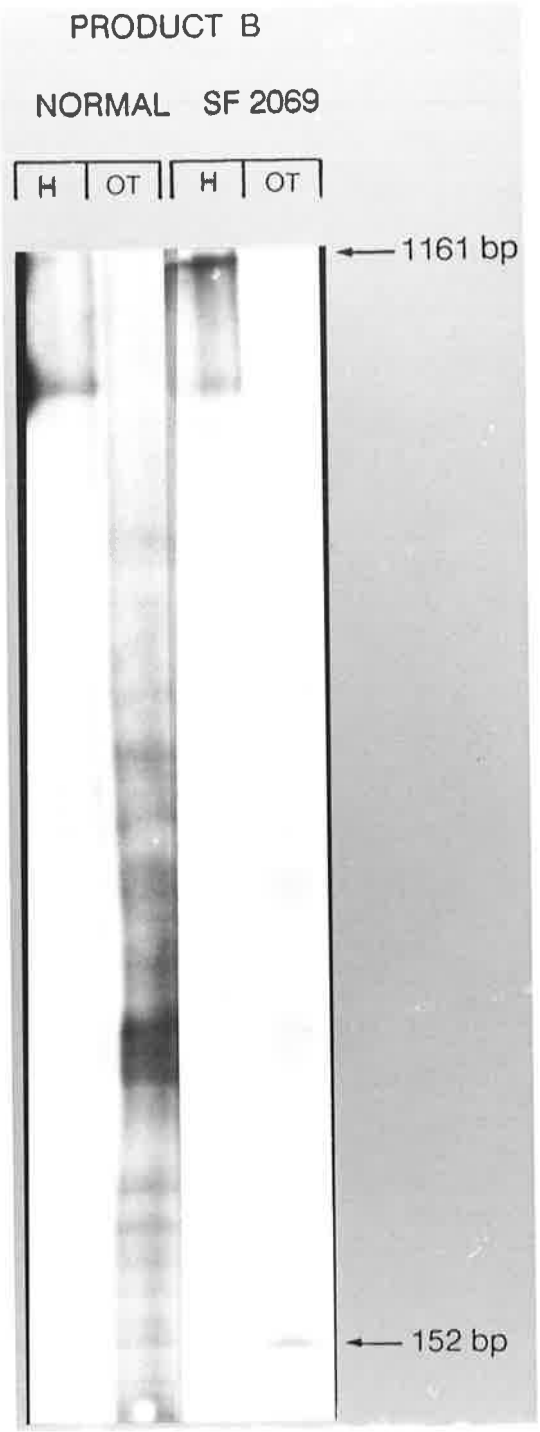

PRODUCT B
NORMAL SF 2069
H
OT
H
OT
1161 bp
152 bp

Figure 5.16 **DIRECT PCR SEQUENCING OF THE MUTATION IN PATIENT SF2069**

Sequence analysis of the normal and patient SF2069 IDS using primer 6. The position of the mutation is indicated with an arrow. The change detected was an A to a C transition at nucleotide position 803.

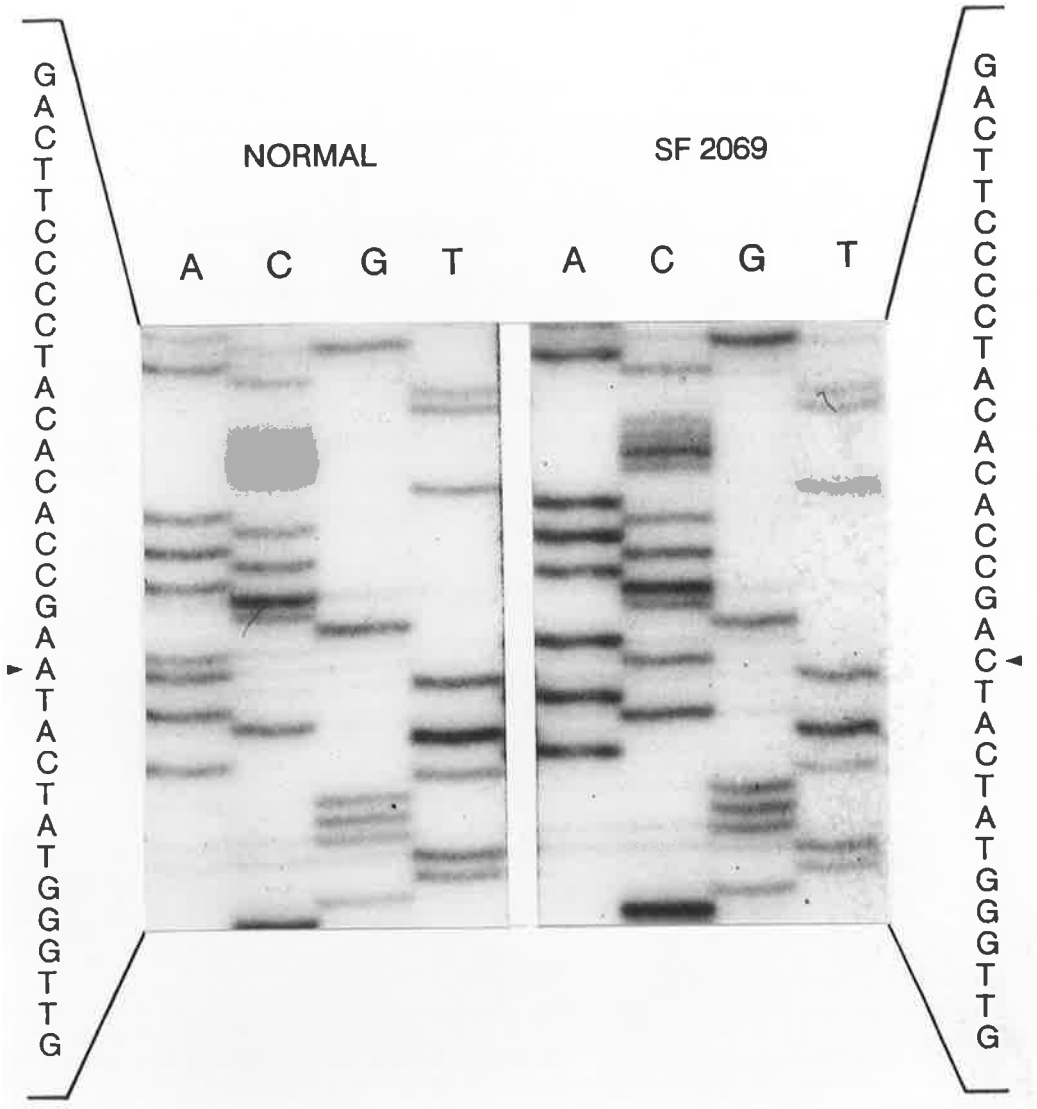

NORMAL
SF 2069
A C G T A C G T
GACTTCCCCCTACACACCGAATACTATGGGTTG
GACTTCCCCCTACACACCGACTACTATGGGTTG

Figure 5.17 **CHEMICAL CLEAVAGE OF PCR PRODUCTS FROM PATIENT SF3055**

Autoradiograph of chemical cleavage method analysis of heteroduplexes formed between normal IDS and patient PCR templates (B, see Figure 5.4). The heteroduplexes were reacted with hydroxylamine (H) for 40 min and with osmium tetroxide (OT) for 5 min. The samples were electrophoresed on a 6% sequencing gel (Section 5.2.4). Indicated on the Figure are the homoduplex (NORMAL-normal/normal) and the heteroduplex (SF3055-normal/patient). The arrows indicate the uncleaved product B (1161-bp) and the size of the patient mutation product (429-bp).

PRODUCT B

NORMAL SF 3055

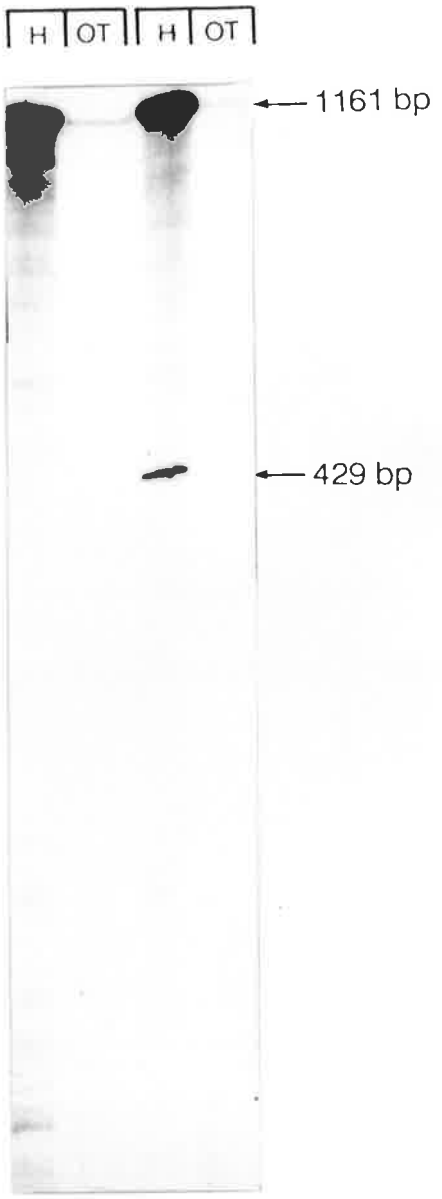
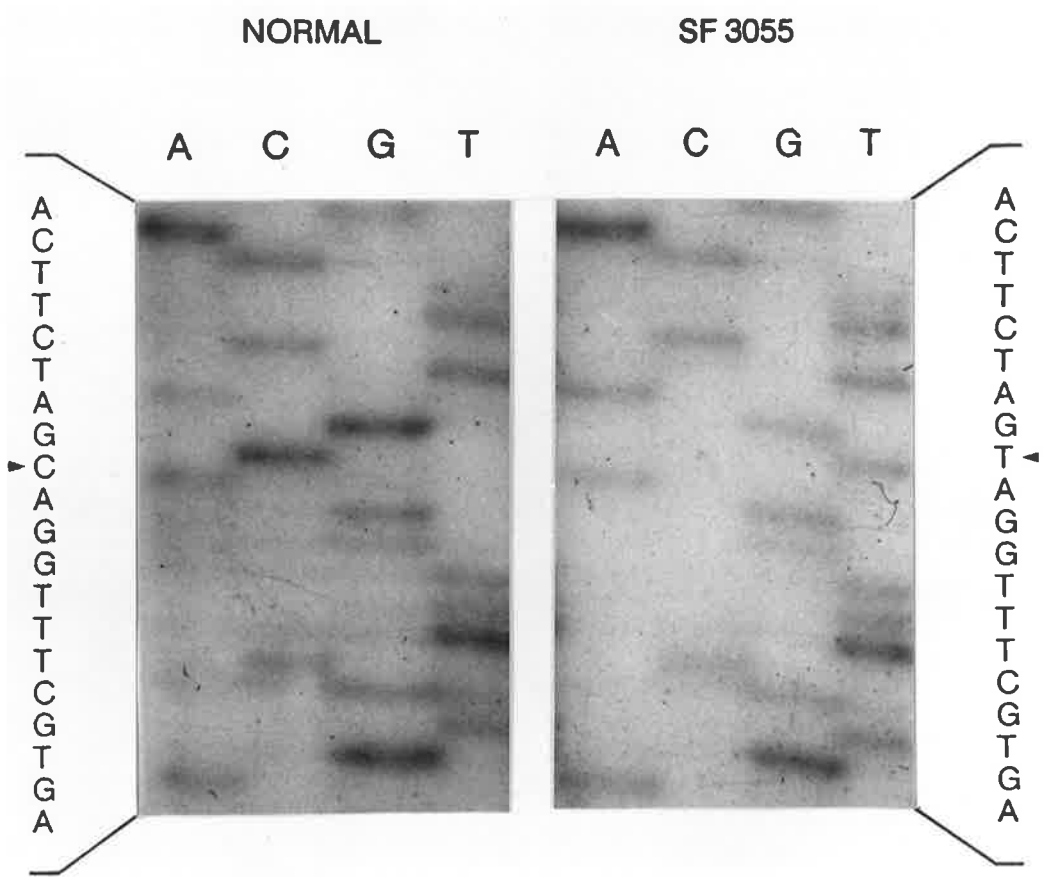


Figure 5.18 **DIRECT PCR SEQUENCING OF THE MUTATION IN PATIENT SF3055**

Sequence analysis of the normal and patient SF3055 IDS using primer 22. The position of the mutation is indicated with an arrow. The change detected was a C to a T transition at nucleotide position 1081.


NORMAL
SF 3055
A C G T
A C G T
ACTTCTAGCAGGTTTCGTGA
ACTTCTAGTAGGTTTCGTGA

sequence (Figure 5.17). This change is in the third base of an aspartic acid codon at position 319 (GAC to GAT) and is therefore most likely to be a polymorphism. The mutation that causes MPS II in patient SF3055 remains undetermined. It is possible that the mutation has been missed due to its being more resistant to chemical attack, i.e. it could be a T.G mismatch (Forrest *et al.*, 1991), as discussed in Section 5.1. This patient will need to be re-screened using both labelled patient IDS and labelled control IDS to establish if this possibility is correct.

Chemical cleavage was performed on patients SF970 and SF3254, revealing strong cleavage sites in product B of both patients. The sequence of the mutation in SF970 was not determined. However, the sequence of the entire PCR product B in patient SF3254 was determined and no change from the normal IDS sequence was detected. It is unknown why the mutation was undetected, although one possiblity is that a PCR contamination may have occurred in the patient DNA sample. Alternatively, it may be that the area of the mutation in the patients IDS was not sequenced. Chemical cleavage will have to be repeated on DNA from patient SF3254 to establish if this cleavage site is real and perhaps if more molecular standards could be included so that the position of the cleavage site in product B may be determined more accurately.

Three other MPS II patients were screened using chemical cleavage and these included, SF181, SF193 and SF3116. However, no reactive sites were observed in either PCR product (A and B). This experiment will have to be repeated, using both labelled normal control DNA and labelled patient DNA, to determine if the mutations are indeed of the resistant type, i.e. T.G mismatches. Until this experiment is performed the nature of the mutations resulting in MPS II in these patients will remain unknown.

The mutations detected in six unrelated MPS II patients are summarised in Table 5.3 and the position of each mutation is indicated in a schematic diagram of the IDS coding region. From this small survey it would appear that the mutations are spread throughout the IDS coding region. The possible involvement of these changes resulting in the MPS II phenotype in these patients will now be discussed.

Compared to the wild type the SF811 patient $E_{375}X$ IDS will be missing 177 amino acids at the C-terminus (Table 5.3). It would be expected that the loss of 26% of the amino acid sequence of IDS would have a detrimental effect on IDS function. Patient SF811 was reported to have a severe clinical phenotype (Table 5.2). The homology of IDS when compared with other sequenced sulphatases is poor at the carboxyl end of the protein this area may not be absolutely essential for IDS function. It is possible that this mutation results in active IDS but the protein is destabilised which results in its rapid degradation in the ER or lysosome (Lodish, 1988). In addition, the tertiary folding of IDS maybe affected, which prevents IDS from either binding substrate or degrading it. When an IDS monoclonal antibody is produced studies of the maturation process of IDS in patient SF811, as has been achieved with another MPS condition (Brooks *et al.*, 1991), may reveal more about the events that result in the absence of IDS activity and a severe clinical phenotype.

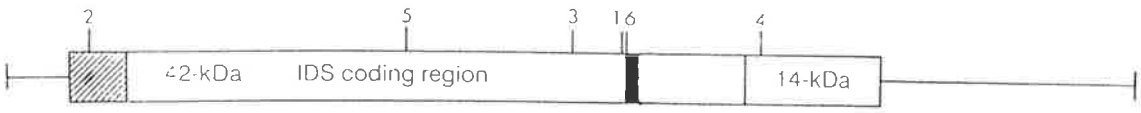
The mutation in patient SF3362 also predicts a truncated IDS (Table 5.3). This patient is missing 208 amino acids from the carboxyl terminus. Given the effect that the absence of 177 amino acids from the carboxyl terminus had on patient SF811 it may be predicted that patient SF3362 would also have a severe clinical phenotype. However, patient SF3362 has a mild clinical phenotype (Table 5.2). Since patient SF3362 has a mild clinical presentation, the C-terminal portion of IDS is clearly not essential for enzyme function, but, in the case of patient SF811, the amount of IDS

Table 5.3 **SUMMARY OF MPS II MOLECULAR MUTATIONS**

A summary of all the mutations found in the MPS II patients studied in this thesis. A schematic representation of the position of each mutation in the coding region of IDS is shown below the table.

	Patient	Mutation	Nucleotide Position	Effect of Mutation	Clinical Phenotype
1.	SF811	G-T	1247	Glu_{376} - Stop	Severe
2.	SF532	del C_{167}	167	Frameshift	Mild*
3.	SF3362	ins GGAACTA	1130	Frameshift	Mild
4.	SF1779	C-G	1526	Arg_{469} - Gly	Severe
5.	SF2069	A-C	803	Lys_{228} - Gln	Severe
6.	SF635	60 bp del/splice	1247	Loss of coding sequence	Mild

**, at this stage the clinical phenotype in this patient is still to be confirmed*



protein sequence removed results in a destabilised protein (perhaps by exposing a protease-sensitive tail) and is rapidly degraded. Whereas in patient SF3362 the larger amount of IDS protein sequence removed may result in a relatively more stable IDS protein with reduced enzyme activity, which enables the reduction of MPS storage in patient SF3362 cells. The residual IDS activity seen in this patient is consistent with a mild MPS II phenotype (see Table 5.2).

The position of the mutation in SF635 is very similar to patient SF811 and patient SF3362 (Table 5.3), however, the mutation that was detected results in the loss of only 20 amino acids in a relatively non-conserved area of the sulphatases. The loss of coding sequence is much more extensive in the other 2 MPS II patients, so it is uncertain what this change would do to IDS function. Patient SF635 was reported to have a mild clinical phenotype with residual IDS activity (Table 5.2). A number of possibilities may explain this result. One possiblity is that, if this mutation is a splicing defect, then a small level of normal IDS mRNA may be produced which is able to correct the majority of storage in this MPS II patients cells and this may also explain the relatively high IDS activity (Table 5.2). To confirm if this is the case the IDS mRNA levels in this patient would need to be measured (no normal PCR product was observed from RNA). Another possibility could be that IDS is still able to function at low levels with the absence of only 20 amino acids which may not play a significant role in the functioning of IDS.

An interesting mutation was found in patient SF532 (Table 5.3). This mutation occurs in the ER signal sequence of IDS, and therefore no IDS should be produced and therefore the patient would be expected to have a severe clinical phenotype. However, the patient was reported to have a mild clinical phenotype and his fibroblasts contained low levels of residual IDS enzyme activity, similar to levels seen in mildly

affected patients (Table 5.2). In some cases it has been reported that cryptic initiation methionines may be used instead of the primary initiation methionine. That is, the ribosome re-initiates translation at the next internal methionine (Kozak, 1985, 1986). However, the next internal methionine occurs at codon 131, well into the protein and beyond several regions of high sequence homology seen in all sulphatases examined (Section 4.3.6, Figure 4.5). The N-terminal third of IDS is therefore believed to be of critical importance for IDS enzymic function, casting doubt on the idea of re-initiation of translation as an explanation. The authenticity of this result was confirmed by repeating the experiment from freshly revived cultured skin fibroblasts, reducing the possibility of a mix-up with patient cell lines in our cell culture facility. The cell line, obtained from the New Jersey Cell Line Repository, may have been mixed at the original source. This cell line, and the mother's cells have been re-ordered to enable the MPS II mutation analysis to be repeated. Also an MPS II cell line from the same contributor and described as having a severe MPS II clinical phenotype where supplied to the New Jersey Cell Line Repository at the same time. This cell line has been requested in case the two cell lines have been accidently mixed. Until these cell lines are obtained and the experiment is repeated the MPS II mutation in this patient remains unexplainable.

Patient SF1779 was observed to have a severe clinical phenotype and was found to have minimal IDS enzyme activity (Table 5.2). It is interesting to note that the position of this mutation is within the 14-kDa N-terminal amino acid sequence and close to the carboxyl terminus of IDS. The observation that patient SF3362 was missing 208 amino acids from the carboxyl terminus yet was observed to have a mild clinical phenotype raises some interesting questions. For example, why does a single amino acid change in an area of IDS thought to have little involvment in enzyme function result in a severe clinical phenotype? It may be possible that the non-

conservative amino acid change acts to destabilise the IDS protein resulting in degradation within the ER or the protease environment of the lysosome prior to maturation to 42 and 14-kDa polypeptides. Alternatively, the amino acid change may result in aberrant tertiary folding in IDS, causing loss of catalytic activity or the loss of substrate binding recognition domains.

Patient SF2069 was also found to have a single base change that resulted in an amino acid substitution (Table 5.3). Patient SF2069 was also observed to have a severe clinical phenotype (Table 5.2). Both the base changes in MPS II patients SF1779 and SF2069 result in non-conservative amino acid substitutions, from positively-charged amino acids ($R_{469}G$) to polar but uncharged amino acids ($K_{228}Q$). To establish their role in MPS II in these patients both of these mutations will have to be introduced into normal IDS. This can be achieved by using site-directed mutagenesis and expression in cell culture. With the use of IDS monoclonal antibodies the effect that these base changes have on the kinetics of IDS and IDS levels can be determined. If these base changes are found to be responsible for the loss of IDS enzyme function in these patients it poses interesting questions relating to the importance of specific protein sequences in IDS.

From the MPS II patients screened in this study it is not possible to calculate the frequencies of these mutations in the MPS II population. It is very likely that they all will turn out to be unique. Of the twelve MPS II patients screened with the chemical cleavage method, all patients with reactive sites showed unique cleavage patterns. This, taken with the variation shown in Southern blot patterns of MPS II patients with gross structural alterations confirms the heterogeneous nature of mutations in MPS II (Wilson *et al.*, 1991). Approximately one-third of MPS II mutations will be new according to the theory of Haldane, (1935) and the remaining mutations will be either

unique or rare, as has been found with other X-linked genetic disorders, i.e. Haemophilia A (Tuddenham *et al.*, 1991). A more extensive survey of MPS II patients will be required before an accurate correlation can be drawn between the molecular nature of individual mutations and the resultant MPS II patient phenotypes. It is also clear that a multidisciplinary approach will be required to help with the prediction of phenotype from the molecular data. A study of the maturation process of mutant IDS may indicate areas of the protein that are not required for activity. Also, the tertiary structure of IDS may indicate the amino acid domains important for substrate binding and catalytic activity. The protein alignment of all the known sulphatases has greatly reduced the number of universally conserved amino acid residues and therefore the residues that are most likely to be involved in IDS function (Section 4.3.6, Figure 4.5). Using this information it may be possible to carry out site-directed mutagenesis studies on these conserved residues and possibly identify these critical IDS residues. For further discussion on the possibilities for future work using the data collected in this study see Chapter 6.

CHAPTER SIX

CONCLUDING DISCUSSION

The work presented in this thesis was primarily aimed at the isolation and characterisation of the IDS gene, which is the gene responsible for the lysosomal storage disorder, MPS II. The isolation of IDS cDNA enabled the characterisation of the molecular defect in MPS II patients and enabled preliminary attempts to relate MPS II mutations to the different clinical phenotypes associated with MPS II patients. The predicted IDS amino acid sequence was shown to have significant homology with other known sulphatases, suggesting that the sulphatases are evolutionarily related and arose from a common ancestral sulphatase by gene duplication and divergent evolution. It was shown that the IDS cDNA clone contained the full coding potential of IDS when it was inserted into an expression vector and the expressed protein showed activity towards specific IDS substrates. This cDNA sequence was also used to demonstrate structural alterations to the IDS gene and to characterise specific point mutations and other small mutations in MPS II patients. The following discussion will include a brief synopsis of what was known about the IDS gene before this thesis was commenced. Also a general discussion of the contribution that this work has made to the concept of a sulphatase gene family. The work has for the first time has allowed a means to determine the genotype detail of the mutations that result in the MPS II phenotype. The number of heterogenous mutations in MPS II was found to be consistent with the broad clinical phenotype seen in this disorder. The availability of a full-length IDS cDNA also opens up the prospects of investigating MPS II patient treatment through the avenues of enzyme-replacement and gene-replacement therapy.

At the commencement of the work described in this thesis the enzyme defect in MPS II was known to be a deficiency of IDS (Bach *et al.*, 1973; Coppa *et al.*, 1973 and Sjoberg *et al.*, 1973). Although MPS II was known to be X-linked since 1946 (Nja, 1946) it was not until the 1980s that the location of the IDS gene was for the first time

pinpointed to the long arm of the X-chromosome (Mossman *et al.*, 1983; Upadhyaya *et al.*, 1986 and Roberts *et al.*, 1989). This information was extremely useful in confirming the identity of isolated IDS clones (discussed in Chapter 3). Many attempts had been made to purify IDS but for the most part, the low abundance prevented the isolation of homogeneous protein preparations or the analysis of IDS amino acid sequence (this is discussed fully in Section 1.4.2). Bielicki *et al.* (1990) reported the half million fold purification of IDS from human liver and were the first to obtain N-terminal amino acid sequence from an IDS 42-kDa polypeptide species. This very limited amount of peptide sequence data, shown to have some homology to N-terminal amino acid sequence obtained from other sulphatases, provided the impetus to commence the study described in this thesis. Using this amino acid sequence information it was possible to isolate a cDNA that encoded IDS (Wilson *et al.*, 1990; see Chapter 4 in this thesis). The cDNA described in this thesis was used for *in situ* hybridization of metaphase chromosomes and independently confirmed the localization of IDS to Xq28 (Wilson *et al.*, 1991).

The work in this thesis describes for the first time the isolation of IDS cDNA and genomic clones and the deduction of the primary amino acid sequence of IDS. These objectives have been the goal of several groups and indicates the internationally competitive nature of the work presented in this thesis.

To date nine sulphatases have been identified in higher eukaryotic cells and all but one of these (steroid sulphatase, or arylsulphatase C (ASC)-Yen *et al.*, 1987; Stein *et al.*, 1989b) are found in the lysosome. Sulphatases have been classed according to their ability to bind and degrade artifical arylsulphate substrates such as 4-methylumbelliferyl sulphate. The sulphatase enzymes that can bind and degrade this

substrate are known as the arylsulphatases, including ASA (Stein *et al.*, 1989a), ASB (Peters *et al.*, 1990) and ASC (Stein *et al.*, 1989b) and those that have a considerably reduced ability to degrade this substrate are known as the non-arylsulphatases i.e. IDS (Bielicki *et al.*, 1990, 1992) and G6S (Robertson *et al.*, 1988, 1992) (Section 4.3.6). However, it should be noted that each of the nine sulphatases whether aryl- or nonarylsulphatases have absolute specificity towards their natural substrates and it is generally more accurate to describe them according to the structure of their natural substrates. It is intriguing to know why so many sulphatases each with their own exquisite substrate specificity have evolved in order to degrade a single bond, i.e. the sulphate ester bond and also what is the basis for their substrate specificity.

It was shown that the N-terminal third of the IDS amino acid sequence showed strong sequence homology to the other known sulphatases. However, the homology in the remainder of the sequence was generally not as high as this, although it was still significant. This was particularly true when conservative amino acid substitutions were taken into account the homology was improved further (Section 4.3.6). The homology displayed between IDS and the other sulphatases, including an arylsulphatase from sea urchin (Sasaki *et al.*, 1988), suggests that they are a family of evolutionarily related proteins that have evolved from a primordial sulphatase by gene duplication and divergent evolution. The addition of the IDS and G6S sequences to the homology diagram (Figure 4.5) of the arylsulphatases has reduced the number of universally conserved residues seen among all the sulphatases. As IDS shows a greater degree of homology to the arylsulphatases than G6S it can be speculated that IDS is perhaps closer to the arylsulphatases in evolutionary terms (Figure 4.5).

It is possible that the primordial sulphatase had a broad specificity allowing it to degrade multiple subtrates with varying degrees of efficiency. As the complexity of

organisms increased a single sulphatase may not have being sufficient to efficiently degrade the newly evolved complex substrates containing sulphate-ester bonds. Another possibility could be that the ancestoral sulphatase was a single enzyme with a single specific substrate and was required to diverge into many different sulphatases to cope with the increasing variety of substrates brought about by the evolutionary increase in organisational complexity.

The reduction of the number of rigidly conserved residues among all the sulphatases has simplified the selection of residues that may be involved in specific functions. It is possible to predict that the residue involved in the sulphate ester bond hydrolysis will be conserved across all the sulphatases and must be capable of nucleophilic attack. Some of these residues may be targetted for site-directed mutagenesis and the kinetics of the mutant IDS could be studied to determine what role the specific residue has in IDS function, e.g does it function as a catalytic residue. In addition, regions important for defining substrate specificity may be identified by comparing IDS a non-arylsulphatase to an arylsulphatase such as ASB. It may be predicted that particular regions responsible for arylsulphatase specificity should be strongly conserved among the arylsulphatases but not in the non-arylsulphatases so these regions could also be targetted for mutagenesis.

For convenience (see Brooks *et al.*, 1991) these studies are reliant on the production of IDS specific monoclonal antibodies to measure the various mutant IDS properties, i.e. K_m, V_{max} , substrate binding studies and the analysis of the effect of substrate structure variants. The production of IDS monoclonal antibodies has been a goal of our group for many years and has met with little success. However, now that expressed IDS is available in large amounts the probability of producing monoclonal antibodies will be greatly improved (Section 4.3.7; Bielicki *et al.*, 1992). The production of large

amounts of IDS may also allow crystals to be grown for X-ray crystallography to allow the tertiary structure of IDS to be determined. Information about the tertiary structure of IDS could be used in conjunction with sequence homology and the genotype to phenotype relationship for MPS II patients may provide detail about the regions responsible for sulphate ester hydrolysis and the binding of specific IDS substrates.

The isolation of IDS cDNA enabled, for the first time, the characterisation of the nature of MPS II mutations at the Southern analysis and nucleotide level. Southern analysis of MPS II patient DNA revealed a wide variety of structural alterations to the IDS gene, which was consistent with what would be predicted for an X-linked disease (Chapter 5). All MPS II patients with a change at the Southern analysis level had severe clinical phenotypes. Two MPS II patients revealed a complete deletion of IDS, both of these patients were reported to have extremely severe clinical phenotypes (Section 5.3.1; Figure 5.1; Wilson *et al.*, 1990; Wraith *et al.*, 1991). This suggested the possiblity that other gene sequences may contribute to be the extreme clinical phenotype seen in these MPS II patients. If the extent of the deletions could be mapped it may be possible to identify the location of these other genes and to correlate the clinical severity with the extent of the deletion. The IDS cDNA probe was distributed to other investigators to begin a collection of MPS II data and this collaboration so far has revealed that 18% of MPS II patients surveyed revealed IDS gene alterations (Section 5.3.1 and Table 5.1). Before the IDS probe was distributed the percentage observed in work reported in this thesis suggested that this percentage was much higher. This was probably because the MPS II patients that were initially selected for analysis tended towards the severe end of the clinical spectrum. The isolation of IDS cDNA has also allowed the characterisation of some of the molecular

mutations that lead to MPS II (Section 5.3.4). All of the MPS II patients that were characterised at the nucleotide level showed unique mutations which included, point mutations, deletions and an insertion. This again is consistent with what would be expected for an X-linked disease, i.e. 1/3 of the mutations will be new (Haldane, 1935), and the remaining mutations will be unique or rare. Compared to autosomal linked diseases this makes the correlation of clinical phenotype with the type of mutation complicated. Obviously, the number of MPS II patients screened in this study is low but with distribution of the IDS probe to other investigators the number of MPS II patients characterised at the DNA level will increase. An obvious question to ask before the feasibility of MPS II genotype/phenotype correlation is possible is, do the same MPS II mutations occurring in different patients result in similar clinical and biochemical phenotypes? Until this question is answered correlation of genotype with clinical phenotype will remain unpredictable. However, studies from this thesis have revealed that some mutations result in either mild or severe clinical phenotypes, it would therefore be helpful if these same mutations could be found in other MPS II patients and whether these mutations were shown to result in similar MPS II clinical and biochemical phenotypes.

As was indicated in Section 5.3.4, the method employed in this study was limited and not all the MPS II patients screened showed a cleavage pattern consistent with a genetic change. It will be necessary to continue developing techniques that ensure 100% mutation detection. After these studies were completed it was reported that by using a labelled normal control and labelled patient DNA heteroduplex the possibility existed for the detection of all mutations (Forrest et al., 1991). The work described in this thesis provides a routinue procedure for the detection of the majority of mutations in MPS II particularly if the above modification is incorporated into the procedure.

In addition the IDS gene structure needs to be determined to aid in mutational analysis and screening. A limitation of this work is that all the mutations described have been determined from patient cDNA. With the availability of the intron/exon structure of IDS it would allow the amplification of exons containing mutations which could then be used as an added confirmation of the mutation. Also, knowledge of the intron/exon structure could be used to confirm the nature of one of the mutations described in Section 5.3.4, the patient with what appeared to be a 60 bp deletion had surrounding sequence that was suggestive of consensus mRNA splicing sites.

The work described in this thesis also opens up the possibilities for the development of patient therapy strategies. The potential of large quantities of expressed IDS will enable the evaluation of enzyme-replacement therapy in MPS II patients to begin. Expressed IDS was shown to correct storage in MPS II fibroblasts after M6P dependent endocytosis and the recombinant enzyme was shown to be localized to the correct sub-cellular fraction (Bielicki *et al.*, 1992). As mental retardation is a significant symptom in the more severely affected MPS II patients an essential question in these earlier studies will be to determine if the precursor form of IDS, which carries M6P residues, can cross the blood-brain barrier. At the very least enzyme-replacement therapy may alleviate the somatic symptoms of MPS II. A disadvantage with the enzyme-replacement therapy and gene replacement therapy protocols in MPS II is the lack of a suitable model system. If an animal model can be developed through homologous recombination of embryonic stem cells then the various treatment regimes may be tested. The information gained from the study of other MPS with animal models will be beneficial when eventually applied to MPS II. In addition, after a suitable number of MPS II patients have been screened for their specific mutations it may be possible to identify patients that will respond more

favourably to treatment and therefore have an improved prognosis. These extensions to the work pioneered in this thesis promise the potential for the eventual treatment of MPS II as well as providing an understanding of the structure/function relationship of IDS and its substrates and its role within the lysosome.

APPENDIX

PUBLICATIONS FROM THESIS WORK

Wilson, P.J., Morris, C.P., Anson, D.S., Occhiodoro, T., Bielicki, J., Clements, P.R. and Hopwood, J.J. (1990). Hunter syndrome: isolation of an iduronate-2-sulphatase cDNA clone and analysis of patient DNA. Proc. Natl. Acad. Sci. USA. **87**: 8531-8535.

Suthers, G.K., Hyland, V.J., Callen, D.F., Oberle, I., Rocchi, M., Thomas, N.S., Morris, C.P., Shwartz, C.E., Schmidt, M., Ropers, H.H., Baker, E., Oostra, B.A., Dahl, N., Wilson, P.J., Hopwood, J.J. and Sutherland, G.R. (1990). Physical mapping of new DNA probes near the fragile X mutation (FRAXA) by using a panel of cell lines. Am. J. Hum. Genet. **47**: 187-195.

Wilson, P.J., Suthers, G.K., Callen, D.F., Baker, E., Nelson, P.V., Cooper, A., Wraith, J.E., Sutherland, G.R., Morris, C.P. and Hopwood, J.J. (1991). Frequent deletions at Xq28 indicate genetic heterogeneity in Hunter syndrome. Hum. Genet. **86**: 505-508.

Wraith, J.E., Cooper, A, Thornley, M., Wilson, P.J., Nelson, P.V., Morris, C.P. and Hopwood, J.J. (1991). The clinical phenotype of two patients with a complete deletion of the iduronate-2-sulphatase gene (mucopolysaccharidoses II; Hunter syndrome). Hum. Genet. **87**: 205-206.

Suthers, G.K., Oberle, I., Nancarrow, J., Mulley, J.C., Hyland, V.J., Wilson, P.J., McCure, J., Morris, C.P., Hopwood, J.J., Mandel, J.L. and Sutherland, G.R. (1991). Genetic mapping of new RFLP's at Xq27-q28. Genomics. **9**: 37-43.

Wilson, P.J. *et al.* (1992). Analysis of two Hunter syndrome patients reveal frame shift mutations. Am. J. Hum. Genet. in preparation.

Clarke, J.T.R., Wilson, P.J., Morris, C.P., Hopwood, J.J., Richards, R.I., Sutherland, G.R., Mulley, J.C., Yu, S. and Ray, P.N. (1992). Partial characterization of a mutationinvolving IDS (Xq28) and associated with unbalanced inactivation of the non-mutant X chromosome. Am. J. Hum. Genet. in press.

Steen-Bondeson, M-L., Dahl, N., Tonnesen, T., Klejer, W.J., Seldlitz, G., Gustavson, K-H., Wilson, P.J., Morris, C.P., Hopwood, J.J. and Pettersson, U. (1992). Molecular analysis of patients with Hunter syndrome: implication of a region prone to structural alterations within the IDS gene. Human Medical Genetics. in press.

Faust, C.J., Verkerk, A.J.M.H., Wilson, P.J., Morris, C.P., Hopwood, J.J., Oostra, B. and Herman, G.E. (1992). Genetic mapping on the mouse X chromosome of Human cDNA clones for the Fragile X and Hunter syndromes. Genomics. **12**: 814-817.

Bielicki, J., Wilson, P.J., Hopwood, J.J. and Anson, D.S. (1992). Recombinant human iduronate-2-sulphatase: correction of MPS II fibroblasts and characterisation of purified enzyme. Biochem J. submitted for publication.

BIBLIOGRAPHY

Anderson, W.F. (1984). Prospects of human gene therapy. Science. **226**: 401-409.

Anson, D.S., Taylor, J.A., Bielicki, J., Harper, G.S., Peters, C., Gibson, G. and Hopwood, J.J. (1992). Correction of human mucopolysaccharidoses type VI fibroblasts with recombinant N-acetylgalactosamine-4-sulphatase. Biochem. J. 284: in press.

Anson, D.S., Bielicki, J. and Hopwood, J.J. (1992). Correction of mucopolysaccharidosis type I by retroviral-mediated transfer of the human α-L-iduronidase gene. Hum. Gene Therapy. in press.

Archer, I.M., Harper, P.S. and Wusteman, F.S. (1982). Multiple forms of iduronate 2-sulphate sulphatase in human tissues and bodily fluids. Biochimica et Biophysica Acta. **708**: 134-140.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K. (1989). Current Protocols in Molecular Biology. Wiley-Interscience, New York.

Bach, G., Eisenberg, F., Cantz, M. and Neufeld, E.F. (1973). The defect in the Hunter syndrome: deficiency of sulphoiduronate sulphatase. Proc. Natl. Acad. Sci. USA. **70**: 2134-2138.

Ballanger, C.E., Swift, T.R., Leshner, R.T., El Gammal, T.A. and McDonald, T.F. (1979). Myelopathy in mucopolysaccharidosis type II. Ann. Neurol. **7**: 382-385.

Baudhium, P., Hers, H.G. and Loeb, H. (1964). An electron microscopic and biochemical diagnosis of type II glucogenesis. Lab. Invest. **13**: 1139-1143.

Beck, M. (1983). Papillodema in association with Hunter's syndrome. Br. J. Ophthalmol. **67**: 174-177.

Beck, M. and Cole, G. (1984). Disc oedema in association with Hunter syndrome: ocular histopathological findings. Br. J. Ophthalmol. **68**: 590-594.

Benoist, C. and Chambon, P. (1981). *In vivo* sequence requirements of the SV40 early promoter region. Nature. **290**: 304-310.

Bielicki, J., Freeman, C., Clements, P.R. and Hopwood, J.J. (1990). Human liver iduronate-2-sulphatase: purification, characterisation and catalytic properties. Biochem. J. **271**: 75-86.

Bielicki, J., Wilson, P.J., Hopwood, J.J. and Anson, D.S. (1992). Recombinant human iduronate-2-sulphatase: correction of MPS II fibroblasts and characterisation of purified enzyme. Biochem J. submitted for publication.

Brante, G. (1952). Gargolylism: A mucopolysaccharidosis. Scand. J. Clin. Lab. Invest. **4**: 43-46.

Braulke, T., Hille, A., Huttner, W.B., Hasilik, A. and von Figura, K. (1987). Sulphated oligosaccharides in human lysosomal enzymes. Biochem. Biophys. Res. Commun. **143**: 178-185.

Breider, M.A., Shull, R.M. and Constantopoulos, G. (1989). Long-term effects of bone marrow transplantation in dogs with mucopolysaccharidosis I. Am. J. Path. **134**: 677-692.

Broadhead, D.M., Kirk, J.M., Burt, A.J., Gupta, V., Ellis, P.M. and Besley, G.T.N. (1986). Full expression of Hunter's disease in a female with an X-chromosome deletion leading to non-random inactivation. Clin. Genet. **30**: 392-398.

Brooks, D.A., McCourt, P.A.G., Gibson, G.J., Ashton, L.J., Shutter, M. and Hopwood, J.J (1991). Analysis of N-acetylgalactosamine-4-sulphatase protein and kinetics in mucopolysaccharidosis type VI patients. Am. J. Hum. Genet. **48**: 710-719.

Bucher, P. (1990). Weight matrix descriptions of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequences. J. Mol. Biol. **212**: 563-578.

Bucher, P. and Trifinov, E.N. (1986). Compilation and analysis of eukaryotic POL II promoter sequences. Nucl. Acids Res. **14**: 10009-10026.

Bucholtz, C.A. and Reisner, A.H. (1986). MBIS-an integrated system for the retrieval and analyses of sequence data from nucleic acids and proteins. Nucl. Acids Res. **14**: 265-272.

Cantz, M., Chrambach, A., Bach, G. and Neufeld, E. (1972). The Hunter corrective factor: purificaton and preliminary characterisation. J. Biol. Chem. **247**: 5456-5462.

Capecchi, M.R. (1989). Altering the genome by homologous recombination. Science. **244**: 1288-1292.

Capobianchi, M.R. and Romeo, G. (1976). Mosaicism for sulphoiduronate sulphatase deficiency in carriers of Hunter's syndrome. Experimentia. **32**: 459-460.

Caruso, R.C., Kaiser-Kupfer, M.I., Muenzer, J., Ludwig, I.H., Zasloff, M.A. and Mercer, P.A. (1986). Electroretinographic findings in the mucopolysaccharidoses. Ophthalmology. **93**: 1612-1616.

Chakravarti, A. and Bale, S.J. (1983). Differences in the frequency of X-linked deleterious genes in human populations. Am. J. Hum. Genet. **35**: 1252-1257.

Chomczynski, P. and Sacchi, N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. **162**: 156-159.

Clarke, J.T.R., Willard, H.F., Teshima, I., Chang, P.L. and Skomorowski, M.A. (1990). Hunter disease in a karotypically normal girl. Clin. Genet. **37**: 355-362.

Clements, P.R., Brooks, D.A., Saccone, G.T.P. and Hopwood, J.J. (1985). Human α-L-iduronidase. purification, monoclonal antibody production, native and subunit molecular mass. Eur. J. Biochem. **152**: 21-28.

Coppa, G.C., Singh, N.J. and Di Ferrante, N. (1973). Urinary excretion of disulphated disaccharides in Hunter syndrome: correction by infusion of a serum fraction. Anal. Lett. **6**: 225-233.

Cotton, R.G.H. (1989). Detection of single base changes in nucleic acids. Biochem. J. **263**: 1-10.

Cotton, R.G.H. and Campbell, R.D. (1989). Chemical reactivity of matched cytosine and thymine bases near mismatched and unmatched bases in a heteroduplex between DNA strands with multiple differences. Nucl. Acids Res. **17**: 4223-4233.

Cotton, R.G.H., Rodrigues, N.R. and Campbell, R.D. (1988). Reactivity of cytosine and thymine in single-base-pair mismatches with hydroxylamine and osmium tetroxide and its application to the study of mutations. Proc. Natl. Acad. Sci. USA. **85**: 4397-4401.

Cullen, B.R. (1988). Expression of a cloned human interleukin-2 cDNA Is enhanced by the substitution of a heterologous mRNA leader region. DNA. **7**: 645-650.

Dahl, H.H., Lamande, S.R., Cotton, R.G.H. and Bateman, J.F. (1989). Detection and localization of base changes in RNA using a chemical cleavage method. Anal. Biochem. **183**: 263-268.

Dahl, H.HM., Maragos, C., Brown, R.M., Hansen, L.L. and Brown, G.K. (1990). Pyruvate dehydrogenase deficiency caused by deletion of a 7-bp repeat sequence in the E1a gene. Am. J. Hum. Genet. **47**: 286-293.

Dahl, N., Hammarstrom-Heeroma, K., Van Ommen, G.B. and Pettersson, U. (1989). A polymorphic locus at Xq27-28 detected by the probe U6.2 (DXS304). Nucl. Acids Res. **17**: 2884.

Danes, B.S. and Bearn, A.G. (1967). Hurler's syndrome: A genetic study of clones in cell culture with particular reference to the Lyon hypothesis. J. Exp. Med. **126**: 509-522.

Daniele, A. and Di Natale, P. (1987). Hunter syndrome: presence of material cross-reacting with antibodies against iduronate sulphatase. Hum. Genet. **75**: 234-238.

Di Natale, P. and Ronsisvalle, L. (1981). Identification and partial characterisation of two enzyme forms of iduronate sulfatase from Human placenta. Biochim. Biophys. Acta. **661**: 106-111.

Di Natale, P. and Daniele, A. (1985). Iduronate sulphatase from human placenta. Biochim. et Biophys. Acta. **839**: 258-261.

Dianzani, I., Forrest, S.M., Camaschella, C., Saglio, G., Ponzone, A. and Cotton, R.G.H. (1991). Screening for mutations in the phenylalanine hydroxylase gene in Italian patients with phenylketonuria by using the chemical cleavage method: a new splice mutation. Am. J. Hum. Genet. **48**: 631-635.

DiLella, A.G., Marvit, J., Brayton, K. and Woo, S.L. (1987). An amino acid substitution involved in phenylketonuria is in linkage disequilibrium with DNA haplotype 2. Nature. **327**: 333-336.

DiLella, A.G., Marvit, J., Lidsky, A.S., Guttler, F. and Woo, S.L. (1986). Tight linkage between a splicing mutation and a specific DNA haplotype in phenylketonuria. Nature. **322**: 799-803.

Dorfman, A. and Lorincz, A.E. (1957). Occurence of urinary acid mucopoysaccharides in the Hurler syndrome. Proc. Natl. Acad. Sci. USA. **43**: 443-446.

Dynan, W.S. (1986). Promoters for housekeeping genes. Trends Genet. **2**: 196-197.

Eckner, R., Ellmeier, W. and Birnsteil, M.L. (1991). Mature mRNA 3' end formation stimulates RNA export from the nucleus. EMBO. **10**: 3513-3522.

Eperon, I.C. (1986). Rapid preparation of bacteriophage DNA for sequence analysis in sets of 96 clones, using filtration. Anal. Biochem. **156**: 406-412.

Erickson, A.H., Walter, P. and Blobel, G. (1983). Translocation of a lysosomal enzyme across the microsomal membrane requires signal recognition particle Biochem. Biophys. Res. Commun. **115**: 275-280.

Fluharty, A.L., Yutaka, T., Stevens, R.L. and Kihara, H. (1977). Hair root analysis of sulfoiduronate sulfatase in a Hunter syndrome heterozygote. Am. J. Hum. Genet. **29**: 43A.

Forrest, S.M., Dahl, H.H., Howells, D.W., Dianzani, I. and Cotton, R.G.H. (1991). Mutation detection in phenylketonuria by using chemical cleavage of mismatch: importance of using probes from both normal and patient samples. Am. J. Hum. Genet. **49**: 175-183.

Fransson, L-A. (1989). Heparan sulphate proteoglycans: Structure and properties. In Heparin: Chemical and Biological Properties, Clinical Applications (Lane, D.W. and Lindahl, U., eds.), Edward Arnold, London, pp. 115-133.

Fratantoni, J.C., Hall, C.W. and Neufeld, E.F. (1968). Hurler and Hunter Syndromes: mutual correction of the defect in cultured fibroblasts. Science. **162**: 570-572.

Fratantoni, J.C., Hall, C.W. and Neufeld, E.F. (1968). The defect in Hurler's and Hunter's syndromes: faulty degradation of mucopolysaccharide. Proc. Natl. Acad. Sci. USA. **60**: 699-706.

Fratantoni, J.C., Hall, C.W. and Neufeld, E.F. (1969). The Defect in Hurler and Hunter syndromes, II. deficiency of specific factors involved in mucopolysaccharide degradation. Proc. Natl. Acad. Sci. USA. **64**: 360-366.

Freeman, C., Clements, P.R. and Hopwood, J.J. (1983). Acetyl-CoA: alpha-glucosaminide N-acetyltransferase: partial purification from human liver. Biochem. Internat. **6**: 663-671.

Freeman, C. and Hopwood, J.J. (1991). Glucuronate-2-sulphatase activity in cultured human skin fibroblast homogenates. Biochem. J. **279**: 399-405.

Freeman, C. and Hopwood, J.J (1992). Lysosomal degradation of heparin and heparan sulphate In: 'Heparin and related polysaccharides' (Lane, D. and Lindahl, U. eds.). Plenum press, U.K., in press.

Friedman, T. and Robin, R. (1972). Gene therapy for human genetic disease. Science. **175**: 949-955.

Friedman, T. (1989). Progress toward human gene therapy. Science. **244**: 1275-1281.

Friedmann, T. and Brown, D.M. (1978). Base-specific reactions useful for DNA sequencing: methylene blue-sensitized photooxidation of guanine and osmium tetroxide modification of thymine. Nucl. Acids Res. **5**: 615-622.

Gasper, P.W., Thrall, M.A., Wenger, D.A., Macy, D.W., Ham, L., Dornsife, R.E., McBiles, K., Quackenbush, S.L., Kesel, M.L., Gillette, E.L. and Hoover, E.A. (1984). Correction of feline arylsuphate B deficiency (mucopolysaccharidosis VI) by bone marrow transplantation. Nature. **312**: 467-469.

Gill, J.P., Hobson, R., Hanley, W.B. and McKusick, V.A. (1965). Electroretinography and fundus occuli findings in Hurlers disease and allied mucopolysaccharidoses. Arch. Ophthal. **74**: 596-603.

Ginzberg, A.S., Onal, E., Aronson, R.M., Schild, J.A., Mafee, M.F. and Lopata, M. (1990). Successful use of nasal CPAP for obstructive sleep apnea in Hunter Syndrome with diffuse airway involvement. Chest. **97**: 1496-1498.

Goldberg, M.F. and Duke, J.R. (1967). Ocular histopathology in Hunter syndrome. Arch. Ophthal. **77**: 503-512.

Grompe, M., Muzny, D.M., Caskey, C.T. (1989). Scanning detection of mutations in human ornithine transcarbamylase by chemical mismatch cleavage. Proc. Natl. Acad. Sci. USA. **86**: 5888-5892.

Haldane, J.B.S. (1935). The rate of spontaneous mutation of a human gene. J. Genet. **31**: 317-326.

Hall, C.W., Liebaers, I., Di Natale, P. and Neufeld, E.F. (1978). Enzymic diagnosis of the genetic mucopolysaccharide storage diseases. Meth. Enzymol. **50**: 439-456.

Harrison, J.R. (1982). Detection of heterozygotes for X-linked enzymopathies by individual hair root analysis. Thesis for Fellowship of the Australian Institute of medical Laboratory Scientists.

Hasilik, A. and von Figura, K. (1981). Oligosaccharides in lysosomal enzymes: distribution of high-mannose and complex oligosaccharides in cathepsin D and b-hexosaminidase. Eur. J. Biochem. **121**: 125-129.

Hasilik, A., von Figura, K., Conzelmann, E., Nehrkorn, H. and Sandhoff, K. (1982). Lysosomal enzyme precursors in human fibroblasts. Activation of cathepsin D precursor *in vitro* and activity of β-hexosaminidase A precursor towards ganglioside GM2. Eur. J. Biochem. **125**: 317-321.

Hasilik, A. and von Figura, K. (1984). Lysosomes in biology and pathology, ed. J.T. Tingle, R.T. Dean, W.S. Sly, **7**: 3-16, Amsterdam, Elsevier.

Haskins, M.E., Desnick, R.J., DiFerrante, N., Jesyk, P.F. and Patterson, D.F. (1979). α-L-iduronidase deficiency in a cat: a model of mucopolysaccharidosis I. Ped. Res. **13**: 1294-1297.

Haskins, M.E., Desnick, R.J., DiFerrante, N., Jesyk, P.F. and Patterson, D.F. (1984). β-Glucuronidase deficiency in a dog: A model of human mucopolysaccharidosis VII. Ped. Res. **18**: 980-984.

Hers, G. (1965). Inborn lysosomal diseases. Gastroenterology. **48**: 625-633.

Hobbs, J.R., Hugh-Jones, K., Barret, A.J., Byrom, N., Chambers, D., Henry, K., James, D.C.O., Lucas, C.F., Rogers, T.R., Benson, P.F., Tansley, L.R., Patrick, A.D., Mossman, J. and Young, E.P. (1981). Reversal of clinical features of Hurler's disease and biochemical improvement after treatment by bone-marrow transplantation. Lancet. **ii**: 709-712.

Hopwood, J.J. (1979). a-L-Iduronidase, b-D-glucuronidase and 2-sulfo-L-iduronate 2-sulphatase: preparation and characterisation of radioactive substrates from heparin. Carbohydr. Res. **69**: 203-216.

Hopwood, J.J. and Harrison, J.R. (1985). Urinary excretion of sulphated N-acetylhexosamine in patients with various mucopolysaccharidoses. Biochem. J. 229: 579-586.

Hopwood, J.J. (1989). Enzymes that degrade heparin and heparan sulphate. In Heparin: Chemical and Biological Properties, Clinical Applications (Lane, D.W. and Lindahl, U., eds.), Edward Arnold, London, pp. 191-229.

Hopwood, J.J. and Morris, C.P. (1990). The mucopolysaccaridoses: diagnosis, molecular genetics and treatment. Mol. Biol. Med. **7**: 381-404.

Howells, D.W., Forrest, S.M., Dahl, H-H.M. and Cotton, R.G.H. (1990). Insertion of an extra codon for threonine is a cause of dihydropteridine reductase deficiency. Am. J. Hum. Genet. **47**: 279-285.

Hunter, C. (1917). A rare disease in two brothers. Proc. R. Soc. Med. **10**: 104-115.

Hurler, G. (1919). Uber eninen type multipler abartungen, vorwiegend am skelettsystem. Kinderheilkd. **24**: 220-234.

Huynh, T.V., Young, R.A. and Davis, R.W. (1985). Construction and screening cDNA libraries in lambda gt10 and lambda gt11. Cloning techniques: a practical approach (D. Glover), pp. 49-78, IRL Press, Oxford.

James, G.T. (1979). Essential arginine residues in human liver arylsulphatase A. Arch. Biochem. Biophys. **197**: 57-62.

Jezyk, P., Haskins, M.E., Patterson, D.F., Mellman, W.J. and Greenstein, M. (1977). Mucopolysaccharidosis in a cat with arylsulphatase B deficiency: a model of Maroteaux-Lamy syndrome. Science. **198**: 834-836.

Jobling, S.A. and Gehrke, L. (1987). Enhanced translation of chimaeric messenger RNA's containing a plant viral untranslated leader sequence. Nature. **325**: 622-625.

Kadonaga, J.T., Jones, K.A. and Tijian, R. (1986). Promoter-specific activation of RNA polymerase II transcription by Sp1. Trends Biochem. **11**: 20-23.

Klein, U., Kresse, H. and von Figura, K. (1978). Sanfilippo syndrome type C: deficiency of acetyl-CoA: α-glucosaminide N-acetyltransferase in skin fibroblasts. Proc. Natl. Acad. Sci. USA. **75**: 5185-5189.

Kidd, K.K., Bowcock, A.M., Schmidtke, J., Track, R.K., Ricciuti, F., Hutchings, G., Bale, A., *et al.* (1989). Report of the DNA committee and catalogs of cloned and mapped genes and DNA polymorphisms. Human Gene Mapping 10 (1989): Tenth international workshop on human gene mapping. Cytogenet. Cell Genet. **51**: 622-947.

Kornfeld, R and Kornfeld, S. (1985). Assembly of asparagine-linked oligosaccharides. Annu. Rev. Biochem. **54**: 631-664.

Kornfeld, S. and Mellman, I. (1989). The biogenesis of lysosomes. Annu. Rev. Cell Biol. **5**: 483-525.

Kozak, M. (1984). Point mutations close to the AUG initiator codon affect the efficency of translation of rat preproinsulin in vivo. Nature. **308**: 241-246.

Kozak, M. (1986a). Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell. **44**: 283-292.

Kozak, M. (1986b). Bifunctional messenger RNA's in eukaryotes. Cell. **47**: 481-483.

Kozak, M. (1987). A analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. Nucl. Acids Res. **15**: 8125-8131.

Kozak, M. (1991). Structural features in eukaryotic mRNAs that modulate the initiation of translation. J. Biol. Chem. **266**: 19867-19870.

Krchnak, V., Mach, O. and Maly, A. (1987). Computer prediction of potential immunogenic determinants from protein amino acid sequence. Anal. Biochem. **165**: 200-207.

Kresse, H., von Figura, K., Klein, U., Glossl, J., Paschke, E. and Pohlmann, R. (1982). Enzymic diagnosis of the genetic mucopolysaccharide storage disorder. Meth. Enzymol. **83**: 559-572.

Krivit, W. and Paul, N.W. (1986). Bone marrow transplantation for treatment of lysosomal storage disease. Birth Defects: original article series, Alan R. Liss, Inc., New York, vol. 22, no. 1.

Kyle, J.W., Birkenmeier, E.H., Gwynn, B., Vogler, C., Hoppe, P.C., Hoffmann, J.W. and Sly, W.S. (1990). Correction of murine mucopolysaccharidosis VII by a human β-glucuronidase transgene. Proc. Natl. Acad. Sci. USA. **87**: 3914-3918.

Lee, G.D. and van Etten, R.L. (1975). Evidence for an essential histidine residue in rabbit liver aryl sulphatase A. Arch. Biochem. Biophys. **171**: 424-434.

Lemansky, P., Hasilik, A., von Figura, K., Helmy, S., Fishman, J., Fine, R.E., Kedersha, N.L. and Rome, L.R. (1987). Lysosomal enzyme precursors in coated vesicles derived from the exocytic and endocytic pathways. J. Cell Biol. **104**: 1743-1748.

Liebaers, I. and Neufeld, E.F. (1976). Iduronate sulphatase in serum, lymphocytes and fibroblasts-simplified diagnosis of the Hunter syndrome. Pediatr. Res. **10**: 733-736.

Lindahl, U., Feingold, D.S. and Roden, L. (1986). Biosynthesis of heparin. Trends Biochem. Sci. **11**: 221-225.

Lipman, D.J., Altschul, S.F. and Kececioglu, J.D. (1989). A tool for multiple sequence alignment. Proc. Natl. Acad. Sci. USA. **86**: 4412-4415.

Lissens, W., Zenati, A. and Liebaers, I. (1984). Polyclonal antibodies against iduronate-2-sulphate sulphatase from human urine. Biochim. Biophys. Acta. **801**: 365-371.

Litjens, T. *et al.* (1992). Molecular basis of Marotreux-Lamy syndrome in a severely affected patient. in preparation.

Lodish, H.F. (1988). Transport of secretory and membrane glycoproteins from the rough endoplasmic reticulum to the Golgi. J. Biol. Chem. **263**: 2107-2110.

Loenen, W.A.M. and Blattner, F.R. (1983). λ Charon vectors (ch 32,33,34 and 35) adapted for DNA cloning in recombination deficient hosts. Gene. **26**: 171-179.

Lowry, R.B. and Renwick, D.H.G. (1971). Relative frequency of Hurler and Hunter syndrome. N. Engl. J. Med. **284**: 221-222.

Lyon, M. (1962). Sex chromatin and gene action in the mammalian X-chromosome. Am. J. Hum. Genet. **14**: 135-148.

Mandel, J.L., Willard, H.F., Nussbaum, R.L., Romeo, G., Puck, J.M. and Davies, K.E. (1989). Report of the committee on genetic constitution of the X chromosome. (10th International workshop on human gene mapping) Cytogenet. Cell Genet. **51**: 384-437.

Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982). Molecular cloning: a laboratory manual. Cold Springs Harbor Laboratory, New York.

Maniatis, T., Goodbourn, S. and Fischer, J.A. (1987). Regulation of inducible and tissue-specific gene expression. Science. **236**: 1237-1245.

Masuyama, T., Gojobori, T., Aota, S. and Ikemura, T. (1986). Codon usage tabulated from the Genbank genetic sequence data. Nucl. Acids Res. **14**: r151-r197.

McGovern, M.M., Mandell, N., Haskins, M. and Desnick, R.J. (1985). Animal model studies of allelism: characterization of arylsulphatase B mutations in homoallelic and heteroallelic (genetic compound) homozygotes with feline mucopolysaccharidosis VI. Genetics. **110**: 733-749.

McKusick, V.A. (1966). Heritable disorders of connective tissue. ed. 3, St. Louis, C.V. Mosby Co.

McKusick, V.A. (1972). Heritable disorders of connective tissue. St. Louis, C.V. Mosby Co., pp. 521-664.

McKusick, V.A. and Neufeld, E.F. (1983). The mucopolysacchridoses. In The metabolic basis of inherited disease, third ed. (J.B. Stanbury, J.B. Wyngaardan, D.S. Fredrickson, J.L. Golstein and M.S. Brown, eds.), McGraw-Hill, New York, pp. 751.

Merle, P. and Kadenbach, B. (1980). The subunit composition of mammalian cytochrome C oxidase. Eur. J. Biochem. **105**: 499-507.

Migeon, B.R., Sprenkle, J.A., Liebaers, I., Scott, J.F. and Neufeld, E.F. (1977). X-linked Hunter syndrome: the heterozygous phenotype in cell culture. Am. J. Hum. Genet. **29**: 448-454.

Mizushima, S. and Nagata, S. (1990). pEF-BOS, a powerful mammalian expression vector. Nucl. Acids Res. **18**: 5322.

Montandon, A.J., Green, P.M., Gianelli, F. and Bentley, D.R. (1989). Direct detection of point mutations by mismatch analysis: application to haemophilia B. Nucl. Acids Res. **17**: 3347-3358.

Morreau, H., Galjart, N.J., Gillemans, N., Willemsen, R., van Horst, T.J. and d'Azzo, A. (1989). Alternative splicing of β-galactosidase mRNA generates the classic lysosomal enzyme and a β-galactosidase-related protein. J. Biol. Chem. **264**: 20655-20663.

Mossman, J., Blunt, S., Stephens, R., Jones, E.E. and Pembrey, M. (1983). Hunters sisease in a girl: association with X:5 chromosomal translocation disrupting the Hunter gene. Arch. Dis. in Childhood. **58**: 911-915.

Murray, V. (1989). Improved double-stranded DNA sequencing using the linear polymerase chain reaction. Nucl. Acids Res. **17**: 8889.

Myers, R.M., Larin, Z. and Maniatus, T. (1985a). Detection of single base substitutions by ribonuclease cleavage at mismatches in RNA:DNA duplexes. Science. **230**: 1242-1246.

Myers, R.M., Lumelsky, N., Lerman, L.S. and Maniatus, T. (1985b). The detection of single base substitutions in total genomic DNA. Nature. **313**: 495-498.

Neufeld, E.F. (1991). Lysosomal storage disorders. Annu. Rev. Biochem. **60**: 257-280.

Neufeld, E.F. and Muenzer, J. (1989). in The Metabolic Basis of Inherited Disease, eds. Scriver, C.R., Beaudet, A.L.,Sly, W.S. and Valle, D. (McGraw-Hill, New York), pp. 1565-1587.

Neufeld, E.F., Liebaers, I., Epstein, C.J., Yatziv, S., Milunsky, A. and Migeon, B.R. (1977). The Hunter syndrome in females: Is there an autosomal recessive form of iduronate sulphatase deficiency. Am. J. Hum. Genet. **29**: 455-461.

Nja, A. (1946). Sex-linked type of gargoylism. Acta. Paediatr. **33**: 267-286.

Nwokoro, N., Liebaers, I. and Neufeld, E.F. (1977). Iduronate sulphatase activity in hair roots of Hunter heterozygotes. Am. J. Hum. Genet. **29**: 82A.

Nwokoro, N. and Neufeld, E.F. (1979). Detection of Hunter heterozygotes by enzymatic analysis of hair roots. Am. J. Hum. Genet. **31**: 42-49.

Oostra, B.A., Hupkes, P.E., Perdon, L.F., van Bennekom, C.A., Bakker, E., Halley, D.J.J., Schmidt, M. *et al.* (1990). New polymorphic DNA marker close to the fragile site FRAXA. Genomics. **6**: 129-132.

Oshima, A., Kyle, J.W., Miller, R.D., Hoffman, J.W., Powell, P., Grubb, J.H., Sly, W.S., Tropak, M., Guise, S. and Gravel, R.A. (1987). Cloning, sequencing, and expression of cDNA for human β-glucuronidase. Proc. Natl. Acad. Sci. USA. **84**: 685-689.

Palmieri, G., Capra, V., Romano, G., D'Urso, M., Johnson, S., Schlessinger, D., Morris, C.P., Hopwood, J.J., Di Natale, P., Gatti, R. and Ballabio, A. (1992). The iduronate sulphatase gene: isolation of a 1.2-Mb YAC contig spanning the entire gene and identification of heterogeneous deletions in patients with Hunter syndrome. Genomics. **12**: 52-57.

Peck, J.E. (1984). Hearing loss in Hunter's syndrome-mucopolysaccharidosis II. Ear Hear. **5**: 243-246.

Peters, C., Schmidt, B., Rommerskirch, W., Rupp, K., Zuhlsdorf, M., Vingron, M., Meyer, H.E., Pohlmann, R. and von Figura, K. (1990). Phylogenetic conservation of arylsulphatases: cDNA cloning and expression of human arylsulphatase B. J. Biol. Chem. **265**: 3374-3381.

Pilz, R.B., Steglich, C. and Scheffler, I.E. (1990). Molecular and genetic characterization of an ornithine decarboxylase deficient Chinese hamster cell line. J. Biol. Chem. **265**: 8880-8886.

Poole, B.P. (1982). Developmental role of hyaluronate. Conn. Tiss. Res. **10**: 93-100.

Pyeritz, R.E. (1983). Cardiovascular manifestations of heritable disorders of connective tissue, in Steinberg, A.G., Bearn, A.G., Motulsky, A.G. and Childs, B. (eds.): Progress in Medical Genetics. Philadelphia, Saunders, vol. **5**., pp. 191-302.

Rayleigh, E.A., Murray, N.E., Revel, H., Blumnthal, R.M., Westaway, D., Reith, A.D., Rigby, P.W.J., Elhai, J. and Hanahan, D. (1988). McrA and McrB restriction phenotypes of some *E. coli* strains and implications for gene cloning. Nucl. Acids Res. **16**: 1563-1575.

Reisner, A.H. and Bucholtz, C.A. (1986). The MTX package of computer programmes for the comparison of sequences of nucleotides and amino acid residues. Nucl. Acids Res. **14**: 233-238.

Roberts, S.H., Upadhyaya, M., Sarfarazi, M. and Harper, P.S. (1989). Further evidence localizing the gene for Hunter's syndrome to the distal region of the X chromosome long arm. J. Med. Genet. **26**: 309-313.

Robertson, D.R., Freeman, C., Nelson, P.V., Morris, C.P. and Hopwood, J.J. (1988). Human glucosamine-6-sulphatase cDNA reveals homology with steroid sulphatase. Biochem. Biophys. Res. Commun. **157**: 218- 224.

Robertson, D.A., Freeman, C., Morris, C.P. and Hopwood, J.J. (1992). A cDNA clone for human glucosamine-6-sulphatase reveals differences between aryl- and non-arylsulphatases. submitted for publication.

Roden, L. (1980). Structure and metabolism of connective tissue proteoglycans. In The Biochemistry of Glycoproteins and Proteoglycans (Lennarz, W.J. ed.), Plenum Press, New York, pp. 267-371.

Rubin, C.M. and Schmid, C.W. (1980). Pyrimidine-specific chemical reactions useful for DNA sequencing. Nucl. Acids Res. **8**: 4613-4619.

Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. and Erlich, H.A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science. **239**: 487-491.

Sanger, F., Nicklen, S. and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA. **74**: 5463-5467.

Sasaki, C.T., Ruiz, R., Gaito, Jr. R., Kirchner, J.A. and Seshi, B. (1987). Hunter's syndrome: a study in airway obstruction. Laryngoscope. **97**: 280-285.

Sasaki, H., Yamada, K., Akasaka, K., Kawasaki, H., Suzuki, K., Saito, A., Sato, M. and Shimada, H. (1988). cDNA cloning, nucleotide sequence and expression of the gene for arylsulphatase in the sea urchin (Hemicentrotus pulcherrimus) embryo. Eur. J. Biochem. **177**: 9-13.

Sawdogo, M. and Van Dyke, M.W. (1991). A rapid method for the purification of deprotected oligodeoxynucleotides. Nucl. Acids Res. **19**: 674.

Schaap, T. and Bach, G. (1980). Incidence of Mucopolysaccharidoses in Israel: Is Hunter disease a jewish disease?. Hum. Genet. **56**: 221-223.

Schmickel, R.D. (1986). Contiguous gene syndromes: a component of recognizable syndromes. J. Pediatr. **109**: 231-241.

Scott, H.S., Litjens, T., Hopwood, J.J. and Morris, C.P. (1992). A common mutation for mucopolysaccharidoses type I associated with a severe Hurler syndrome phenotype. Hum. Mut. in press.

Sheffield, V.C., Cox, D.R., Lerman, L.S., Myers, R.M. (1989). Attachment of a 40-base-pair G+C-rich (GC-clamp) to genomic DNA fragments by the polymerase chain reaction results in improved detection of single-base changes. Proc. Natl. Acad. Sci. USA. **86**: 232-236.

Shull, R.M., Munger, R.J., Spellacy, E., Hall, C.W., Constantopoulos, G. and Neufeld, E.F. (1982). Canine α-L-iduronidase deficiency, a model of mucopolysaccharidosis I. Am. J. Path. **109**: 244-248.

Simmers, R.N., Stupans, I. and Sutherland G.R. (1986). Localisation of the human haptoglobin genes distal to the fragile site at 16q22 using *in situ* hybridisation. Cytogenet. Cell Genet. **41**: 38-41.

Simmers, R.N., Smith, J., Shannon, M.F., Wong, G., Lopez, A.F., Baker, E., Sutherland, G.R. and Vadas, M.A. (1988). Localisation of the human G-CSF gene to the region of a breakpoint in the translocation typical of acute promyelocytic leukemia. Hum. Genet. **78**: 134-136.

Sjoberg, I., Fransson, L-A., Matalon, R. and Dorfman, A. (1973). Hunter's syndrome: a deficiency of L-iduronosulphate sulphatase. Biochem. Biophys. Res. Comm. **54**: 1125-1132.

Spellacy, E., Shull, R.M., Constantopoulos, G. and Neufeld, E.F. (1983). A canine model of human α-L-iduronidase deficiency. Proc. Natl. Acad. Sci. USA. **80**: 6091-6095.

Spranger, J., Cantz, M., Gehler, J., Liebaers, I. and Theiss, W. (1978). Mucopolysaccharidosis II with corneal opacities. Eur. J. Pediatr. **129**: 11-16.

Staden, R. (1980). A new computer method for the storage and manipulation of DNA gel reading data. Nucl. Acids Res. **8**: 3673-3694.

Steen-Bondeson, M-L., Dahl, N., Tonnesen, T., Klejer, W.J., Seldlitz, G., Gustavson, K-H., Wilson, P.J., Morris, C.P., Hopwood, J.J. and Pettersson, U. (1992). Molecular analysis of patients with Hunter syndrome: Implication of a region prone to structural alterations within the IDS gene. Human Medical Genetics. in press.

Stein, C., Gieselmann, V., Kreysing, J., Schmidt, B., Pohlmann, R., Waheed, A., Meyer, H.E., O'Brien, J.S. and von Figura, K. (1989a). Cloning and expression of human arylsulfatase A. J. Biol. Chem. **264**: 1252-1259.

Stein, C., Gieselmann, V., Kresying, J., Schmidt, B., Pohlmann, R., Waheed, A., Meyer, H.E., O'Brien, J.S. and von Figura, K. (1989b). Cloning and expression of human steroid-sulphatase: membrane topology, glycosylation, and subcellular distribution in BHK-21 cells. J. Biol. Chem. **264**: 1252-1259.

Stigson, M. and Kjellen, L. (1991). Large disulfide-stabilized proteoglycan commplexes are synthesized by the epidermis of Axolotl embryos. Arch. Biochem. Biophys. **290**: 391-396.

Suthers, G.K., Callen, D.F., Hyland, V.J., Kozman, H.M., Baker, E., Eyre, H., Harper, P.S., Roberts, S.H., Hors-Cayla, M.C., Davies, K.E., Bell, M.V. and Sutherland, G.R.(1989). A new DNA marker tightly linked to the fragile X locus. Science. **246**: 1298-1300.

Suthers, G.K., Hyland, V.J., Callen, D.F., Oberle, I., Rocchi, M., Thomas, N.S., Morris, C.P., Schwartz, C.E., Schmidt, M., Ropers, H.H., Baker, E., Oostra, B.A., Dahl, N., Wilson, P.J., Hopwood, J.J. and Sutherland, G.R. (1990). Physical mapping of new DNA probes near the fragile X mutation (FRAXA) by using a panel of cell lines. Am. J. Hum. Genet. **47**: 187-195.

Suthers, G.K., Oberle, I., Nancarrow, J., Mulley, J.C., Hyland, V.J., Wilson, P.J., McCure, J., Morris, C.P., Hopwood, J.J., Mandel, J.L. and Sutherland, G.R. (1991). Genetic mapping of new RFLP's at Xq27-q28. Genomics. **9**: 37-43.

Taylor, R.M., Stewart, G.J., Farrow, B.R., Byrne, J. and Healy, P.J. (1989). Histological improvement and enzyme replacement in the brains of fucosidosis dogs after bone marrow engraftment. Trans.Proc. **21**: 3074-3075.

Tomatsu, S., Fukada, S., Maue, M., Sukegawa, K., Fukao, T., Yamagishi, A., Hori, T., Iwata, H., Ogawa, T., Nakashima, Y., Hanyu, Y., Hashimoto, T., Titani, K., Oyama, R., Suzuki, M., Yagi, K., Hayashi, Y. and Orri, T. (1991). Morquio Disease: Isolation, characterization and expression of full-length cDNA for human N-acetylgalactosamine-6-sulphate sulphatase. Biochem. Biophys. Res. Comm. **181**: 677-683.

Topping, T.M., Kenyon, K.R., Goldberg, M.F. and Maumenee, A.E. (1971). Ultrastructural ocular pathology of Hunter's syndrome. Arch. Ophthal. **86**: 164-177.

Trifonov, E.N. (1987). Translation framing code and frame-monitoring mechanism as suggested by the analysis of mRNA and 16S rRNA nucleotide sequences. J. Mol. Biol. **194**: 643-652.

Tuddenham, E.G.D., Cooper, D.N., Gitschier, J., Higuchi, M., Hoyer, L.W., Yoshioka, A., Peake, I.R., Schwaab, R., Olek, K., Kazazian, H.H., Lavergne, J.-M., Giannelli, F. and Antonarakis, S.E. (1991). Haemophilia A: database of nucleotide substitutions, deletions, insertions and rearrangements of the factor VIII gene. Nucl. Acids Res. **19**: 4821-4833.

Upadhyaya, M., Sarfarazi, M., Bamforth, J.S., Thomas, N.S.T., Oberle, I., Young, I. and Harper, P.S. (1986). Localization of the gene for Hunter syndrome on the long arm of X chromosome. Hum.Genet. **74**: 391-398.

Van Hoof, F. and Hers, H.G. (1964). The ultrastructure of hepatic cells in Hurler's disease. C.R. Hebd Seances Acad. Sci. Ser. D. **259**: 1281-1283.

Vincent, A., Dahl, N., Oberle, I., Hanauer, A., Mandel, J.L., Malmgren, H. and Pettersson, U. (1989). The polymorphic marker DXS304 is within 5 centimorgans of the fragile X locus. Genomics. **5**: 797-801.

von Figura, K. and Hasilik, A. (1986). Lysosomal enzymes and their receptors. Annu. Rev. Biochem. **55**: 167-193.

von Figura, K., Hasilik, A., Pohlmann, R., Braulke, T., Lemansky, P. and Stein, M. (1987). Mutations affecting transport and stability of lysosomal enzymes. Enzyme. **38**: 144-153.

von Heijne, G. (1985). Signal sequences: the limits of variation. J. Mol. Biol. **184**: 99-105.

von Heijne, G. (1986). A new method for predicting signal sequence cleavage sites. Nucl. Acids Res. **14**: 4683-4690.

Wasteson, A. and Neufeld, E.F. (1982). Iduronate sulphatase from human plasma. Methods Enzymol. **83**: 573-578.

Wilson, P.J., Morris, C.P., Anson, D.S., Occhiodoro, T., Bielicki, J., Clements, P.R. and Hopwood, J.J. (1990). Hunter syndrome: isolation of an iduronate-2-sulphatase cDNA clone and analysis of patient DNA. Proc. Natl. Acad. Sci. USA. **87**: 8531-8535.

Wilson, P.J., Suthers, G.K., Callen, D.F., Baker, E., Nelson, P.V., Cooper, A., Wraith, J.E., Sutherland, G.R., Morris, C.P. and Hopwood, J.J. (1991). Frequent deletions at Xq28 indicate genetic heterogeneity in Hunter syndrome. Hum. Genet. **86**: 505-508.

Wilson, P.J. *et al.* (1992). Analysis of two Hunter syndrome patients reveal frame shift mutations. Am. J. Hum. Genet. in preparation.

Wolfe, J.H., Schuchman, E.H., Stramm, L.E., Concaugh, E.A., Haskins, M.E., Aguirre, G.D., Patterson, D.F., Desnick, R.J. and Gilboa, E. (1990). Restoration of normal lysosomal function in mucopolysaccharidosis type VII cells by retroviral vector-mediated gene transfer. Proc. Natl. Acad. Sci. USA. **87**: 2877-2881.

Wraith, J.E., Cooper, A, Thornley, M., Wilson, P.J., Nelson, P.V., Morris, C.P. and Hopwood, J.J. (1991). The clinical phenotype of two patients with a complete deletion of the iduronate-2-sulphatase gene (mucopolysaccharidoses II; Hunter syndrome). Hum. Genet. **87**: 205-206.

Yanish-Perron, C., Vieira, J. and Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene. **33**: 103-119.

Yatziv, S., Erickson, R.P. and Epstein, C.J. (1977). Mild and severe Hunter syndrome (MPS II) within the same sibships. Clin. Genet. **11**: 319-326.

Yen, P.H., Allen, E., Marsh, B., Mohandas, T., Wang, N., Taggart, R.T. and Shapiro, L.J. (1987). Cloning and expression of steroid sulphatase cDNA and the frequent occurrence of deletions in STS deficiency: implications for X-Y interchange. Cell. **49**: 443-454.

Young, I.D. and Harper, P.S. (1982). Incidence of Hunter's syndrome. Hum. Genet. **60**: 391-392.

Young, I.D. and Harper, P.S. (1983). The natural history of the severe form of Hunter syndrome: A study based on 52 cases. Dev. Med. Child. Neurol. **25**: 481-489.

Young, I.D., Harper, P.S., Archer, I.M. and Newcombe, R.G. (1982b). A clinical and genetic study of Hunter's syndrome. 1. Heterogeneity. J. Med. Genet. **19**: 401-407.

Young, I.D., Harper, P.S., Newcombe, R.G. and Archer, I.M. (1982a). A clinical and genetic study of Hunter's syndrome. 2. Differences between the mild and severe forms. J. Med. Genet. **19**: 408-411.

Young, R.A and Davis, R.W. (1983). Efficient isolation of genes by using antibody probes. Proc. Natl. Acad. Sci. USA. **80**: 1194-1198.

Yutaka, T., Fluharty, A.L., Stevens, R.L. and Kihara, H. (1982). Purification and some properties of human liver iduronate sulphatase. J. Biochem. (Tokyo). **91**: 433-441.

Zechner, G. and Moser, M. (1987). Otosclerosis and mucopolysaccharidosis. Acta. Otolaryngol. **103**: 384-386.

Zlotogora, J., Schaap, T., Zeigler, M. and Bach, G. (1985). Hunter syndrome among Ashkenazi Jews in Israel; evidence for prenatal selection favoring the Hunter allele. Hum. Genet. **71**: 329.

ADDENDUM

1. In chapter 5, section 5.4 (p99-104) the MPS II mutations detected by direct PCR sequencing were confirmed in both directions. In some cases, for example, patient SF 532, SF 811, SF 3055 the PCR sequencing was repeated more than once.

2. The molecular standards used to determine the size of restriction fragments in figure 5.1A-D (chapter 5, section 5.3.1, p93-95) were present in the ethidium bromide stained gel.

3. The ethidium bromide stained gel of figure 5.1B (chapter 5, section 5.3.1, p93-95) reveals that there was no DNA present in lane 5.

4. Figure 5.7 (chapter 5, section 5.4.1). This figure has lost resolution during the photographic reproduction process. Due to the limitations of the photographic resources this reproduction was the best possible. The original autoradiograph is quite clear and indicates that the bands exposed in the hydroxylamine lane are definitely present in the normal cleavage pattern.

5. Figure 5.9 (chapter 5, section 5.4.1, p101-102). The cleavage band in the osmium tetroxide lane is most likely to be secondary non-specific cleavage at a paired T adjacent to the mismatch. However, it is uncertain which of the flanking T bases surrounding the deleted C base is responsible.

6. Figure 5.12 (chapter 5, section 5.4.1, p102-103). This sequence was performed in both directions (primer 23 and 24). The sequence from primer 23 extended to nucleotide position 1360 well beyond the site of the deletion, leaving no doubt as to the nature of this mutation.

7. Figure 1.4 and 1.5 (chapter 1, section 1.4.3, 1.4.4) are from Hopwood and Morris, 1990.

8. Figure 1.2 (chapter 1, section 1.4.1) was modified from Neufeld and Meunzer, 1989 (p1563).

Typographical errors.

page 8, line 15-integrated for intergrated

page 14, line 14-exacerbating for excacerbating

page 80, line 6-produce for poduce

page 107, line 14-were for where.

page 113, line 3-ancestral for ancestoral.