# Functional Aspects of a Mutation in the *PLP2* Promoter Region of Patients with Non-Syndromic X-Linked Mental Retardation


| | |
|---|---|
| **Diploma Thesis** of | Achim Salamon |
| Matriculation Number | 200083 |
| Studying at the | Technische Universität Berlin<br>Straße des 17. Juni 135<br>10623 Berlin<br>Germany<br>Faculty III: Process Sciences,<br>Institute of Biotechnology,<br>Branch of Medical Biotechnology,<br>Prof. Dr. Roland Lauster |
| Thesis Realization at the | Max Planck Institute for Molecular Genetics<br>Ihnestraße 63 – 73<br>14195 Berlin<br>Germany<br>Department Human Molecular Genetics,<br>Prof. Dr. H.-Hilger Ropers<br>Mutation Screening Group,<br>Dr. Andreas W. Kuß |



Max Planck Institut
for Molecular Genetics

# Contents

## Zusammenfassung

Geistige Behinderung wird klinisch charakterisiert durch einen Intelligenzquotienten von weniger als 70, der in Verbindung mit Defiziten im adaptativen Verhalten der betroffenen Person auftritt und 1 – 3% der Bevölkerung betrifft. Die Ursachen von geistiger Behinderung können sowohl genetisch als auch umweltbedingt sein. Zu den äußeren Faktoren zählen z.B. Mangelernährung, Infektionen oder Kontakt mit Toxinen, während genetische Faktoren in chromosomalen Umordnungen oder in Mutationen einzelner Gene zu finden sind. In einem Drittel aller Fälle ist geistige Behinderung mit anderen Merkmalen gekoppelt (syndromische Formen), während zwei Drittel aller Patienten an nicht-syndromischen Formen leiden, in denen die geistige Behinderung das einzige Krankheitszeichen ist.

In dieser Arbeit wurde im Genom von Patienten mit nicht-syndromischer geistiger Behinderung nach ursächlichen Veränderungen gesucht. In 2,6% der Patienten wurde 188 bp oberhalb des Translationsstartes des *PLP2*-Gens ein Austausch von Cytosin durch Adenin gefunden. *PLP2* kodiert ein integrales Membranprotein des endoplasmatischen Retikulums, das noch nicht im Zusammenhang mit geistiger Behinderung gefunden wurde.

Da die gefundene Transversion den Promotorbereich von *PLP2* betrifft und Promotoren die Transkription von Genen regulieren, wurde die *PLP2*-Expression per Northern blot überprüft. In neun Individuen mit dem Genotyp $PLP2^{-188(A)}$ wurde gegenüber fünf Kontrollen mit $PLP2^{-188(C)}$ eine Reduktion der Expression um mehr als den Faktor 2,7 festgestellt.

Gleichzeitig wurde die Häufigkeit des $PLP2^{-188(A)}$-Genotyps überprüft. In männlichen Blutspendern wurde er in 2,6% aller Individuen gefunden. Da diese bezüglich des Intelligenzquotienten nicht charakterisiert waren, wurde die Promotorveränderung auch in Kontrollen mit mindestens durchschnittlichem Intelligenzquotienten gesucht – und in keiner Person gefunden. Das Auftreten der Veränderung in den Blutspendern mit der Häufigkeit eines seltenen Polymorphismus spricht jedoch dagegen, daß $PLP2^{-188(A)}$ unvermeidlich geistige Behinderung verursacht, sondern eher für eine Rolle als genetischer Modifier, der den Ausbruch der Krankheit begünstigt oder ihre Form verstärkt.

Um den Einfluß der Promotorveränderung auf die Transkription näher zu beleuchten, wurde in einem elektrophoretischen Mobilitätsshift Assay untersucht, ob die die Veränderung enthaltene Region von Transkriptionsfaktoren gebunden wird, welche so einen Einfluß auf die Geschwindigkeit des Zusammenbaus des Transkriptionsapparates und dessen Stabilität ausüben. *In silico* wurde dabei vorhergesagt, daß die Transkriptionsfatoren REL und ETS1 binden, aber nur zu $PLP2^{-188(C)}$. *In vitro* zeigte sich tatsächlich Proteinbindung, die aber bei beiden Varianten gleichstark war. Da sich ein sehr großer Protein-DNA-Komplex bildete, war die Analyse in diesem Assay nur mit Einschränkungen möglich.

Zusammenfassend kann man sagen, daß die Abwesenheit von $PLP2^{-188(A)}$ in Individuen mit mindestens durchschnittlichem Intelligenzquotienten für eine kognitive Rolle der Veränderung spricht, was dadurch erhärtet wird, daß sie in einer Transkriptionsfaktorbindungsstelle liegt. Weitere Untersuchungen sind nötig, um die Rolle der Veränderung in der Ätiologie von geistiger Behinderung aufzudecken.

## Summary

Mental retardation is defined by an overall intelligence quotient lower than 70 that is accompanied by functional deficits in adaptive behaviour. 1 – 3% of the population are affected by this disability. It can be caused by environmental and genetic factors. Environmental factors can be found in e.g. malnutrition or exposure to toxins, whereas genetic factors are represented by chromosomal rearrangements or single-gene mutations. In one third of the cases, mental retardation is associated with other clinical features and is therefore called syndromic, as opposed to non-syndromic forms where mental retardation is the only obvious sign.

When screening a panel of patients with non-syndromic mental retardation for causative mutations, in 2.6% of the patients, a cytosine to adenine exchange was found 188 bp upstream of the translation start of the gene *PLP2*. This $PLP2^{-188(A)}$ genotype is not represented in the RefSeq. *PLP2* codes for an integral membrane protein of the endoplasmic reticulum and was so far not implicated in mental retardation.

As the promoter region of a gene regulates its transcription, the cytosine to adenine change might affect gene expression. Thus, I investigated *PLP2* expression on the mRNA level by Northern blotting and found that it was more than 2.7 fold reduced in nine carriers of $PLP2^{-188(A)}$ as compared to five controls with $PLP2^{-188(C)}$.

In parallel, I investigated the occurence of the *PLP2* promoter change in control panels containing genetic material from blood donors that were not characterized for their intelligence quotient and from people that were selected to have at least an average intelligence quotient. In the male blood donors, the A-variant was found with a frequency of 2.6%, whereas it was completely absent in people with at least average intelligence quotient. However, as the variation of the *PLP2* promoter occurs with the frequency of a rare polymorphism, it cannot be an inevitable cause for mental retardation, but might still influence an individual's susceptibility for being affected by this disability. This is reflected in the concept of genetic modifiers.

Since regulation of transcription in eucaryotes mostly is achieved through modifying the assembly rate of the transcriptional pre-initiation complex by transcription factors binding to the promoter region of the regulated gene, I analyzed the consequences of the cytosine to adenine change of the promoter sequence by electrophoretic mobility shift assays. *In silico*, I found that binding of the transcription factors REL and ETS1 was not predicted for the A-variant in contrast to $PLP2^{-188(C)}$. *In vitro*, protein-binding was confirmed, but not found to be differential between both variants. However, this might have been due to limitations of the electrophoretic mobility shift assay, for the large protein-DNA complex formed could not be analyzed well on the polyacrylamide and agarose gels used.

In conclusion, the absence of $PLP2^{-188(A)}$ in people with at least average intelligence quotient suggests a cognitive implication of this polymorphism. This is substantiated by the finding that this promoter site contains a transcription factor binding site. However, as no difference concerning protein binding could be found between both variants, future research is needed to elucidate the role of $PLP2^{-188(A)}$ in the etiology of mental retardation.

# 1 Introduction

## 1.1 Mental Retardation

### 1.1.1 Definition of Mental Retardation

Mental retardation (MR) is a disability characterized by patients having an Intelligence Quotient (IQ) of less than 70 (see table 1) and showing behavioural and adaptational problems in daily life [20]. To exclude the normal decline of mental abilities in aging people, MR is defined to occur before the age of 18. With a prevalence of 1 – 3% in the population and extraordinary high lifetime costs in healthcare support (estimated to one to two million dollars per patient) [51], MR is one of the major socio-economic issues.

MR can affect people in differing severety, measured by means of the Intelligence Quotient (IQ). Depending on the IQ measured, mental retardation patients can be assigned to five subclasses (see table 1). Especially for borderline IQ values, it is difficult to differentiate between affected and non-affected individuals. This is why their ability to handle everyday life must also be considered. To simplify matters, MR classification is often reduced to distinguishing between mild ($50 < \text{IQ} < 70$) and severe ($\text{IQ} \leq 50$) forms.

Table 1: **Definition of classes of severety of mental retardation (after [95])**

| *Terminology* | *Intelligence quotient* |
|---|---|
| Profound | $< 20$ |
| Severe | 20 – 35 |
| Moderate | 35 – 50 |
| Mild | 50 – 70 |
| Borderline | 70 – 85 |

In some cases, MR is correlated with clinical symptoms or syndroms. Therefore, these types of MR are called syndromic, contrasting with the non-syndromic cases were no clinical features can be found despite of the mental disabilities of the patients. Due to increasing knowledge about disease phenotypes, many non-syndromic forms of MR in the past have been found to be actually syndromic.

Due to its diffuse character, MR cannot be considered as a clearly defined medical disease nor as a mental disorder. MR is a clinical state that reflects the loss of "fit" between the capabilities of individuals and the structure and expectations of their environment [3]. The disability can be caused by a heterogenous variety of factors of both environmental and genetic origin.

### 1.1.2 Causes of Mental Retardation

MR results from an extremely heterogeneous background of environmental, chromosomal and monogenic causes.

**1.1.2.1 Environmental causes of MR** include alcohol and drug abuse as well as malnutrition during pregnancy, perinatal complications, exposure to infecting agents and

toxins [95]. In the devoloping countries, injuries and deprivation of oxygen at birth and brain infections during childhood are more common causes of MR than in the industrialized countries.

Much is known about avoiding environmental causes for MR. For example, iodine-deficiency mental retardation (cretinism) can be prevented by iodization of salt, and mental retardation in people with phenylketonuria can be prevented by dietary control. In addition, social and educational interactions are important environmental factors that modulate a person's mental abilities, especially during childhood.

**1.1.2.2 Genetic causes of MR** can be subdivided into chromosomal and monogenic types [20]. Chromosomal aberrations include variants of aneuploidy, e.g. trisomy 21, and balanced chromosomal rearrangements, e.g. inversions and reciprocal translocations, abolishing a gene's function if the chromosomal breakpoint lies within that gene, as well as deletions and duplications of chromosomal segments.

Monogenic forms affect a single gene only and can be caused by small insertions or deletions as well as by nucleotide substitutions. Such changes affect mRNA expression, splicing or the amino acid sequence of the protein product.

### 1.1.3 X-Linked Mental Retardation

Elucidation of the causative gene defects for MR began exclusively with X-linked genes (X-Linked Mental Retardation, XLMR), as there is strong evidence for a major role of X-linked recessive traits in this disability: the ratio of affected males to females is roughly 1.4 for severe forms of MR and 1.9 for mild forms [70]. In the pedigrees of XLMR families, often only males are affected, heterozygous females acting as carriers and being unaffected or only mildly affected.

In addition to the pros for an involvement of the X chromosome, linkage analysis of the typically small pedigrees of affected families in the developed countries is not possible for autosomal genes. So, X-linked genes were a good starting point and so far revealed 61 genes for X-linked mental retardation (see figure 1), reviewed in [69]. These genes have been found to be implicated in diverse processes (see figure 2). This clearly shows the heterogeneity of the disorder, MR genes seemingly having only one property in common: to be expressed in the brain.

### 1.1.4 Finding Mental Retardation Genes

Due to technical limitations, it is currently not possible to sequence the entire genome of each MR patient to find the causative mutation. Therefore, one has to find it in another way, the method depending on the nature of the gene defect.

**1.1.4.1 Chromosomal Rearrangements** can be found initially by karyotyping. If a chromosomal rearrangement, such as a deletion or inversion or a balanced translocation

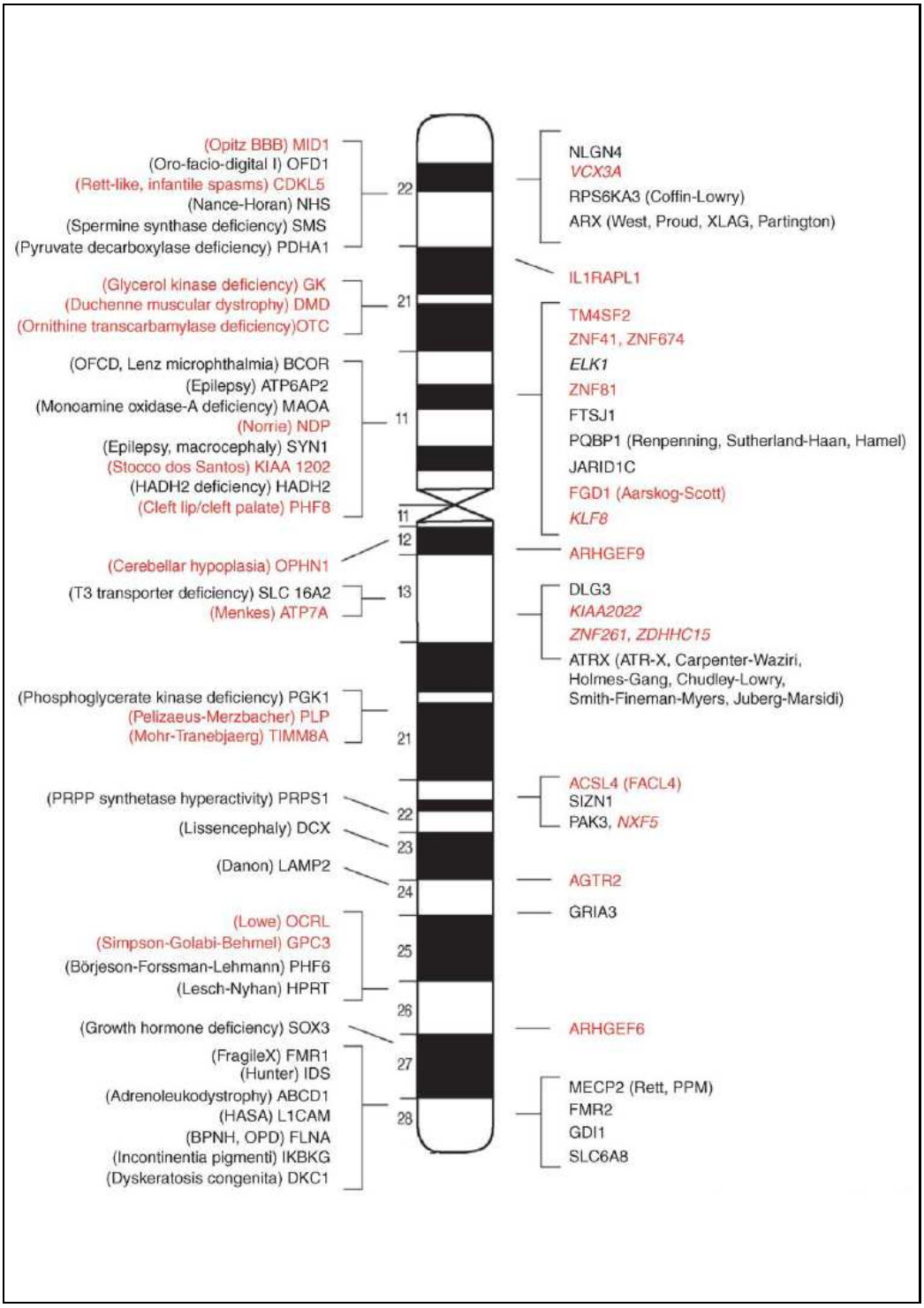



Figure 1: **Distribution of the 61 known XLMR genes on the X chromosome [69].** All XLMR genes that were identified by mutation screening are shown in black, whereas those that resulted from studying patients with chromosomal rearrangements are marked in red. Candidate genes are denoted in italics. Genes implicated in syndromic XLMR are shown on the left, and those implicated in non-syndromic XLMR are shown on the right.

| Gene | Functions | Year[b] |
|---|---|---|
| *FMR2* | Transcriptional regulator; possible role in long-term memory and enhanced long-term potentiation | 1996 [11] |
| *GDI1* | Regulation of synaptosomal Rab4 and Rab5 pools; possible role in endocytosis | 1998 [69] |
| *PAK3* | Regulation of actin cytoskeleton; stimulation of neurite outgrowth | 1998 [70] |
| *IL1RAPL* | Regulator of dense-core granule exocytosis; possible modulator of neurotransmitter release | 1999 [71] |
| *RPS6KA3* (*RSK2*)[a] | Serine–threonine protein kinase; CREB phosphorylation; role in long-term memory formation | 1999 [72] |
| *MECP2*[a] | Transcriptional silencer of neuronal genes, role in splicing | 1999 [73] |
| *ARHGEF6* | Integrin-mediated activation of Rac–cdc42; stimulation of neurite outgrowth | 2000 [24] |
| *TM4SF2* | Modulation of integrin-mediated signalling; neurite outgrowth; possible role in synapse formation | 2000 [74] |
| *SLC6A8*[a] | Creatine transporter; required for maintenance of (phospho)creatine pools in the brain | 2001 [75] |
| *ARX*[a] | Transcription factor with possible role in the maintenance of specific neuronal subtypes in the cerebral cortex and axonal guidance in the floorplate. Neuronal proliferation, differentiation of GABA-ergic neurons | 2002 [76,77] |
| *ATRX*[a] | DNA-binding helicase, involved in chromatin remodelling, DNA methylation and regulation of gene expression; intrinsic regulator of cortical size | 2002 [78] |
| *FGD1*[a] | RhoGEF; possible role in stimulation of neurite outgrowth | 2002 [79] |
| *ACSL4* (*FACL4*) | Long-chain fatty acid synthase; possible role in membrane synthesis and/or recycling | 2002 [80] |
| *AGTR2* | Brain-expressed angiotensin receptor 2 | 2002 [81] |
| *PQBP1*[a] | Polyglutamine-binding; putative role in transcription and mRNA splicing | 2003 [39•] |
| *ZNF41* | KRAB domain-containing zinc finger protein; putative transcriptional regulator; possible involvement in chromatin remodeling | 2003 [82] |
| *NLGN4*[a] | Postsynaptic membrane protein; involved in induction of presynaptic structures; linked to NMDA-type glutamatergic receptors | 2004 [83•] |
| *FTSJ1* | RNA methyltransferase, possible role in tRNA modification and translation | 2004 [84•] |
| *DLG3* | Post-synaptic scaffolding protein linked to NMDA-type glutamatergic receptors | 2004 [85•] |
| *JARID1C*[a] (*SMCX*) | Role in chromatin remodelling | 2004 [47•] |
| *ZNF81* | KRAB domain-containing zinc finger protein; related to ZNF41 and ZNF674 | 2004 [86] |
| *GRIA3* | AMPA receptor GLUR3; mediates fast, synaptic transmission in central nervous system | 2005 [22] |
| *ARHGEF9* | Cdc42 guanine nucleotide exchange factor; pivotal role in formation of postsynaptic glycine and GABA(A) receptor clusters | 2005[c] [24] |
| *ZNF674* | KRAB domain-containing zinc finger protein; related to ZNF41 and ZNF81 | 2006 [26] |

[a] Also mutated in S-XLMR.
[b] Year when first implicated in NS-XLMR.
[c] VM Kalscheuer *et al*., personal communication.

Table 2: **Functions of NS-XLMR genes [69].** XLMR is not only caused by mutations in a lot of different genes, but also the functions of these genes diverge largely, from transcriptional regulation over mRNA splicing, synaptosomal endocytosis, regulation of actin cytoskeleton, neurotransmitter release and signal transduction to methylation of tRNA. Therefore, exclusion of genes as MR candidates due to distinct biochemical functions is essentially precluded.

is present in an MR patient, the breakpoint can be mapped more precisely using Fluorescent *in situ* Hybridization (FISH). This method is based on hybridization of labeled Bacterial Artificial Chromosome (BAC) probes to human metaphase chromosomes, each BAC containing approximately 100 kb of human genomic DNA. As the sequences of the BACs as well as their location in human genome is known, it is possible to narrow down the breakpoint to a resolution of around 100 kbp. The search for the precise breakpoint, interrupting the MR gene, can then be persued using Southern blot hybridisations.

Comparative Genome Hybridization (CGH) is another way of finding genomic rearrangements. This method relies on the hybridization of equal quantities of DNA fragments from both a MR patient and a healthy control individual to the chromosome of a normal cell. Labeling the patient- and control-derived DNA fragments with different fluorescence dyes prior to hybridization, genetic imbalances in the patient genome, due to deletions or duplications, can be detected by the ratio of the intensities of both fluorescence signals diverging from one at the affected site of the chromosome. However, inversions cannot be detected. With help of computational analysis, array-based CGH can be done with high resolution and throughput [15].

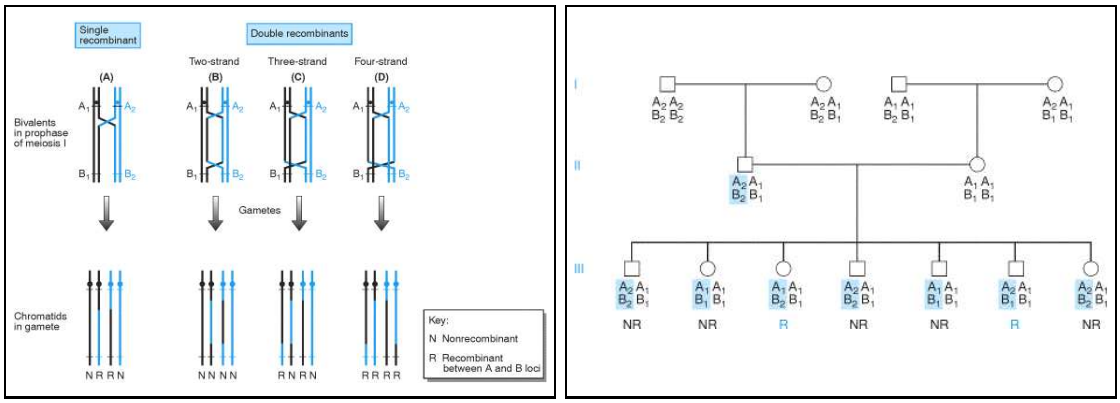
**1.1.4.2 Linkage Analysis** is performed when no genomic rearrangements can be observed. In a family where only males are affected and females act as mildly affected or unaffected carriers, X-linkage of the mutated gene can be assumed. To not have to sequence the entire X chromosome to find the mutated gene, linkage analysis is performed: Genotyping of polymorphic markers spread allover the chromosome is done first. Then, a computer software is used to find out which marker alleles are always present when the MR phenotype occurs and calculates the likelihood of the marker alleles being linked to the disease. This leads to the identification of a so-called linkage interval, which is a region that most likely contains the mutation. Finally, this region has to be screened for mutations in all affected individuals, e.g. by sequencing of all exons in all genes within it.

However, linkage analysis is a statistical method based on finding recombinations (see figure 2 for an overview) and thus provides results that are true with a probability of less than 100%. To increase the likelihood of finding a causative mutation, the linkage intervals of all families sharing the same phenotype can be merged. The shortest overlap of all linkage intervals will be screened subsequently.

When dealing with non-syndromic XLMR where no obvious phenotype accompanies MR, pooling of linkage data to narrow down the chromosomal location of the mutation is not possible. As two thirds out of all XLMR cases are non-syndromic, a way out of this dilemma is needed. By merging the linkage intervals from 125 unrelated families with NS-XLMR, a region in the proximal Xp was identified that was expected to contain approximately 30% of all NS-XLMR genes [71]. In a subsequent mutation analysis of 22 XLMR families with linkage to this region, mutations in the coding regions of four different genes were found.

### 1.1.5 Strategies of Curing Mental Retardation

With regard to the high abundance of mental retardation in the population and the resulting ethic and economic problems, curative strategies are desired. These strategies have to be adapted to the factors causing MR. Concerning environmentally caused MR, avoidance of the above-mentioned risk factors is relatively easy. This is not the case for genetic causes of the disability. In such cases, patient treatment with a drug specifically designed to counteract the aberrant MR gene is mandatory. Such drugs are not available for human use yet, although there is promosing report about treatment of Fragile X (FRAX) MR in adult *Drosophila* flies. Here, the effect of the causative gene mutation, leading to overexpression of metabotropic glutamate receptors (mGluR), could be countersteered by administration of mGluR antagonists, restoring short-term memory and courtship behaviour of the flies [52]. As in this animal model, treatment of inborn human neural disorders seems to be possible even after birth, as the nervous system keeps its plasticity in adult stages of life (see [10] and [97]).




(a) The process of recombination (b) Recombinants and non-recombinants in a family

Figure 2: **Recombination detection needs heterozygous crossover-flanking markers [79].** Recombination occurs between two chromatides, each of one homologous chromosome, during meiosis (see figure 2(a)). The individual the recombination takes place in is heterozygous for two markers A and B. (A) One crossover between A and B causes 50% recombination frequency in the gametes of this individual. If two crossovers happen, the recombination frequency depends on the chromatides involved in recombination: (B) The same chromatides crossover both times: both markers remain unrecombined, but recombination is found between both chiasms. In the same manner, three- and four-strand crossovers produce 50% and 100% recombinant chromatides, respectively ((C) and (D)). By consequence, a dense map of heterozygous markers is essential to detect recombinations efficiently. Single Nucleotide Polymorphisms (SNPs) satisfy this request, being evenly spread over the whole genome, having small distances to each other and being typeable quickly by high-throughput array methods. In figure 2(b), recognition of recombinants is shown for the above-mentioned markers. Each descendant inherits one chromosome and thus one combination of these markers from the mother and from the father. For the mother $II_2$ is homozygous for both markers, one cannot detect any recombinations during oogenesis, but knows that the mother will always inherit $A_1B_1$. Descendants that are recombinant (R) for the paternally inherited chromosome (boxed in blue) can be distinguished unambigously from the non-recombinants (NR).

## 1.2 Experimental Outline of this Diploma Thesis

Through a search for mutations causing NS-XLMR in a 7.4 Mb interval in the Xp11 region, a cytosine to adenine change was found in nine patients, located 188 bp upstream of the translation start site of the *PLP2* gene.

In this diploma thesis, I tried to find out if this transversion has an impact on the expression level of the gene and how frequent it is in the common population and in people with at least average IQ. Finally, after having found reduced expression levels, I searched for altered binding of transcription factors which could explain this observation, since for a long time, transcription rates are known to be controlled by binding of proteins to regulatory elements located mostly upstream of the transcription start site of the controlled gene [6]. The experimental outline is illustrated in figure 3.

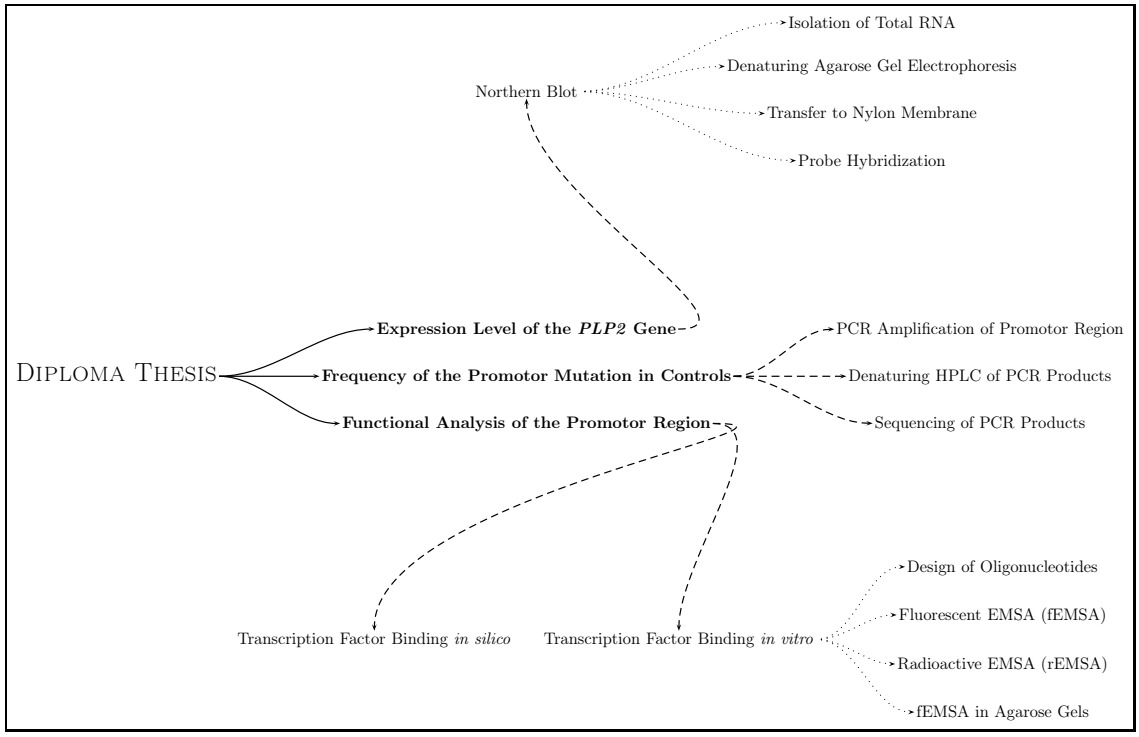



Figure 3: **Flow chart giving an overview of the projects and sub-projects of this diploma thesis.** First, *PLP2* expression was analyzed in patients and controls by Northern blotting. In parallel, the frequency of the promoter mutation was investigated in an indirect sequencing approach, using Denaturing HPLC to discard candidates prior to sequencing. Last, binding of transcription factors to the investigated promoter site was analyzed *in silico* and *in vitro* with help of transcription factor databases and the Electrophoretic Mobility Shift Assay (EMSA) technique.

## 1.3 Literature Review on the *PLP2* Gene

Until now, seven articles were published on the *PLP2* gene and its protein product, sometimes alternatively called A4 protein. None of these articles establishes a direct connection to mental retardation.

In 1993, Oliva *et al.* published a paper about a differentiation-dependent gene in the human colonic cell line HT29-18 [57]. By differential screening of a subtraction cDNA library of a highly differentiated variant of that cell line in comparison to the undifferentiated one, they found a full-length cDNA clone of 945 bp, containing an open reading frame (ORF) of 456 bp, corresponding to 152 amino acids that are grouped into five exons in the PLP2 protein. The protein contains four potential membrane-spanning domains, several sites for posttranslational modification and is structurally similar to the family of proteolipids. Southern blot analysis revealed that *PLP2* is a single-copy gene and conserved in the rat. Northern blot analysis and *in situ* hybridization showed high expression in both human and rat colonic epithelium. The promoter region was found to lack a TATA box, but to contain several regulatory elements also present in the cystic fibrosis transmembrane conductance regulator gene.

Two years later, Oliva *et al.* published another paper on *PLP2* [56], this time characterizing the promoter region of the gene. The goal of the research was to find an explanation for the observed high abundance of *PLP2* transcripts in differentiated cells. By transfecting colonic cells with constructs of a bacterial reporter gene fused to varying lengths of sequence from upstream of the *PLP2* transcription start site, they found one DNA element that silenced transcription rates and two that enhanced them. Using DNase I footprinting essays and EMSAs, one enhancer, containing the consensus Sp1 binding sequence, could be demonstrated to maintain a basal level of PLP2 in both differentiated and undifferentiated colonic cells. Interestingly, the other enhancer was found to bind a 50 kDa protein that is much more abundant in the differentiated cells, thus, by binding of this stimulatory protein, the higher *PLP2* expression level in the differentiated cells was seen to be explainable. The identity of the protein could not be revealed.

In 1997, Breitwieser *et al.* (under co-work of M.M. Oliva) demonstrated by immunocytochemistry that PLP2 localizes to the endoplasmic reticulum (see figure 4) and by coimmunoprecipitation that it is able to multimerize [11], both pointing towards the abovementioned idea of PLP2 belonging to the group of proteolipids. To substantiate this, the authors performed electrophysiological studies in nuclei isolated from microinjected *Xenopus laevis* oocytes transiently expressing *PLP2*, and indeed found PLP2 multimerizing to form a 28-pS ion channel.

The next two papers, published in 1997 and 2002, respectively, did not add any biochemical knowledge about the PLP2 protein, but provide insights into the gene's chromosomal location and its precise intron-exon structure. By constructing YAC contigs spanning Xp11.23-p11.22 and sequencing them subsequently, Fisher *et al.* [25] were able to localize the *PLP2* gene to the X chromosome. Five years later, Strausberg *et al.* [80] presented a paper describing the generation of high quality sequences for complete ORF cDNA clones

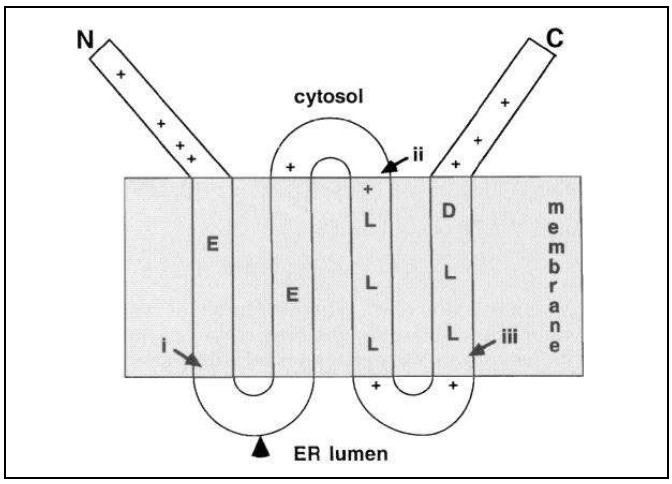



Figure 4: **Membrane topology of the PLP2 protein [11].** The protein has four transmembrane domains, the N and C terminal end remaining on the cytosolic side of the Endoplasmic Reticulum (ER) membrane. The first three exon($n$)-exon($n$+1) junctions are indicated by i, ii and iii. The arrowhead denotes a putative cleavage site for signal peptidases.

for any human and mouse gene. One of those genes was *PLP2*.

A 2003 paper by Wang *et al.* [89] contained a link from *PLP2* to apoptosis. PLP2 was found to associate with BAP31, a protein that is also present in the membrane of the endoplasmic reticulum and that is a substrate of caspase 8. Upon cleavage of BAP31 by caspase 8, p20BAP31 is generated. p20BAP31 has proapoptotic features, promoting cellular condensation and cytochrome $c$ release from the mitochondria. In contrast, full-length BAP31, maintained by mutation of the caspase 8 cleavage site of BAP31, acts as an inhibitor of apoptosis. Interestingly, this feature was found to be independent from the BAP31 interaction partner p20BAP31, such that another binding partner of BAP31 was searched in a yeast two-hybrid screen – and found in the PLP2 protein.

The latest paper on *PLP2*, published in 2004 by Lee *et al.* [48], demonstrates, as a result of a yeast two-hybrid library screen, an interaction of PLP2 with the CC chemokine Receptor 1 (CCR1). Chemokines are proinflammatory cytokines that bind to G Protein-Coupled Receptors (GPCRs) of the seven transmembrane domain family. One of these receptors is CCR1. It is known to play important roles immune responses and inflammatory processes, such as chemotaxis, that is the migration of immune cells towards certain chemical compounds, as e.g. that of neutrophils towards bacterial cells, recognized by their unique formylated methionine residues in membrane proteins. The PLP2-CCR1 interaction was confirmed by mammalian two-hybrid and coimmunoprecipitation analyses as well as by colocalization following indirect immunofluorescence. Overexpression of *PLP2* lead to a twofold increase of migration of HOS/CCR1 cells, implicating a functional role for PLP2 in the chemotactic processes via CCR1.

# 2 Materials and Methods

## 2.1 Expression Level of the *PLP2* Gene

### 2.1.1 The Principle of Northern Blotting

The technique of Northern blotting is widely used to detect the expression levels of genes. As a first step, total RNA has to be isolated from cells, removing all DNA and proteins. The isolated RNA is then submitted to denaturing agarose gel electrophoresis. Formaldehyde, contained in the electrophoresis buffer and in the gel, inhibits the single-stranded RNA from establishing intramolecular Watson-Crick basepairs due to partial self-complementarities. These secondary structures would influence the electrophoretic mobility of the RNA and the subsequent hybridization of the probes. After electrophoresis, all RNAs are transferred to a membrane by means of capillary forces (see figure 5). This immobilizes the RNAs and makes them accessible at the same time. Placing the membrane in hybridization buffer, a single-stranded and radioactively labeled DNA probe, designed to be complementary to an RNA of interest, is added. Having bound its complementary counterpart, which is mostly the messenger RNA of a gene under investigation, a washing step removes all remaining unbound probe from the membrane. The radioactive label of the probe allows the localization of the RNA of interest on the membrane. Visualization and semi-quantitation of the signal is possible using X-ray film or a PhosphorImager system.

After RNA transfer to the membrane, UV crosslinking immobilizes the RNA on the membrane. The primary amino groups, present on the nylon membrane, react with the UV-activated thymines in the RNA, establishing a covalent and thus irreversible binding of the RNA to the membrane [16].

When using a nylon membrane for RNA transfer, a prehybridization step is necessary: Due to the membrane's positive charge, all sites that received RNA during the blotting procedure will be uncharged, because the negative charges of dissociated RNA phosphate groups compensate for the membrane's charge. All other sites remain positively charged. When hybridizing, these sites can bind the probe unspecifically, resulting in a high background intensity. To prevent this, the membrane is incubated in prehybridization buffer, compensating all charges, prior to probe hybridization. Additionally, some protocols propose to add denatured, single-stranded salmon sperm DNA to the buffer, which will bind unspecifically all over the membrane. In both ways, the probe will only be able to bind specifically to the complementary transferred RNA on the blot.

Radioactive labeling of the probe for Northern blot hybridization is done using a randomly primed fill-in reaction, catalyzed by the Klenow fragment of DNA polymerase I (see figure 6).

Generally, when investigating gene expression and dealing with low signal intensities – due to low copy numbers of a gene's mRNA transcript –, a purification step, proceeding gel electrophoresis, makes sense: using a column whose matrix consists of beads carrying poly-T-tails, poly-A-containing eucaryotic mRNAs can be isolated from the total RNA, discarding all tRNAs, rRNAs and other RNA variants. This procedure usually increases

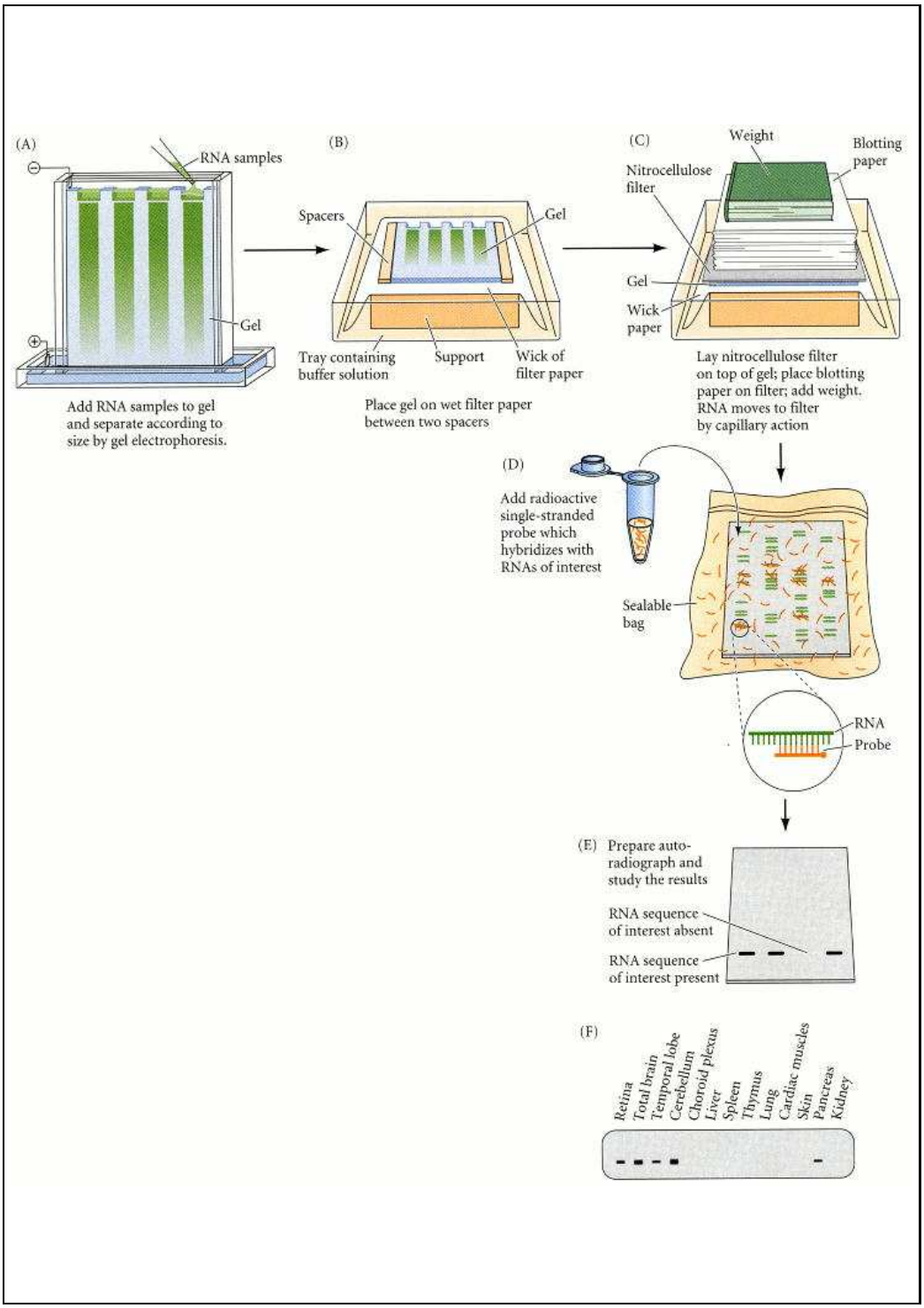



Figure 5: **Principle of Northern blotting [32].** Total RNA, isolated from cells, is separated by denaturing agarose gel electrophoresis (A). In the blotting sandwich, several layers of blotting paper suck the transfer buffer through the agarose gel and thereby transfer all RNAs in the gel to the membrane (B and C). The RNA of interest, immobilized on the membrane lying in hybridization buffer, is bound by its complementary, single-stranded and radioactively labeled probe (D), giving a clear radioactive signal at the position where the RNA is bound to the membrane (E). The amount of radioactivity represents the amount of RNA present (F).

signal intensity by two orders of magnitude [32], for mRNAs represent only a small fraction out of the total RNA present in a cell.

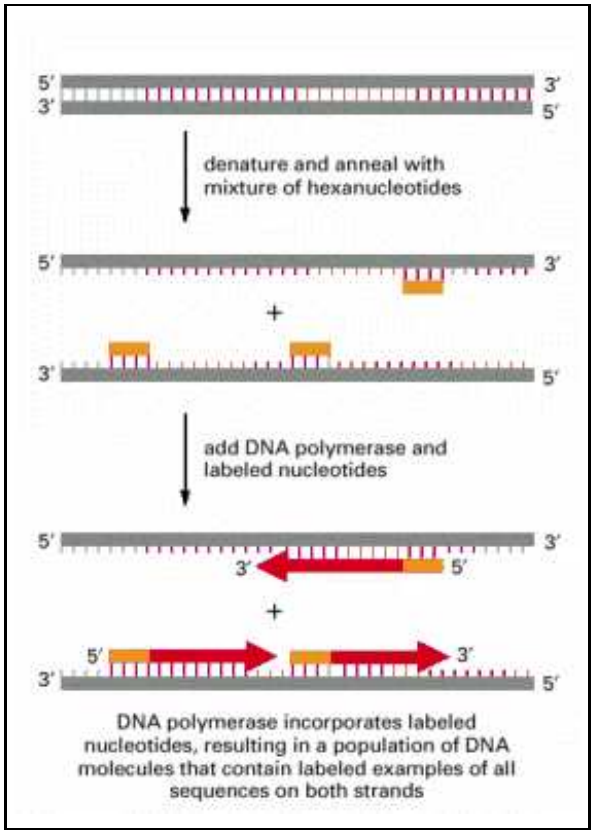



Figure 6: **Principle of random hexanucleotide labeling (after [1]).** After initial denaturation, hexanucleotides of random sequence are added to the melted DNA strands and bind to them via complementary base pairing. These hexanucleotides are prolonged by a DNA polymerase, incorporating radioactively labeled and non-labeled nucleotides. Thus, both single-strands are filled up more or less completely to give double-strands, depending on the processivity of the enzyme. As a polymerase, the Klenow fragment of DNA polymerase I is usually used. Besides the fact that it exhibits no exonuclease activity allowing it neither to proof-read the last incorporated nucleotides nor to excise those polynucleotide stretches lying ahead in 3' direction, it is not able to ligate any strands together. Therefore, newly synthesized strands will normally not cover the length of the whole template. In the subsequent hybridization steps, this will not play a role, because the short radioactive fragments will find their target anyway.

### 2.1.2 RNA Quantitation

To calculate the volume of RNA solution needed to use exactly 15 µg of each RNA isolate in the subsequent blotting procedure, all isolated RNA samples had to be quantified. This was achieved using a spectral photometer at a wavelength of 260 nm. At this wavelength, DNA and RNA absorbe UV irradiation maximally. Thus, the concentration of the RNA is directly proportional to the measured optical density, following the law of Lambert-Beer

$$OD_{260\,\mathrm{nm}} = \epsilon c d,$$

where $OD$ is the optical density, $\epsilon$ the extinction coefficient, $c$ the concentration of the RNA and $d$ the diameter of the cuvette containing the RNA solution. $\epsilon$ and $d$ being constant, the concentration of the RNA can directly be calculated by measuring the $OD$ and solving the upper equation for $c$. For 1 OD of RNA corresponds to an RNA concentration of 40 µg/ml, one can determine the RNA concentration as a multiple of the OD:

$$c_{\mathrm{RNA}}\,[\mu\mathrm{g/ml}] = 40 \cdot Dilution\ factor \cdot OD_{260\,\mathrm{nm}}.$$

At 280 nm, the proteins present in the RNA solution exhibit their absorption maximum, meaning that, by calculating the quotient

$$\frac{OD_{260\,\mathrm{nm}}}{OD_{280\,\mathrm{nm}}},$$

one determines the purity of the RNA isolate as the amount of RNA (wanted) in relation to amount of protein (unwanted). Ratios indicating a clean, nearly preotein-free isolation can be found in a range from 1.5 to 2.0, making 1.8 the average value specifying a clean RNA preparation.

### 2.1.3 Experimental Protocol

**2.1.3.1 RNA Isolation** was performed from the pellets of lymphoblastoid cell lines, using phenol and chloroform and following the subsequent protocol.

1. Cell lysis
   (a) Dissolve pellets of approximately 1 – 10 million cells in 10 ml TRIzol by shaking and incubate dissolved cells for 1 h at room temperature (RT)
2. RNA extraction
   (a) Transfer dissolved cells to centrifuge tube and add 2 ml of chloroform
   (b) Shake for 15 s to mix both phases and leave on the bench for 2 – 3 min
   (c) Centrifuge at 10,000 g and 4 °C for 15 min
   (d) Transfer upper aqueous and RNA containing phase to new centrifuge tube
3. RNA precipitation and washing
   (a) Add 5 ml of isopropanol to each tube and shake vigorously
   (b) Incubate for 10 min at RT
   (c) Centrifuge at 10,000 g and 4 °C for 10 min
   (d) Discard supernatant and add 10 ml of 75% ethanol
   (e) Mix vigorously and centrifuge for 5 min at 7,500 g and 4 °C
4. RNA solubilization
   (a) Carefully discard supernatant (pellet might attach only weakly to the tube)
   (b) Place centrifuge tube, head first, on towels and let pellet dry completely
   (c) Dissolve pellet in 500 µl DEPC-treated Aqua ad iniectabilia
   (d) Spin down RNA solution and transfer to microcentrifuge tube
   (e) Store tubes at –20 °C
5. Determination of RNA concentration
   (a) Thaw one RNA aliquot of each sample on ice
   (b) Clean nanodrop sample loading area with Aqua ad iniectabilia and disposable paper
   (c) Start DNA/RNA quantitation program and zero device with 2 µl of Aqua ad iniectabilia
   (d) Load 2 µl of each sample and record concentration of RNA and purity in the sample

**2.1.3.2 Verification of Isolation Success** was done not only by photometric quantitation, but also by agarose gel electrophoresis under non-denaturating conditions, following the protocol stated below.

1. Gel preparation
   (a) Prepare a 1.5% solution of agarose in 1x TAE buffer
   (b) Boil solution in a microwave oven to completely dissolve all agarose in the buffer
   (c) Cool down at RT to approximately 70 °C while stirring and add ethidium bromide to a final concentration of 0.5 µg/ml
   (d) Pour gel and let it set for 1 h
2. Electrophoresis
   (a) Refresh TAE buffer in electrophoresis chamber
   (b) Place gel-containing tray in chamber
   (c) Add loading buffer to the RNA samples to give a final concentration of 1x
   (d) Load samples to the gel and apply 100 V direct voltage
   (e) Run gel until lower gel loading buffer band reaches last third of the gel
   (f) Document the run on a gel documentation system

**2.1.3.3 Denaturing Agarose Gel Electrophoresis** followed the estimation of the RNA concentration. 15 µg of each RNA isolate was run on seperate lanes of a denaturing, formaldehyde-containing 1% agarose gel (see the following protocol).

1. Gel preparation
   (a) Prepare a solution of 1% agarose in DEPC-treated bidistilled water, also containing 6.5% formaldehyde and 1x MOPS buffer
   (b) Boil in a microwave oven and allow to cool down at RT to approximately 60 °C
   (c) Pour gel in gel tray and allow to set for 1 h
   (d) As electrophoresis buffer, fill 1x MOPS containing 6.5% formaldehyde in electrophoresis chamber
   (e) Rinse gel slots with 1000 µl pipette
   (f) Prerun gel at 60 V to equilibrate with the electrophoresis buffer
2. Loading of samples
   (a) Loading dye
      i. Add a few microliters of loading dye to the first and last lane to visualize progression of electrophoretic migration
   (b) RNA ladder

i. Add 10 µl loading buffer and 4 µl Aqua ad iniectabilia to 6 µl RNA ladder
ii. Vortex briefly and spin down
iii. Denature for 5 min at 70 °C and place immediately on ice

(c) RNA samples

i. Add loading buffer to samples
ii. Denature for 10 min at 65 °C and cool down immediately on ice
iii. Load to gel

3. Electrophoresis

(a) Apply 60 V for 30 min to allow especially large RNAs to leave the pockets and enter the gel
(b) Increase voltage to 100 V and run until blue dye has reached lower third of the gel
(c) End run and wash gel 3x in DEPC-treated water to remove most formaldehyde from buffer residues on the gel
(d) Align fluorescence emitting ruler with zero to gel pockets and photograph gel on a gel documentation system

**2.1.3.4 Blotting** onto a nylon membrane was done using the classical wet method (see below-mentioned protocol).

1. RNA Transfer to the membrane

(a) Cut off gel slots and upper left corner of the gel using a scalpel
(b) Fill a chamber with 10x SSC buffer and place a glass tray on the chamber's margins
(c) On the glass tray, place two layers of Whatman paper, avoiding to form air bubbles (this is also crucial for all subsequent membrane-placing activities); paper ends have to reach buffer reservoir in chamber
(d) Place gel on paper
(e) Cut transfer membrane to gel size, put on gel and wet using SSC buffer
(f) Cover Whatman paper not covered by gel with parafilm to prevent short circuit of buffer stream when transferring RNA
(g) Cut Whatman paper to gel size and place directly on transfer membrane
(h) Place half a pack of disposable towels on the paper
(i) Weigh the towels down by no more than 200 g
(j) Incubate over night to let as much RNA as possible move from the gel to the transfer membrane

2. Covalent crosslinking of RNA to the membrane

   (a) Disassemble blotting sandwich and put transfer membrane on disposable towel, RNA side up

   (b) Let membrane dry for 5 min at RT

   (c) Put membrane in crosslinker and apply autocrosslink program at 1200 µJ for two times

   (d) Shrink-wrap membrane in plastic film

   (e) Store at –20 °C until further use, e.g. hybridization of probes

**2.1.3.5 Hybridization** was done with first a single-stranded probe for the *PLP2* gene and, after removing the *PLP2* probe, with a second probe for *β*-actin, in order to reference the *PLP2* expression level. The protocol was as follows.

1. Radioactive labeling of probe

   (a) Add 4 µl OLB buffer (contains random hexamer primers and nucleotides) and 11 µl of Aqua ad iniectabilia to 2 µl double-stranded probe (weak PCR product giving at least 25 ng is sufficient)

   (b) Denature for 5 min at 95 °C and chill on ice

   (c) Add 2 µl $\alpha^{32}$PdCTP and 1 µl Klenow fragment

   (d) Incubate at 37 °C for 45 min

2. Prehybridization for compensation of membrane charge

   (a) Cut transfer membrane out of plastic film, roll up with RNA side inside and put in hybridization tube

   (b) Add 10 – 15 ml pre-warmed hybridization buffer

   (c) Incubate for 45 min at 42 °C in hybridization oven (don't forget to apply balancing tube)

3. Purification of probe

   (a) Equilibrate matrix of mini gel column by vortexing thoroughly

   (b) Remove lower end of column to open it and unscrew lid for half a complete rotation

   (c) Move column to microcentrifuge tube and centrifuge for 1 min at 730 g to remove protective buffer from the column matrix

   (d) Discard microcentrifuge tube and place column in a new one

   (e) Apply probe to the center of the column matrix surface by pipetting and avoiding neither to touch the column wall nor the matrix

   (f) Centrifuge for 2 min at 730 g and check radioactivity of eluate and column

(g) Discard column and save microcentrifuge tube containing purified probe

4. Hybridization

(a) Denature purified probe for 5 min at 95 °C

(b) Pipet probe directly into the hybridization buffer of the prehybridizing membrane

(c) Incubate over night in hybridization oven

**2.1.3.6 Analysis of Northern Blot** was done by exposing X-ray film and a PhosphorImager plate to the hybridized membrane, following the subsequent protocol.

1. Wash membrane

(a) Remove hybridization tube from oven and dispose buffer into liquid radioactive waste

(b) Transfer membrane to plastic tank using a tweezers

(c) Check radioactivity

(d) Wash for 5 min and then for 10 min in washing buffer 1

(e) Check radioactivity again: if not markedly decreased (high background), wash again in washing buffer 2

(f) Shrink-wrap membrane in plastic film and dry with disposable towels

2. Exposure of X-ray film

(a) Attach membrane inside film cassette

(b) In dark chamber, place X-ray film on membrane and close cassette

(c) Incubate over night in –80 °C freezer

(d) Take cassette out of freezer and let thaw at RT

(e) Open cassette in dark chamber and feed film to film developing machine

(f) Analyze developed film:

- expose new film for shorter time or use a less sensitive one if film is overexposed (and *vice versa*)
- wash again in washing buffer 2 if background is still high

3. Exposure of PhosphorImager plate

(a) Place membrane in PhosphorImager cassette

(b) Close cassette and incubate at RT over night

(c) Scan plate on PhosphorImager scanner

(d) Delete image on plate by white light

(e) Analyze scan using PhosphorImager associated software and expose for longer time, if signal to noise ratio is too small

**2.1.3.7 Hybridization of a New Probe** was done after washing off the old probe, using a strip off buffer and the following protocol.

1. Place membrane in plastic tank
2. Boil strip off buffer in microwave oven and pour on membrane
3. Incubate for 30 min at RT
4. Dry membrane on disposable towels
5. Continue as explained above to hybridize another probe

**2.1.3.8 Generation of the *PLP2* Probe** was done by Polymerase Chain Reaction (PCR, invented and firstly described by Kary B. Mullis in [73]), using the primers shown in table 3.

Table 3: **Sequence of the PCR primers used to generate the *PLP2* Northern blot probe.** Both primers were ordered from MWG Biotech.

| *PCR primer name* | *Primer sequence* |
|---|---|
| PLP2 probe forward primer | 5'-ttcagccaatgctttctctc-3' |
| PLP2 probe reverse primer | 5'-gctgtttccactccatttcc-3' |

This gives a PCR product of 533 bp, amplifying the extreme 3' terminal region of the gene, as shown in the Genome Browser window [45] after BLAT alignment [44] (see figure 7).

**2.1.3.9 Obtainment of the $\beta$-Actin Probe** was done by directly purchasing it from Biocat.

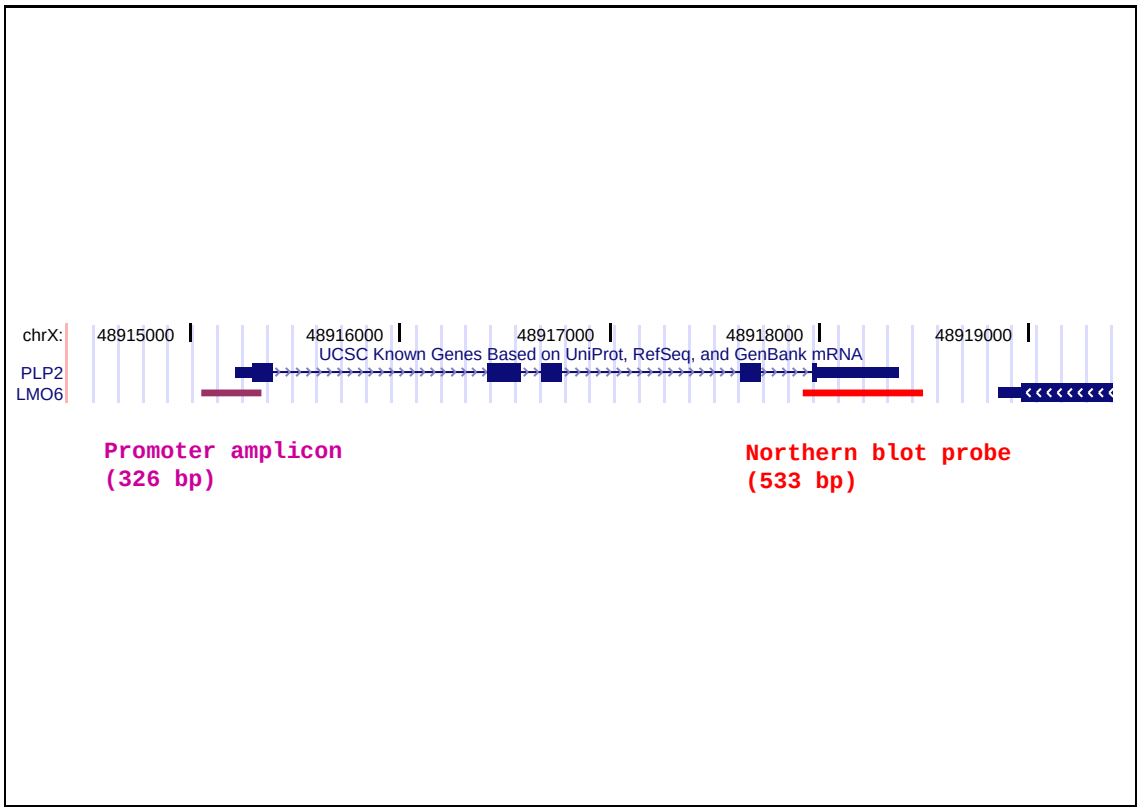



Figure 7: **Location of the *PLP2* probe and the promoter PCR product [40, 45].** A schematic representation of the *PLP2* gene is shown. The Northern probe (in red) binds the extreme 3' end of the genomic *PLP2* sequence, containing the gene's last exon and the complete 3' untranslated region. The PCR amplicon in the promoter region (in purple) has a total length of 326 bp and contains the site of mutation, the complete 5' untranslated region as well as a part of the first exon.

### 2.1.4 The PhosphorImager System

Molecular Dynamics' Storage Phosphor System provides a powerful technique to analyze autoradiographs, such as e.g. the Northern blots described above. Its dynamic range is 1000 times greater than that of any X-ray film, allowing direct comparison of weak and strong signals on the same exposition [31]. For strong signals drive X-ray film very fast into saturation and thereby escape from quantitation, this feature is a great improvement concerning speed and reliability of analyses. In addition, exposure to radioactivity does not happen inside the scanner, but inside a cassette, leaving the scanner free to work. After exposure and scanning, the screens can indefinitely often be restored to an unexposed state by a simple ten minute treatment with visible light. The provided software ImageQuant transforms user-defined bands of the scanner images into numerical values which then can be used to quantify signal intensity.

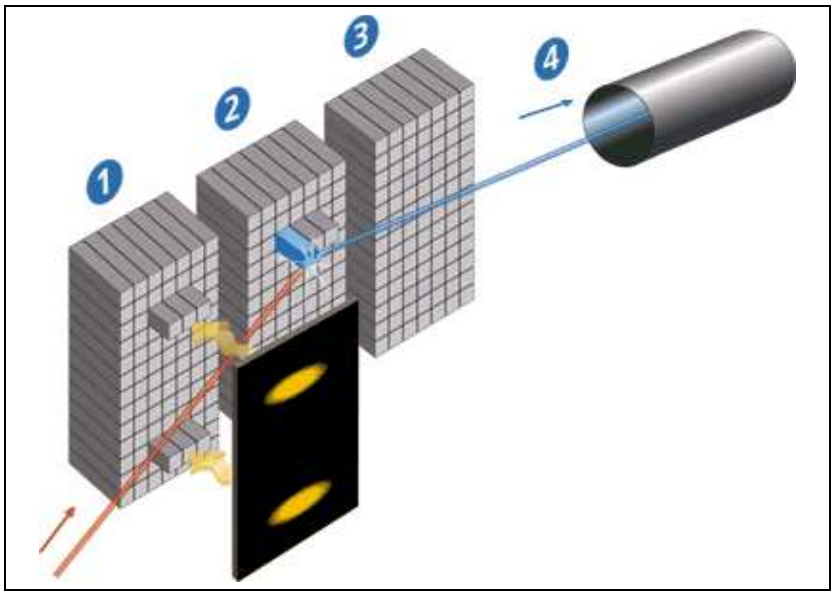



Figure 8: **Principle of the PhosphorImager System [31].** 1) Ionizing radioactive radiation causes a change in the quantum state of crystals in the storage phosphor layer. 2) Upon scanning with the Storm Laser Scanner, the BaFBR:EU+2 crystals in the layer emit blue light and return to the ground state in 80% of all cases (3). 4) The quantity of this blue light correlates directly with the amount of radioactivity previously present.

The PhosphorImager system was used both in order to quantify the radioactivity in the Northern blot, but also to analyze the results of the radioactive Electrophoretic Mobility Shift Assays (EMSAs).

## 2.2 Frequency of the *PLP2* Promoter Mutation

### 2.2.1 PCR Amplification of the Promoter Region

I used Qiagen's *Taq* DNA Polymerase Kit, as a number of other polymerases and systems did not yield any success, although there was a lot of testing. The final composition of each PCR mix is shown in table 4.

Table 4: **Final composition the used Qiagen PCR reaction mix**

| *Substance* | *Concentration stock* [mM] | *Volume per reaction* [µl] | *Final concentration* [mM] |
|---|---|---|---|
| Qiagen Buffer | 10x | 5 | 1x |
| Q Solution | 5x | 10 | 1x |
| $MgCl_2$ | 25 | 7 | 5 |
| dNTPs | 25 (each) | 1 | 0.5 (each) |
| Forward Primer | 0.01 | 1 | 0.0002 |
| Reverse Primer | 0.01 | 1 | 0.0002 |
| *Taq* DNA Polymerase | 5 U/µl | 0.25 | 1.25 U/reaction |
| Aqua ad iniectabilia | | 22.75 | |
| DNA template | ≈ 50 ng/µl | 2 | ≈ 100 ng/reaction |
| **Total** | | **50** | |

The sequence of the primers used for this PCR can be found in table 5.

Table 5: **Sequence of primers for *PLP2* promoter PCR.** All primers were purchased from MWG.

| *Primer name* | *Primer sequence* | *Primer melting temperature* |
|---|---|---|
| PLP2 Forward | 5'-cgagaggcgctcagaatc-3' | 59,8 °C |
| PLP2 Reverse | 5'-tcccgctattactgctcctg-3' | 60,4 °C |

The resulting PCR product has a size of 326 bp and is located on the Crick strand, upstream of the transcription start of the *PLP2* gene (see figure 7). It has a high GC content of 71% [55].

The final cycling program for screening the controls was as follows:

1. Initial denaturation
   (a) 94 °C for 3 min
2. Amplification
   (a) Denaturation: 94 °C for 30 s
   (b) Primer annealing: 55 °C for 30 s
   (c) Primer extension: 72 °C for 1 min
3. Thermal cycling

   (a) 44x step 2

4. Final extension

   (a) 72 °C for 10 min

5. Program end

   (a) Hold 8 °C endlessly

A simplified hot start technique was applied, not loading the PCR plates into the PCR machine until the block temperature had reached more than 80 °C.

Verification of PCR success was done by non-denaturating agarose gel electrophoresis. According to the protocol described in section 2.1.3.2, an 1.5% agarose gel was poured. After polymerization of the gel, all samples were loaded using ready-mixed loading buffer at a final concentration of 1x. The size of the PCR products was estimated by comparison of the position of the PCR product band to the position of the marker bands of known sizes (see figure 17 on page 58).

### 2.2.2 Analysis of the PCR Products on the DHPLC System

**2.2.2.1 The Principle** of denaturing HPLC (DHPLC) relies on the detection of differences between PCR products of the same amplicon, but generated from different template DNA. Thus, DHPLC can be used to detect mutations in the PCR products generated from patient's DNAs without sequencing. Compared to sequencing, clear advantages of DHPLC are its speed and its inexpensiveness, making it the method of choice to prescreen PCR products for mutations prior to sequencing. However, the precise character of a found mutation cannot be determined, wherefore PCR products showing aberrant elution profiles, indicative for sequence variations, still have to be analyzed by sequencing.

The DHPLC method is based on hydrophobic interaction, taking place in a densely packed column. The column material consists of alkylated nonporous poly(styrene-divinylbenzene) particles that do not directly interact with DNA. In order to retain DNA in the column, a hydroorganic mobile phase is used, containing triethylammonium acetate. The triethyl ammonium cation can, due to its hydrophobic ethyl residues, bind to the column beads, thereby distributing ammonia's positive charges over the whole stationary phase. These positive charges interact with the negative charges of the dissociated sugar-phosphate backbone of the DNA. Hence, DHPLC is a type of an ion-pair reversed-phase liquid chromatography. The longer the DNA, the more negatively charged phosphate residues can attach to the ammonia ions, and the stronger the DNA will be retained when eluting it in a gradient of increasing concentration of the organic solvent acetonitrile. Thus, DNA fragments are eluted in the order of increasing length.

The breakthrough on the way to establishing DHPLC for mutation detection came from the finding that single-stranded (ss) DNA is eluted much earlier than double-stranded (ds) DNA of the same length. For ds-oligonucleotides are single-stranded at the position of the mismatch, they elute a little earlier than oligonucleotides of the same length and perfect matching. In addition, increasing the column temperature unwinds mismachted DNAs to a greater extent than perfectly matched ones, thus increasing the time difference between the elution of mismatched and non-mismachted DNAs of identical length. Hence, application of DHPLC for mutation detection requires the combination of matching and mismatching sequences. This is achieved by pairwise pooling of PCR products from different individuals, but identical amplicons (see figure 9).

When using this technique for mutation screening, one has to keep in mind that pooling of PCR products from two individuals that carry the same mutation in a homozygous state will result in a chromatogram without an early homoduplex peak – thus, one will fail to detect them as mutants. However, as mutations causing mental retardation occur at very low frequency, it is rather unlikely to exhibit such circumstance.

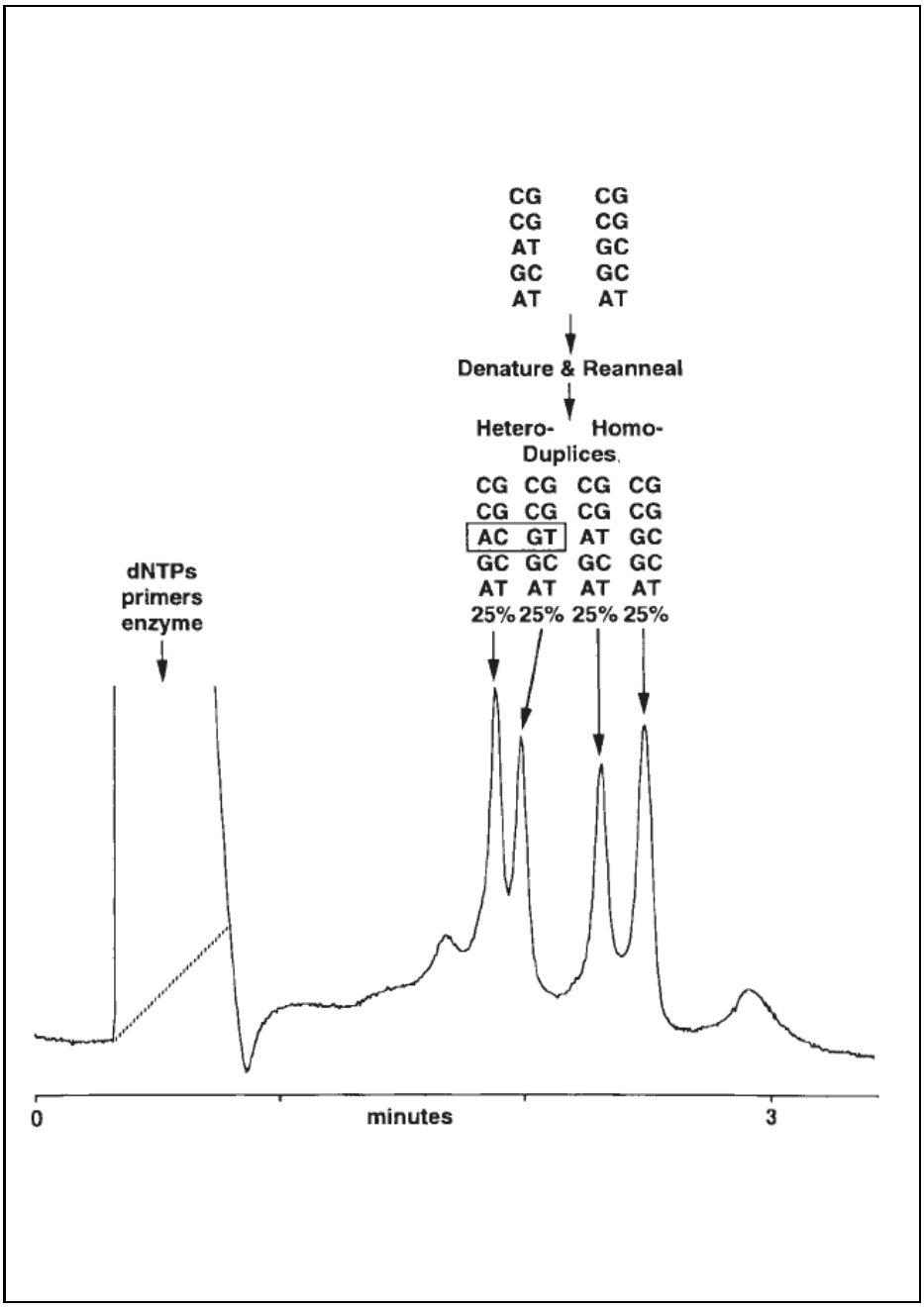



Figure 9: **Principle of Denaturing HPLC [96].** The principle of DHPLC relies on the earlier elution of heteroduplices, compared to homoduplices, in ion-pair reversed-phase liquid chromatography. The PCR products of the same amplicons from two different people differ by only one base. After mixing and heat denaturation of both PCR products, the four different single-strands reanneal to form four different duplices: two heteroduplices, if single-strands meet that were originally not paired, and two homoduplices. Under partial denaturating conditions, heteroduplices will unwind to a larger extent than homoduplices, and therefore interact less strongly with the matrix. They are eluted earlier, resulting in a chromatogram of one early heteroduplex peak and a retarded homoduplex peak. Strictly spoken, both peaks consist of two sub-peaks, one from each of both heteroduplices and homoduplices, respectively.

**2.2.2.2 Preparations for DHPLC** comprized the pooling of PCR products from the same amplicon of two patients and a subsequent de- and renaturation step, which was done after the following protocol.

1. Pooling of PCR products:
   (a) Transfer 15 µl PCR product of one patient and amplicon into a well of a PCR plate for DHPLC analysis
   (b) Add 15 µl PCR product of another patient, but the same amplicon, to the first well (both PCR products have to be of comparable concentration and quality)
   (c) Seal PCR plate with adhesive film
2. De- and renaturation in PCR machine:
   (a) 95 °C for 3 min
   (b) 60 °C for 15 min
   (c) 37 °C for 15 min
   (d) 25 °C for 30 min
3. Store at –20 °C until run on the DHPLC machine
4. DHPLC analysis was not done by me but is done centrally for the whole department

### 2.2.3 Sequencing of Candidate PCR Products

**2.2.3.1 The Principle** of the dideoxy sequencing method, invented by Frederick Sanger [76], relies on a PCR-like approach. This approach requires a single-stranded template DNA that is amplified by a polymerase using template-specific primers. In contrast to PCR, no exponential increase of the number of template copies is desired, although a thermal cycling program and PCR-specific reagents are used. In addition to what a PCR reaction mixture normally contains, a sequencing reaction mix contains also one labeled 2'-3'-dideoxy NTP (ddNTP). In all cases where this ddNTP is incorporated by the polymerase, chain growth by further incorporation of nucleotides is terminated, for there is no 3'-hydroxy group in the ddNTP that could attach to the 5'-hydroxy group of the next dNTP (see figure 10).

In order to visualize the template copies, Sanger originally describes incorporation of one radioactively labeled dNTP in each reaction and subsequent polyacrylamid gel electrophoresis of the resulting chain termination fragments [76]. Of course, it is also possible to use other labeling techniques than described by Sanger, e.g. radioactive labeling of the sequencing primer or of the ddNTP, or – more advanced – fluorescent labeling in either of the positions described above. To date, ddNTPs, each labeled with a different fluorescence dye, and a fully automated capillary gel electrophoresis sequencer are used. Because of the distinguishable dyes, all four termination reactions can be analyzed at the same time in a single capillary.

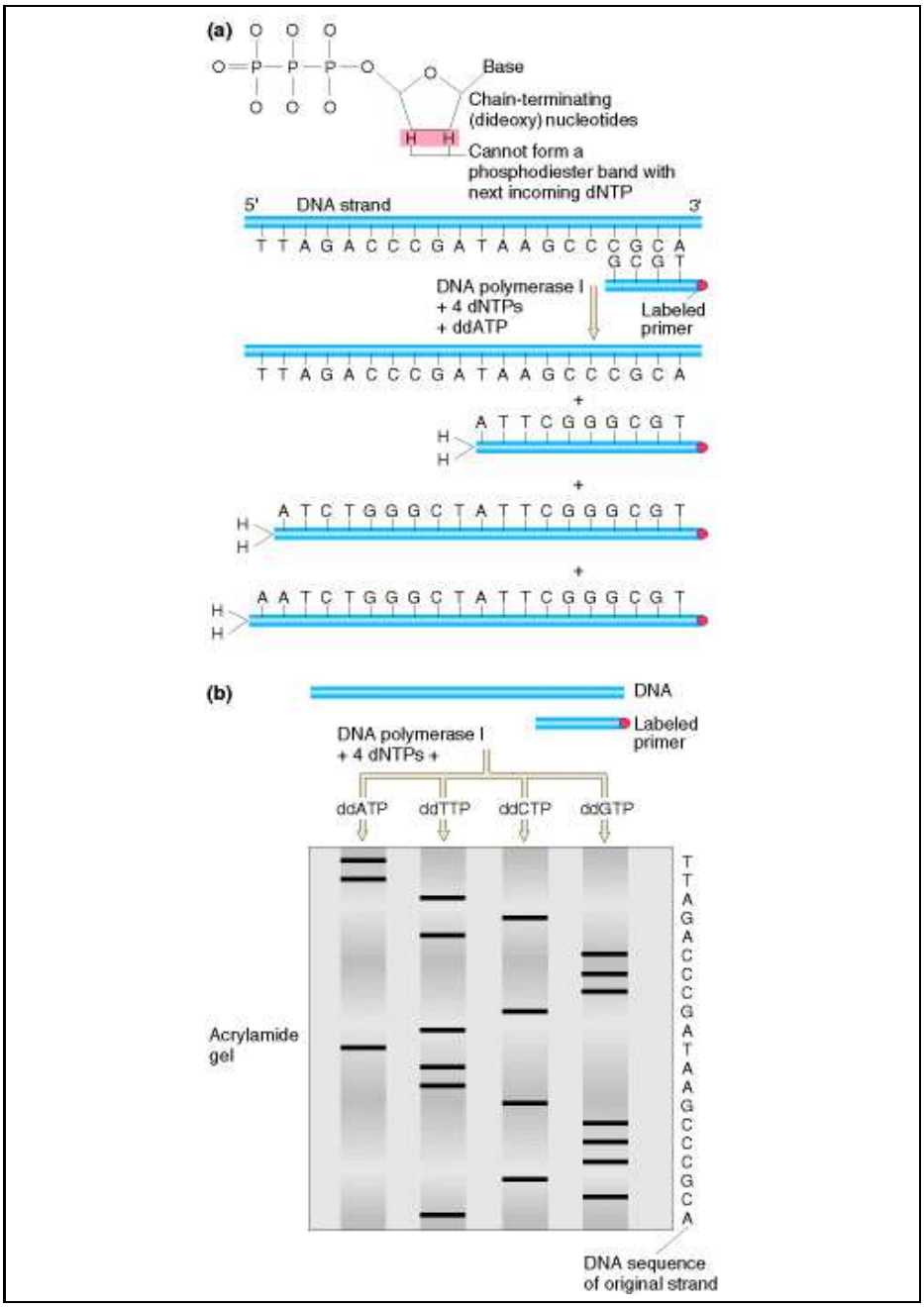



Figure 10: **Principle of Sanger dideoxy sequencing [34].** (a) In a PCR-like reaction, a primer on a single-stranded template is prolonged by incorporation of standard deoxy nucleotides (dNTPs), but also dideoxy nucleotides (ddNTPs) that have no hydroxy residue at the 3' position of the pentose. However, if a ddNTP was incorporated, chain growth is arrested, for no phosphodiester bridge to the next nucleotide can be established. Doing four PCR reactions in parallel, each with a different ddNTP and all four dNTPs, one gets chains of all possible lengths. (b) As the ddNTPs are radioactively or fluorescently labeled, the differing lengths of the termination reactions can be visualized by polyacrylamide gel electrophoresis. The template sequences can then be obtained by moving from the smallest termination fragment to the longest, always recording the ddNTP at the 3' end. The resulting sequence corresponds to the exact reverse-complement of the template's sequence, namely the sequence of the prolonged primer from 5' to 3'.

**2.2.3.2 Preparation for Sequencing** was to determine the mass of PCR product in the sample, normalized by its sequence length, for it was found empirically that an amount of 2 ng of DNA per 100 bp of sequence yields the best sequencing results. This determination was achieved by submitting the PCR products to agarose gel electrophoresis, loading a quantitative DNA ladder in one lane.

When documenting an agarose gel after electrophoresis, the intensity of the band belonging to the PCR product of interest has to be referenced visually to the intensity of the bands of the quantitative DNA ladder. For this ladder, the mass of DNA contained in each band is known (see figure 17 on page 58). Having thereby determined the mass of the PCR product band, one subsequently can normalize it to 100 bp of length.

An example: The intensity of a PCR product band is found to correspond the best to a ladder band whose mass is 20 ng. The length of the PCR product shall be 400 bp, and 5 µl was loaded to the gel. Hence it follows that the concentration of the PCR product is

$$\frac{20\,\mathrm{ng}}{400\,\mathrm{bp}} = \frac{5\,\mathrm{ng}}{100\,\mathrm{bp}}.$$

As described above, no more than 2 ng per 100 bp shall be used in the sequencing PCR. Therefore, only two fifth of the 5 µl originally applied to electrophoresis have to be used in the sequencing PCR, namely 2 µl.

## 2.3 Electrophoretic Mobility Shift Assay

### 2.3.1 Introduction

The Electrophoretic Mobility Shift Assay (EMSA), also known as gel retardation assay or gel shift assay, is a method to study *in vitro* DNA binding of proteins. It therefore mostly serves the investigation of regulatory DNA sequences, so-called *cis* elements, that are target sites for distinct proteins, so-called *trans*-acting factors. The regulatory sequences are called *cis* for they lie in the vicinity of the gene to be regulated, while the *trans*-acting factors are encoded by genes lying far away.

To perform an EMSA, one has to incubate a labeled DNA fragment, containing a putative regulatory site, with an extract of nuclear proteins, i.e. proteins with DNA binding capacity and regulatory function, or with a purified transcription factor. If a protein is able to bind the DNA fragment, a protein-DNA complex is formed. This complex is of larger size and higher molecular weight than the DNA itself and therefore has a lower electrophoretic mobility (see figure 11). Complex formation occurs with high sensitivity.

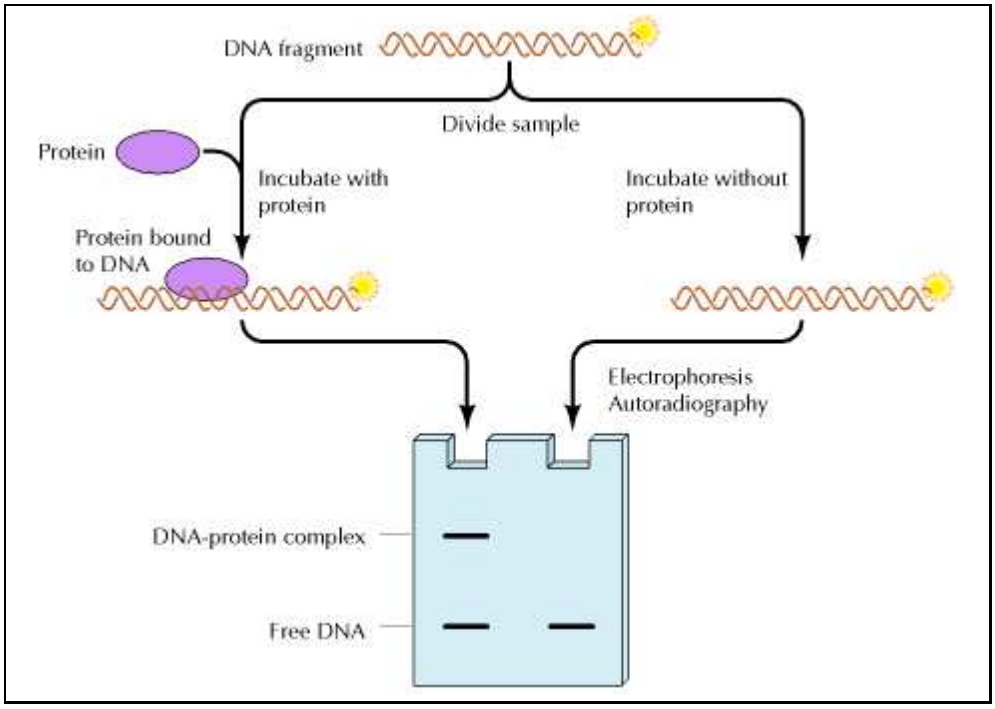



Figure 11: **Principle of Electrophoretic Mobility Shift Assay [17].** The principle of an EMSA relies on the reduction of the electrophoretic mobility of a DNA fragment by binding of regulatory proteins. A labeled DNA fragment is incubated with proteins. If they are able to bind it, this will result in a protein-DNA complex of larger size than that of the DNA fragment alone. Therefore, the electrophoretic mobility of the complex is lower than the mobility of the unbound DNA, resulting in an electrophoretic shift.

Binding of proteins to DNA can occur in a specific and a non-specific manner. Specific binding results from non-covalent interaction of the protein with the DNA in dependance of the DNA's steric structure. It is known that DNA can assume a variety of different and complex shapes, depending on external or internal factors. Dehydration, e.g., is an

external factor and leads to the transition of a classically B-helix shaped DNA to an A-helical structure, exhibiting a higher helix diameter and 11 bases per turn, compared to 10.4 in the B-type. In contrast, a Z-helix is caused by an internal factor, namely a sequence of alternating pyrimidines and purines. Z-DNA has the narrowest helix which is, as opposed to A- and B-DNA, left-handed [6]. Hence, the base sequence has a large influence on the steric structure of the DNA. In order to bind this structure, the protein itself has to have a structure that exactly matches that of the DNA, alike image and mirror image.

In contrast to this sequence-specific binding, also non-specific binding can occur. Under physiological conditions, DNA is, due to its dissociated phosphate groups, a highly negatively charged molecule and therefore can attach to any surface that is positively charged. Thus, positively charged proteins can bind to DNA irrespective of its sequence. To prevent such false-positive results in the EMSA, one has to add an excess of a suitable competitor, e.g. a different DNA fragment with a totally different sequence, or a synthetic DNA-like fragment, consisting of deoxy-inosinic-deoxy-cytidylic acid polymers of random sequence.

Verification of the specificity of the remaining shifts can then be done by adding an excess of an unlabeled DNA fragment to the binding reaction, the unlabeled fragment's sequence being identical with that of the labeled fragment. Due to the lack of label, any binding to this specific competitor is not visible. Thus, the intensity of the shifted band shall decrease with increasing amounts of specific competitor.

Identification of a specifically binding transcription factor can then be done in a super-shift assay, if candidates have been found in an *in silico* EMSA analysis. Adding and binding of an antibody specific for a candidate transcription factor will result in a size increase of the protein-DNA complex, leading to an additional retardation of the shifted band (so-called super-shift). If the antibody binds directly to the DNA binding domain of the protein, any specific DNA binding will be abolished, but can be restored by using another antibody raised against a different part of the protein.

Labeling of the DNA fragment to be bound can be done in different ways. Originally, the EMSA was done with radioactively labeled oligonucleotides (see figure 12). This yielded a very high sensitivity, because even weak signals can be "amplified" by simply increasing the exposition time of the photographic material. However, such treatment can make quantitation of signals of different intensities, present in the same scan, very difficult: a signal of high intensity will drive the screen into saturation very fast, causing weak signals to appear as intense as strong ones, if exposure times are chosen long enough. Therefore, determination of signal intensities is said to be only semi-quantitative. Of course, a general counter-argument against radioactive EMSAs is the inevitable exposure to radioactive radiation.

Hence, as soon as fluorescent dyes became available, they were tried as an alternative label. Quantitation of the the signal intensity can be done precisely, in contrast to the restrictions in radioactive assays, but the overall sensitivity is lower. Nevertheless, both systems should be evaluated.

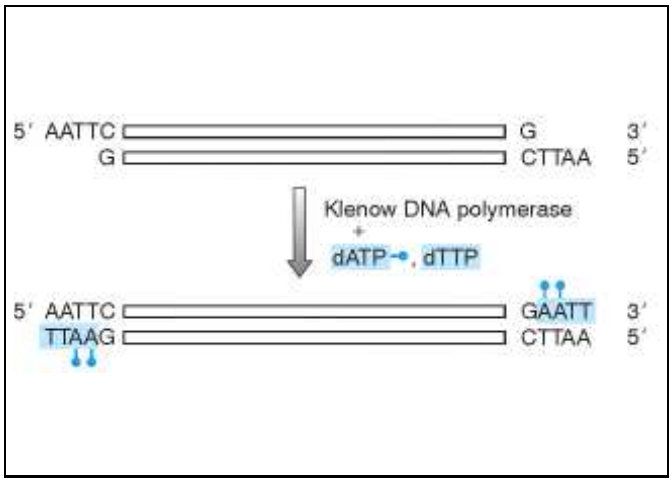



Figure 12: **Radioactive labeling of EMSA oligonucleotides (after [79]).** Single-stranded oligonucleotides were designed in a way that, after annealing, the doublestrand has protruding 5'-ends. These ends can then be filled up with radioactive nucleotides, using the Klenow fragment of DNA polymerase I.

### 2.3.2 EMSA Systems Used

#### 2.3.2.1 Fluorescent EMSA on A.L.F. DNA Sequencer

**The System** principally works is illustrated in figure 13. It consists of an electrophoresis unit, driven by a power unit, and a personal computer which controls all run parameters (see figure 14). In the electrophoresis unit, a polyacrylamide gel glass plate sandwich (see figure 16) is installed between an upper and a lower buffer chamber. The upper chamber contains the cathode, the lower chamber contains the anode, the electrophoresis direction being from top of the gel to the bottom. Upon laser excitation, optical detectors at the lower third of the gel and behind each gel lane allow the detection of fluorescence dyes coupled to the DNA oligonucleotides to be bound (see figure 15). This laser is able to excite fluorochromes such as fluorescein (often abbreviated as FITC), which absorbes laser light at 494 nm and emits light at 518 nm [14]. Quantitation of the amount of oligonucleotide bound by protein can be done by calculating the fraction of the fluorescence intensity of the retarded band with respect to the total amount of fluorescence per lane.

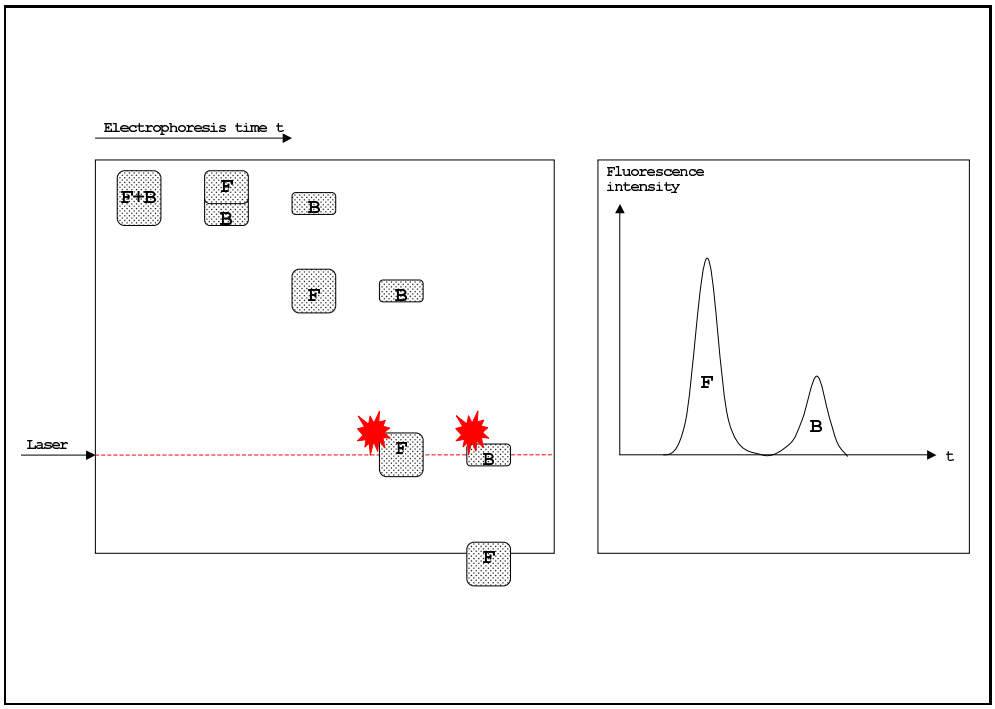
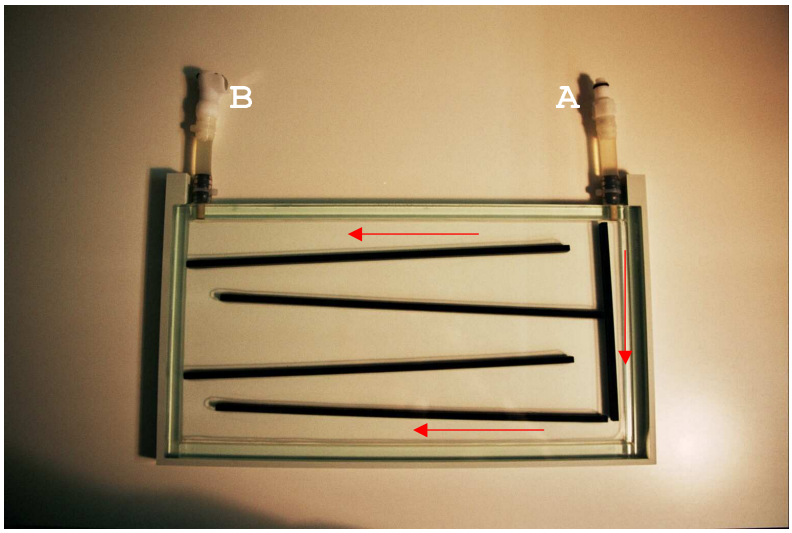


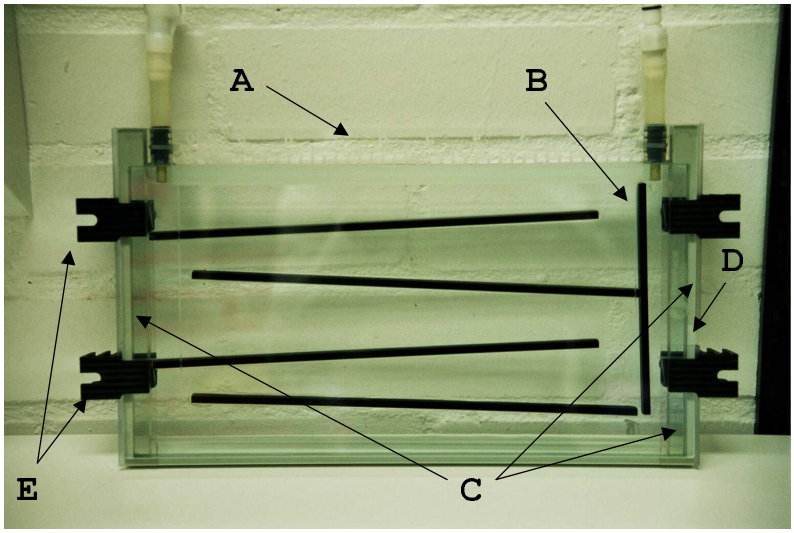

Figure 13: **Detection principle of the A.L.F. EMSA system.** Detection of protein-DNA complexes is based on the lower migration rate of the protein-bound oligonucleotide (B) in comparison to the free oligonucleotide (F). Therefore, the fluorescence dye of the free oligonucleotide is excited first by a laser beam penetrating the gel. Excitement causes a quantifiable fluorescence signal, the area under the fluorescence curve being proportional to the amount of fluorescence dye present and thereby also to the amount of oligonucleotide that has passed the detector. Determination of the amount of oligonucleotide that was bound can be done by calculating the fraction of the retarded band's peak area and the total peak area.


Figure 14: **Components of the A.L.F. EMSA System.** The system consists of a power supply (C) that drives an electrophoresis unit (B). The temperature of the polyacrylamide gel in the electrophoresis unit can be increased by the device itself, but cannot be decreased to room temperature or lower, so that an external thermostat (A) had to be used to prevent thermal denaturation of the proteins during electrophoresis. All parameters of the run were adjusted in the ALFWin software on a notebook (D), despite the temperature of the external water circulation which was directly set at the thermostat.


Figure 15: **View of the inside of the A.L.F. electrophoresis unit.** The polyacrylamide gel for electrophoresis, sandwiched between a cooling plate and a glass plate, is mounted vertically in this unit, using the screws (E). When mounted, the glass pushes down the contacts (F) that prevent the emission of laser radiation when no gel is present. On top, you find the connections for the water circulation (A) that lead to the thermostat integrated into the device. In order to use an external thermostat for cooling, these connections had to be loosened on their opposite end, leading to the unit's thermostat. Afterwards, new connections were mounted to directly connect to the external thermostat. The plugs of the power supply are indicated with (B) and are connected to the upper and lower buffer chamber. The laser beam leaves the unit on the left hand side (see arrow D). By turning a screw (dashed arrow line), one can adjust the angle of a mirror such that the laser beam enters the gel precisely via the light coupler. (C) indicates the panel of detectors, each detector being assigned to an individual lane of the gel.




(a) **The cooling plate.** (A) denotes the connection for water influx (from thermostat), whereas (B) is the connector for water outflow (to thermostat). Red arrows illustrate the way the cooling water flows through the plate.




(b) **The glass plate sandwich.** Between both glass plates, the polyacrylamide gel will be poured. (A) shows the glass plate with the inscription of the gel lane number, (B) is the cooling plate (located behind the glass plate). The position of the three spacers, establishing the distance between both plates and thus defining the thickness of the gel, is indicated by (C). The light coupler, through which the laser beam enters the gel, is located between the medium long and the short spacer on the right hand side (D). The screw clamps used to mount the glass plate sandwich in the electrophoresis unit are indicated by (E).

Figure 16: **Assembly of the "glass plate sandwich".** To pour a polyacrylamide gel, the cooling plate (a) and the glass plate have to be fixed together. Spacers between the plates were used to adjust the thickness of the gel. An assembled "plate sandwich" is shown in (b).

**Preparation of Nuclear Extracts** was done according to the following protocol.

1. Wash cells
   (a) Aliquote cell suspensions to $\approx 20 - 100 \cdot 10^6$ cells
   (b) Centrifuge for 10 min at 2000 rounds per min (rpm) and 4 °C
   (c) Discard supernatant and resuspend cells in 2 ml of ice cold 1x PBS buffer
   (d) Centrifuge for 10 min at 2000 rpm and 4 °C and remove supernatant
2. Isolate cell nuclei
   (a) Resuspend in four times the pellet volume of nuclear extraction buffer A (NEB A)
   (b) Incubate for 1 h on ice and place also some pestles on ice (preferentially one pestle for each sample)
   (c) Transfer cell suspension to pestle and homogenize cells by crushing 25 times
3. Lyse cell nuclei
   (a) Transfer homogenized samples to microcentrifuge tubes
   (b) Centrifuge for 5 min at 2000 rpm and 4 °C, discard supernatant
   (c) Resuspend pellet in 1 ml NEB A and centrifuge for 5 min at 2000 rpm and 4 °C
   (d) Discard supernatant and resuspend pellet in twice the pellet volume of NEB B
   (e) Incubate for 30 min on ice
4. Isolate nuclear proteins
   (a) Centrifuge for 20 min at 13,000 rpm and 4 °C
   (b) Transfer supernatant to new microcentrifuge tube and add the same volume of NEB C as the supernatant has
   (c) Aliquote 100 – 200 µl-wise and quick-freeze on dry ice
   (d) Store aliquots at –80 °C (long term) or –20 °C (usage shortly after isolation)
5. Quantify amount of isolated protein
   (a) Thaw one aliquot of each sample on ice
   (b) Clean nanodrop sample loading area with Aqua ad iniectabilia and disposable paper
   (c) Start protein quantitation program and zero device with 2 µl of an 1:1 mix of NEB B : NEB C
   (d) Load 2 µl of each sample and record concentration of proteins in the sample

**Design of Binding Site Containing Oligonucleotides** was done by taking 21 bp of reference sequence upstream of the *PLP2* translation start, ranging from –202 to –182. In the $PLP2^{-188(A)}$ oligonucleotides, cytosine was exchanged by an adenine. Both oligonucleotides were designed to contain the entire binding sequence of the transcription factors REL and ETS1, as defined by the MatInspector database [66], plus three nucleotides of the *PLP2* sequence at both the 5' and 3' end. For the binding sites of both factors are shifted by only one base, this could be achieved without problems. The resulting oligonucleotides are shown in the following table.

| *Oligonucleotide* | *Sequence* |
|---|---|
| $PLP2^{-188(C)}$ | 5'-ggcgagggggcttcCggggcg-3' |
| $PLP2^{-188(A)}$ | 5'-ggcgagggggcttcAggggcg-3' |

Both oligonucleotides were ordered with fluorescent labeling at MWG Biotech AG (Ebersberg, Germany), together with their single-stranded complement, such that double-stranded oligonucleotides could be obtained after annealing both single-strands in a thermal cycler (see table 6). For specific competition of the protein binding to the FITC-labeled oligonucleotides, unlabeled oligonucleotides of the same sequence were used.

Table 6: **Sequence of fluorescently labeled oligonucleotides used in A.L.F. EMSA**

| *Oligonucleotide* | *Sequence (5'–3')* |
|---|---|
| Version 1: | |
| $PLP2^{-188(C)}$ oligonucleotide, forward, FITC | FITC–ggcgagggggcttccggggcg |
| $PLP2^{-188(C)}$ oligonucleotide, forward | ggcgagggggcttccggggcg |
| $PLP2^{-188(C)}$ oligonucleotide, reverse | cgccccggaagccccctcgcc |
| $PLP2^{-188(A)}$ oligonucleotide, forward, FITC | FITC–ggcgagggggcttcaggggcg |
| $PLP2^{-188(A)}$ oligonucleotide, forward | ggcgagggggcttcaggggcg |
| $PLP2^{-188(A)}$ oligonucleotide, reverse | cgcccctgaagccccctcgcc |
| Version 2: | |
| $PLP2^{-188(C)}$ oligonucleotide, forward, FITC | gggcttccggggcggcagcacctggcccggctccggcc–FITC |
| $PLP2^{-188(C)}$ oligonucleotide, forward | gggcttccggggcggcagcacctggcccggctccggcc |
| $PLP2^{-188(C)}$ oligonucleotide, reverse | ggccggagccgggccaggtgctgccgccccggaagccc |
| $PLP2^{-188(A)}$ oligonucleotide, forward, FITC | gggcttcaggggcggcagcacctggcccggctccggcc–FITC |
| $PLP2^{-188(A)}$ oligonucleotide, forward | gggcttcaggggcggcagcacctggcccggctccggcc |
| $PLP2^{-188(A)}$ oligonucleotide, reverse | ggccggagccgggccaggtgctgccgcccctgaagccc |

**Annealing of Single Stranded Oligonucleotides** produced those double-stranded oligonucleotides that were subsequently used for protein binding. Annealing was done according to the following protocol.

1. Prepare annealing reaction in microcentrifuge tube
   (a) 6 µl forward oligonucleotide (FITC-labeled, 100 pmol/µl, 660 ng/µl)
   (b) + 6 µl reverse oligonucleotide (unlabeled, 100 pmol/µl, 570 ng/µl)
   (c) + 2 µl 10x annealing buffer
   (d) + 6 µl Aqua ad iniectabilia
2. Execute annealing program in thermocycler
   (a) 95 °C for 3 min
   (b) 60 °C for 15 min
   (c) 37 °C for 15 min
   (d) 25 °C for 30 min
3. Dilute double stranded oligonucleotide (30 pmol/µl, 170 ng/µl) 20 fold to 1.5 pmol/µl
   (a) 20 µl annealing reaction
   (b) + 342 µl Aqua ad iniectabilia
   (c) + 38 µl 10x annealing buffer
4. Aliquote 40 µl-wise and store at –20 °C

**Preparation of a Polyacrylamide Gel** was done according to the following protocol.

1. Preparation of a polyacrylamide gel
   (a) Clean gel touching sides of cooling plate (connections for water circulation on the upper right) and glass plate (inscription must not be inverted), all spacers and the light coupler according to following instructions:
      i. Brush with dishwashing detergent, flush detergent and waste away with tap water
      ii. Repeat two times
      iii. Wipe 3x with ethanol and 3x with distilled water, using disposable paper
   (b) Mount cooling plate to glass plate:
      i. Place cooling plate, connections for water circulation on the upper right, on a cube, leaving a free margin to be able to apply the screw clamps later on
      ii. Align long spacer to the left margin, flushing with the upper edge
      iii. Align medium long spacer, light coupler and short spacer up to down at the right margin, also flushing with the upper edge (light coupler is now at correct position to allow the laser to enter the gel)
      iv. Place glass plate on top, inscription inverted
      v. Mount clamp screws at the positions defined by the inscription on the glass plate
   (c) Prepare acrylamide solution in a beaker (see table 25)
   (d) Homogenize solution by pipetting up and down with a syringe
   (e) Pour gel with a syringe between the still horizontally lying glass and cooling plates; tip with fingers on the plates to avoid air bubbles
   (f) Insert comb at the upper edge of the plate sandwich, aligning the teeth with the inscription of the glass plate
   (g) Let gel polymerize for at least one hour
2. Finish mounting of A.L.F. parts
   (a) Place lower buffer chamber to A.L.F device and connect electrode
   (b) Fill in sufficient 1x TBE buffer and install plate sandwich, using the fixation screws in the device (buffer level has to be high enough to reach the gel)
   (c) Insert upper buffer chamber, attach it to the plate sandwich and connect the electrode
   (d) Fill in sufficient 1x TBE buffer to cover the gel slots

(e) Adjust screws for plate sandwich carriers in the lower buffer chamber to let the laser beam enter the gel via the middle of the light coupler and to let it pass the gel exactly horizontally

(f) Move plate sandwich horizontally, if the gel lanes do not fit exactly with the detectors

(g) Adjust laser with laser screw to shine brightly through the gel

(h) Start water circulation by external thermostat

(i) Prerun the gel at 10 W for 30 min, setting parameters with ALFWin Software

**The Binding Reaction** and its composition is mentioned below. The core reagents are denoted in normal letters, whereas supplemental components are indicated in *italics* and are added in dependence of the necessity of competing protein binding specifically or of producing a super-shift with an antibody.

1. Binding reaction

   (a) 1x binding buffer

   (b) 2 µg poly-dIdC

   (c) *x fold specific competitor*

   (d) 3 µg nuclear protein

   (e) *2.5 µg antibody*

   (f) 1.5 pmol oligonucleotide to be bound

   (g) Fill up to 50 µl with Aqua ad iniectabilia

2. Incubate for 30 min at 37 °C in thermomixer

3. Start Run:

   (a) Add EMSA loading buffer to samples and mix by pipetting up and down

   (b) Pause prerun of gel

   (c) Load samples to each second gel lane, using flat pipet tips

   (d) Continue run and stop, when bound oligonucleotide has passed detector

4. End run:

   (a) Disassemble plate sandwich, dispose gel and clean up as described above

   (b) Save data file and analyze with Allele Locator software

**2.3.2.2 Fluorescent EMSA in Agarose Gels** only differ from the gel shift assays of the previous section in the use of agarose gels for separation instead of polyacrylamide gels. Two types of oligonucleotide labeling were investigated: covalent modification of the oligonucleotide with FITC and non-covalent modification with SYBR Green.

SYBR Green was chosen instead of ethidium bromide because of its higher sensitivity. Using the Fuji FLA scanner, as little as 5 pg DNA can be detected using SYBR Green [26], whereas 194 pg are the detection limit when ethidium bromide is used [27], revealing a remarkable increase in sensitivity that is greater than 38 times. Of course, post-staining is also not meaningful with SYBR Green. Thus, I invented a pre-staining and purification procedure explained below. Purification of the labeled oligonucleotide is necessary, because it is very unlikely that the entire amount of SYBR Green in the staining solution will bind to the oligonucleotide. If the stained oligonucleotide is then added to the EMSA reaction, unbound SYBR Green will be able to bind to the specific competitor and thus will add to the fluorescence signal of the free oligonucleotide.

The protocol for the fluorescent agarose gel EMSAs can be found below.

1. Labeling of oligonucleotides
   - FITC-labeling:
     (a) No labeling procedure was necessary, for oligonucleotides were delivered fluorescently (see table 6)
   - SYBR Green-labeling:
     (a) Prepare a 50x concentrated solution of SYBR Green in TBE buffer
     (b) Dilute 12 µl annealed, but not FITC-labeled oligonucleotide (see table 6) in 188 µl Aqua ad iniectabilia to 10.2 ng/µl (dilution factor 16.7)
     (c) Add 4 µl of 50x concentrated SYBR Green solution to 196 µl oligonucleotide dilution (final oligonucleotide concentration: 10 ng/µl)
     (d) Purify labeled oligonucleotides from unbound label in solution by gel filtration chromatography, as described in 2.1.3.5
2. Gel preparation
   (a) Prepare a solution of 1.5% agarose in 1x TBE buffer
   (b) Boil solution in a microwave oven and stirr at RT until solution has cooled down to $\approx$ 70 °C
   (c) Pour gel in tray, place comb and let gel set for 1 h
   (d) Fill electrophoresis chamber with 1x TBE buffer
3. Prepare binding reaction (see section 2.3.2.1)
4. Electrophoresis
   (a) Add EMSA gel loading buffer to samples

(b) Remove comb from gel and place gel tray in electrophoresis chamber

(c) Load 20 µl of each sample to the gel and leave a few lanes empty

(d) Load blue gel loading buffer to the free lanes to see migration during electrophoresis

(e) Apply a voltage not higher than 5 V per cm distance of the electrodes

(f) Let run until blue loading buffer has reached last third of the gel

5. Analysis

(a) Transfer gel to gel scanner and start scanning at 100 µm accuracy, using the blue laser at 473 nm and the LPB filter

(b) Save TIFF file with AIDA Software for further analysis

**2.3.2.3 Radioactive EMSA** was done according to the following protocol.

1. Annealing of single-stranded oligonucleotides
   (a) Anneal oligonucleotides according to section 2.3.2.1, but omit last diluting steps
2. Labeling of oligonucleotides
   (a) Prepare on ice:
      i. 36.5 µl Aqua ad iniectabilia
      ii. + 5 µl 10x reaction buffer for Klenow fragment
      iii. + 3 µl annealed oligonucleotide (≈ 170 ng/µl)
      iv. + 1 µl dTTP, dGTP, dATP each (each 2.5 mM)
      v. + 0.5 µl $\alpha^{32}$PdCTP (3.3 nM)
      vi. + 2 µl Klenow fragment
   (b) Incubate at 37 °C for 1 h
   (c) Store at 4 °C until use and use up within six weeks (half-life of 32P is 14.3 d [79])
3. Preparation of polyacrylamide gel
   (a) Clean both glass plates, spacers and comb according to section 2.3.2.1
   (b) Assemble spacers and place between both glass plates
   (c) Hold both plates together by clamps and tilt plate sandwich to ≈ 45°
   (d) Mix all ingredients for a polyacrylamide gel as described in 2.3.2.1
   (e) Fill in gel mix with a syringe, avoiding air bubbles to form
   (f) Insert comb and let gel polymerize for at least one hour
4. Preparation of the binding reaction
   (a) Prepare on ice:
      i. 1x binding buffer
      ii. 2 µg poly-dIdC
      iii. *x fold specific competitor*
      iv. 3 µg nuclear protein
      v. *2.5 µg antibody*
      vi. 20 ng labeled oligonucleotide
      vii. Fill up to 50 µl with Aqua ad iniectabilia
   (b) Incubate for 30 min at 37 °C in thermomixer
5. Polyacrylamide gel electrophoresis
   (a) Prerun of gel:

i. Fill lower buffer tank of electrophoresis chamber with 1x TBE buffer
ii. Insert glass sandwich after removing comb and lower horizontal spacer
iii. Fill upper buffer tank and rinse slots with syringe
iv. Run gel for 30 min at no more than 8 V per cm gel length to avoid significant heating

(b) Electrophoresis:

i. Load samples with EMSA loading buffer asymmetrically to the gel and leave free lanes
ii. Load blue loading buffer to the free lanes to observe electrophoretic migration
iii. Apply no more than 8 V per cm gel length
iv. Stop run when blue dye has reached last third of the gel

6. Gel drying

(a) Disassemble electrophoresis unit
(b) Seperate both glass plates from each other, carefully preventing the gel from breaking
(c) Take up gel with Whatman paper of appropriate size
(d) Cover the side opposed to the Whatman paper side with clingwrap and transfer this sandwich to the gel dryer
(e) Run drying program for at least one hour
(f) Clean electrophoresis unit, glass plates, spacers and comb with tap water, disposing electrophoresis buffer into radioactive waste, if necessary

7. Analysis of EMSA

(a) Take dried gel out of dryer and place it in PhosphorImager plate cassette
(b) Expose over night or longer and scan on PhosphorImager at an accuracy of 100 µm
(c) Erase screen with white light and save output file for further analysis
(d) Expose storage phosphor screen for a longer time, if necessary, otherwise discard membrane in radioactive waste

## 2.4 Materials and Devices

In tables 7 and 2.4, all substances which were directly purchased are listed alphabetically. From these substances, all buffers and solutions were prepared, composed as documented in the tables starting on page 52. These tables were grouped according to the experiments the particular solution was used in, and in each group, the tables are sorted according to the order in which the solutions were used. Tables 8, 9 and 11 contain lists of all devices and disposable materials that were used.

Table 7: **Liquid and solid substances**

| *Substance name* | *Solid substance: Molar mass* [g/mol] | *Liquid substance: Concentration* | *Manufacturer* |
|---|---|---|---|
| Acetic acid, glacial | | $\geqq$ 99% | Merck |
| Acrylamide/Bisacrylamide (37.5:1) | | 40% | Applichem |
| $\beta$-Actin cDNA Hybridization Probe | | n.s. | Biocat |
| Agarose | n.s. | | Invitrogen |
| Agarose gel loading buffer, blue | | 6x | Fermentas |
| Ammonium persulfate (APS) | 228.2 | | Applichem |
| Antibody against ETS1 (rabbit, polyclonal) | | 0.5 µg/µl | Active Motif |
| Antibody against REL (rabbit, polyclonal) | | 0.5 µg/µl | AbCam |
| Antibody against ZFP42 (rabbit, polyclonal) | | 0.25 µg/µl | Abgent |
| Aqua ad iniectablilia | | | Baxter |
| Boric acid | 61.83 | | Merck |
| Buffer for Klenow Fragment | | 10x | New England Biolabs |
| Diethyl pyrocarbonate (DEPC) | | $\geqq$ 97% | Sigma Aldrich |
| Di-sodium ethylenediaminetetraacetic acid ($Na_2$-EDTA·$2H_2O$, Titriplex III (Merck)) | 372.24 | | Merck |
| Di-sodium hydrogen phosphate ($Na_2HPO_4$) | 141.96 | | Merck |
| $\alpha^{32}$PdCTP, Redivue | | 370 MBeq/ml | Amersham |
| Deoxynucleoside Triphosphate Set, PCR Grade | | 100 mM (each) | Roche |
| Dodecyl sulfate sodium salt (SDS) | 288.38 | | Merck |
| DNA ladder pUC Mix Marker 8 | | 0.5 µg / µl | Fermentas |
| DNA ladder HyperLadder I | | n.s. | Bioline |
| DNA ladder HyperLadder IV | | n.s. | Bioline |
| Ethanol | | absolute | Merck |
| Ethidium bromide solution | | 10 mg/ml | Sigma Aldrich |
| Ethylenediaminetetraacetic acid (EDTA, Titriplex II (Merck)) | 292.25 | | Merck |
| Formaldehyde | | 37% | Merck |
| Formamide | | $\geqq$ 99.5% | Merck |

| *Substance name* | *Solid substance: Molar mass* [g/mol] | *Liquid substance: Concentration* | *Manufacturer* |
|---|---|---|---|
| Glycerol | | 87% | Merck |
| 2-[4-(2-Hydroxyethyl)-1-piperazinyl]-ethanesulfonic acid (HEPES) | 238.30 | | Merck |
| Isopropanol | | $\geqq$ 99.7% | Merck |
| Klenow Fragment (3'-5' exo–) | | 5,000 U/ml | New England Biolabs |
| $\beta$-Mercapto ethanol | | $\geqq$ 99% | Applichem |
| Magnesium chloride ($MgCl_2$) | 95.22 | | Merck |
| 3-Morpholinopropane sulfonic acid (MOPS) | 209.26 | | Merck |
| Nonylphenylpolyethylenglycol (Nonidet P40, NP40) | | $\geqq$ 99% | Applichem |
| Nuclear extract of human adult brain proteins | | 1 µg/µl | BioChain Institute |
| Nuclear extract of human fetal brain proteins | | 1 µg/µl | BioChain Institute |
| Poly-deoxy-inosinic-deoxy-cytidylic acid (Poly-dIdC) | n.s. | | Roche |
| Potassium chloride | 74.55 | | Merck |
| Protease inhibitor cocktail tablets, EDTA free, complete | | | Roche |
| Random control DNA panel HRC-1 and HRC-2, human | | 8 ng/µl | ECACC/Sigma |
| RNA ladder 0,24 – 9,5 kB | | 1 µg/µl | Invitrogen |
| Sodium chloride (NaCl) | 58.44 | | Merck |
| Sodium dihydrogen phosphate ($NaH_2PO_4$) | 119.98 | | Merck |
| SYBR Green I | | 10,000x | Biozym |
| *Taq* DNA Polymerase Kit | | 4 x 250 units *Taq* DNA Polymerase, 10x PCR Buffer, 5x Q-Solution, 25 mM $MgCl_2$ | Qiagen |
| Tetramethylethylendiamin (TEMED) | | $\geqq$ 99% | Applichem |
| Tris(hydroxymethyl)aminomethane (Tris buffer) | 121.14 | | Merck |
| Tris(hydroxymethyl)aminomethane hydrochloride (Tris-HCl) | 157.60 | | Merck |
| Tri-sodium citrate | 258.07 | | Merck |
| TRIzol (= phenol + guanidine isothiocyanate) | | n.s. | Invitrogen |
| ULTRAHyb Northern hybridization buffer | | n.s. | Ambion |

Table 8: **Devices used in particular experiments**

| *Device* | *Manufacturer* |
|---|---|
| **Agarose gel electrophoresis:** | |
| Chambers: Horizon series | Life Technologies |
| Gel documentation system: | |
| Camera: E.A.S.Y. 440K | Herolab |
| UV transilluminator: UVT 28 ME | Herolab |
| Light shield: RH 5.1 Darkroom Hood | Herolab |
| Printer: CP770DW | Mitsubishi |
| **Polyacrylamide gel electrophoresis:** | |
| Chamber: V15 · 17 | Life Technologies |
| Gel drying: | |
| Dryer Modul 583 | Bio-Rad |
| Vapor trap | Bio-Rad |
| Vacuum pump | Bio-Rad |
| **Northern blotting:** | |
| Film developing machine Curix 60 | Agfa |
| Foil heat exchanger Sealboy 320 | Audion Elektro |
| Mini Hybridisation Oven | Appligene |
| UV Stratalinker 1800 | Stratagene |
| **Mutation frequency screening:** | |
| Sequencer Genetic Analyzer 3100 | Applied Biosystems |
| DHPLC: | |
| Device: WAVE Nucleic Acid Fragment Analysis System | Transgenomic |
| Column: DNASep HT Cartridge | Transgenomic |
| **EMSA** | |
| **A.L.F. EMSA:** | |
| A.L.F. DNA Sequencer | Amersham |
| Thermostat F25+HC | Julabo |
| Notebook Z-Note 1000 | Zenith Data Systems |
| **Agarose gel EMSA:** | |
| Gel fluorescence scanner FLA-5100 | Fujifilm |

Table 9: **Devices used in multiple experiments**

| *Device* | *Manufacturer* |
|---|---|
| Balance BP 2100 S | Sartorius |
| G.M.-counter LB 1210 B (for flat surfaces) | berthold |
| G.M.-counter mini-monitor, series 900 (for spots) | Wellhöfer |
| Magnetic stirrer MR 1000 | Heidolph |
| Microcentrifuge EBA 12 and EBA 12R (cooled) | Hettich |
| Microwave oven "Quarzgrill 9023G" | privileg |
| Nanodrop ND-1000 Spectrophotometer | NanoDrop Technologies |
| PCR device DNA Engine Tetrad | MJ Research |
| Pestle "L" (for cell homogenization) | B.Braun |
| PhosphorImager Image Eraser | Molecular Dynamics |
| PhosphorImager scanner Storm 820 | Molecular Dynamics |
| PhosphorImager screen and cassette | Molecular Dynamics |
| Pipets, single-channel: research series | Eppendorf |
| Pipets, single-channel: pipetman | Gilson |
| Pipets, multi-channel: Pipet-Lite LTS 12-channel, 2 – 20 µl | Rainin |
| Power supply for electrophoresis: PheroStab 0652 | Biotec Fischer |
| Thermomixer 3831 and 5426 | Eppendorf |

Table 10: **Software applications**

| *Software* | *Manufacturer* |
|---|---|
| AIDA Imager Analyzer (for Fuji FLA scanner) | raytest |
| ALFWin Software (for A.L.F. DNA sequencer) | Amersham Pharmacia Biotech |
| Allele Locator (for A.L.F. DNA sequencer) | Amersham Pharmacia Biotech |
| Array Gauge and Multi Gauge (for FLA 5100) | Fujifilm |
| CodonCode Aligner (for sequence evaluation) | CodonCode Corporation |
| E.A.S.Y. Win32 (for agarose gel documentation) | HeroLab |
| ImageQuant (for storage phosphor screen evaluation) | Molecular Dynamics |
| Scan control software (for PhosphorImager) | Molecular Dynamics |

Table 11: **Disposable material used**

| *Material* | *Manufacturer* |
|---|---|
| Cling-wrap | Frappan |
| Dishwashing detergent "zack" | August Wencke |
| Disposable gloves safety + comfort | LATECH |
| Disposable paper KIMTECH Science | Kimberly Clark |
| Disposable paper astrein AS67 | igefa |
| Disposable towels Kleenex Ultra | Kimberly Clark |
| Flexible plates for gel loading, 96 well, U-bottom without lid | Becton Dickinson |
| Hybond XL nylon membrane | Amersham Biosciences |
| Hybridization tube GL 45 | Schott |
| Hypercassette film cassette | Amersham Biosciences |
| Microcentrifuge tube Mμlti 1.7 ml | Roth |
| Mini column MicroSpin G-50 | Amersham Biosciences |
| Parafilm | American National Can |
| PCR plates Thermo-Fast 96, Non-Skirted | ABgene |
| PCR plate adhesive film | ABgene |
| PCR plate for DHPLC analysis "Costar Thermowell" | Corning |
| Pipet tips SAFESEAL Premium | Biozym |
| Pipet tips, flat MμlTiFlex (E), 2 mm | Sorenson BioScience |
| Pipet tips, multichannel: Spacersaver LTS 20 µl Tip 960/10 | Rainin |
| Plastic film | Severin |
| Polypropylen tubes Cellstar, sterile, 15 and 50 ml | Greiner |
| Syringe "Plastipak" | Becton Dickinson |
| Whatman 3MM Chr paper | Whatman |
| X-Ray Film 100 NIF, 18 cm x 24 cm | Fujifilm |
| X-Ray Film Biomax MS, 18 cm x 24 cm | Kodak |

### 2.4.1 Solutions for Northern Blotting

Table 12: **Composition of 10x MOPS buffer**

| *Substance* | *Mass substance for 500 ml 10x buffer* [g] | *Concentration 10x buffer* [M] |
|---|---|---|
| 3-Morpholinopropane sulfonic acid | 41.85 | 0.4 |
| Sodium acetate | 4.1 | 0.1 |
| Ethylenediaminetetraacetic acid | 1.86 | 0.01 |
| DEPC-treated, bidistilled water | fill up to 500 ml | |

Table 13: **Composition of 2x RNA loading buffer**

| *Substance* | *Volume substance for 5 ml 2x buffer* [µl] | *Concentration 2x buffer* |
|---|---|---|
| Formamide | 2500 | 50 % v/v |
| Formaldehyde | 500 | 1.3 M |
| Ethidium bromide | 20 | 40 µg/ml |
| DEPC-treated, bidistilled water | 1980 | |

Table 14: **Composition of 5x Oligonucleotide Labeling Buffer (OLB)**

| *Volume needed for 20 ml 5x OLB solution* | *Composition stem solutions* | *Composition of additional solutions* |
|---|---|---|
| 5 ml solution A | Solution A: 5 ml solution O<br>+ 50 µl of each dATP, dGTP, dTTP (each 100 mM)<br>+ 90 µl $\beta$-mercapto ethanol | Solution O:<br>1.25 mM Tris-HCl (pH=8)<br>+ 125 mM $MgCl_2 \cdot 6H_2O$ |
| 12.5 ml solution B | 2 M HEPES (pH=6.6) | |
| 7.5 ml solution C | 50 OD random hexamers in 7.5 ml Tris-EDTA (pH=7.5) | |

Table 15: **Composition of Northern blot washing buffers**

| *Substance* | *Volume substance needed for 1000 ml buffer* | *Concentration in buffer* |
|---|---|---|
| Washing buffer 1 | 100 ml SSC (20x), 10 ml SDS (10%), 890 ml bidistilled water | 2x SSC, 0.1% SDS |
| Washing buffer 2 | 10 ml SSC (20x), 10 ml SDS (10%), 980 ml bidistilled water | 0.2x SSC, 0.1% SDS |

Table 16: **Composition of 20x SSC buffer**

| *Substance* | *Amount substance needed for 500 ml buffer* [g] | *Concentration in buffer* [M] |
|---|---|---|
| Tri-sodium citrate | 38.7 | 0.3 |
| Sodium chloride | 87.7 | 3 |
| Bidistilled water | fill up to 500 ml | |

Table 17: **Composition of Northern blot strip off buffer**

| *Substance* | *Amount substance needed for 500 ml buffer* [ml] | *Concentration in buffer* |
|---|---|---|
| Dodecyl sulfate sodium salt | 5 | 0.1% |
| SSC buffer | 1 | 0.04x |
| Bidistilled water | 494 | |

### 2.4.2 Solutions for PCR

Table 18: **Composition of 50x Tris Acetic acid EDTA (TAE) buffer**

| *Substance* | *Amount substance needed for 1000 ml 50x buffer* [g] | *Concentration in buffer* [M] |
|---|---|---|
| $Na_2EDTA \cdot 2H_2O$ | 18.61 | 0.05 |
| Acetic acid | 60.05 | 1 |
| Tris buffer | 242.28 | 2 |
| Bidistilled water | fill up to 1000 ml | |

### 2.4.3 Solutions for EMSA

Table 19: **Composition of 10x Phosphate Buffered Saline (PBS) buffer**

| *Substance* | *Amount substance needed for 1000 ml 10x buffer* [g] | *Concentration in buffer* [mM] |
|---|---|---|
| NaCl | 90 | 1,540 |
| $NaH_2PO_4$ | 3.2 | 90.8 |
| $Na_2HPO_4$ | 10.9 | 76.8 |
| Bidistilled water | fill up to 1000 ml | |

Table 20: **Composition of nuclear extraction buffer A (NEB A).** As protease inhibition, 2 tablets of complete protease inhibitor cocktail were added to 100 ml of buffer.

| *Substance* | *Concentration stock* [M] | *Volume for 100 ml NEB A* [ml] | *Concentration NEB A* [mM] |
|---|---|---|---|
| HEPES (pH=7.9) | 1 | 1 | 10 |
| KCl | 1 | 1 | 10 |
| $MgCl_2$ | 1 | 0.150 | 1.5 |
| Bidistilled water | | 97.85 | |

Table 21: **Composition of nuclear extraction buffer B (NEB B).** As protease inhibition, 2 tablets of complete protease inhibitor cocktail were added to 100 ml of buffer.

| *Substance* | *Concentration stock* [M] | *Volume for 100 ml NEB B* [ml] | *Concentration NEB B* [mM] |
|---|---|---|---|
| HEPES (pH=7.9) | 1 | 2 | 20 |
| Glycerol | 87% | 11.49 | 10% |
| NaCl | 5 | 8.4 | 420 |
| $MgCl_2$ | 1 | 0.150 | 1.5 |
| EDTA | 0.5 | 0.04 | 0.2 |
| Bidistilled water | | 77.92 | |

Table 22: **Composition of nuclear extraction buffer C (NEB C).** As protease inhibition, 2 tablets of complete protease inhibitor cocktail were added to 100 ml of buffer.

| *Substance* | *Concentration stock* [M] | *Volume for 100 ml NEB C* [ml] | *Concentration NEB C* [mM] |
|---|---|---|---|
| HEPES pH=7.9 | 1 | 2 | 20 |
| Glycerol | 87% | 34.48 | 30% |
| $MgCl_2$ | 1 | 0.150 | 1.5 |
| EDTA | 0.5 | 0.04 | 0.2 |
| Bidistilled water | | 63.33 | |

Table 23: **Composition of 10x oligonucleotide annealing buffer**

| *Substance* | *Concentration stock* [mM] | *Volume stock for 1 ml 10x buffer* [µl] | *Concentration 10x buffer* [mM] |
|---|---|---|---|
| Tris-HCl (pH=8) | 500 | 200 | 100 |
| NaCl | 2500 | 400 | 1000 |
| $MgCl_2$ | 20 | 1000 | 20 |
| EDTA | 250 | 40 | 10 |
| Bidistilled water | | 340 | |

Table 24: **Composition of 4x binding buffer**

| *Substance* | *Concentration stock solution* [mM] | *Concentration 4x buffer* [mM] | *Volume stock for 10 ml 4x buffer* [µl] |
|---|---|---|---|
| HEPES (pH=7.5) | 1000 | 80 | 800 |
| Nonidet P40 | 1% | 0.04% | 400 |
| Glycerol | 87% | 20% | 2299 |
| $MgCl_2$ | 1000 | 10 | 100 |
| Bidistilled water | | | 6401 |

Table 25: **Composition of a 4% polyacrylamide gel**

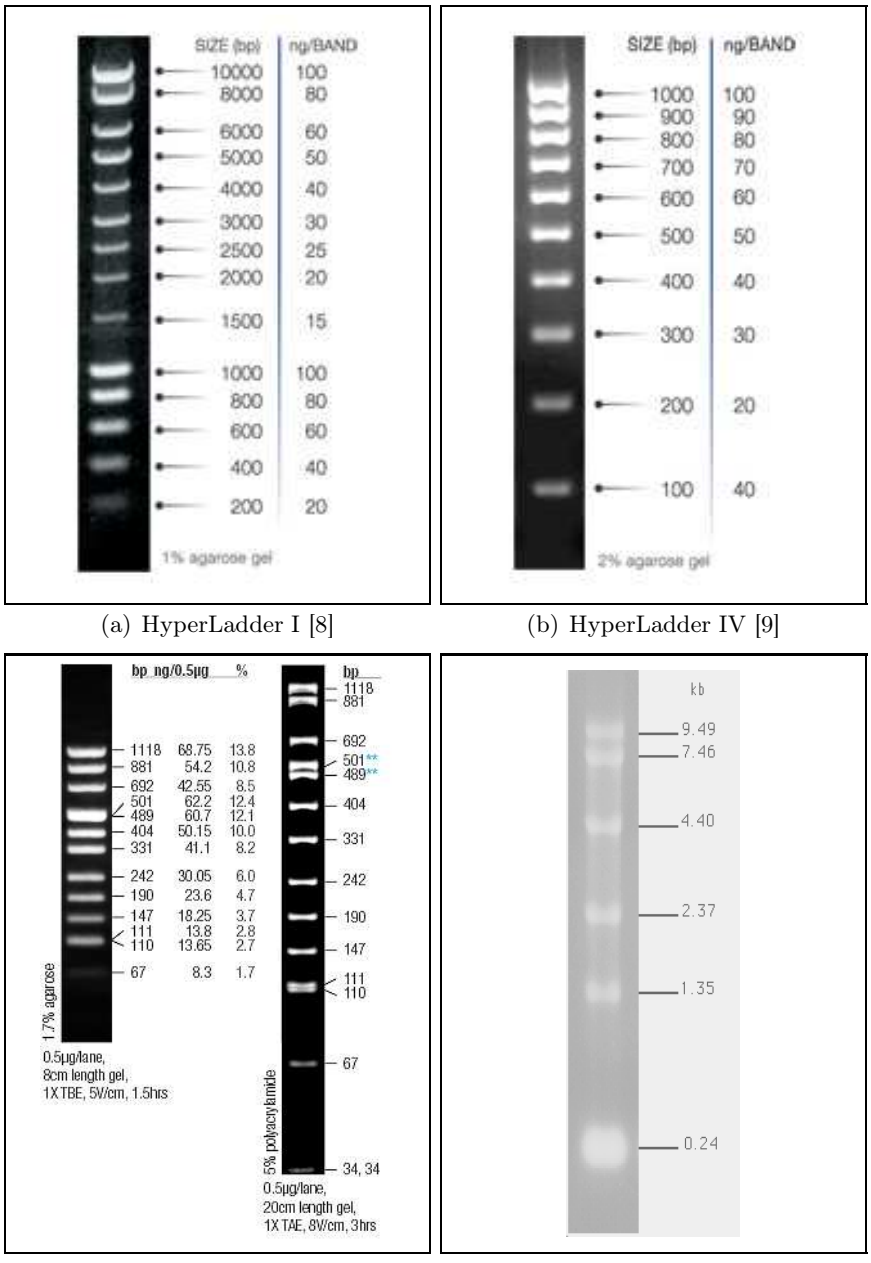
| *Substance* | *Concentration stock* | *Volume stock for 40 ml gel* [ml] | *Concentration in the gel* |
|---|---|---|---|
| TBE | 5x | 8 | 1x |
| Acrylamide:Bisacrylamide (37.5:1) | 40% | 3.5 | 3.5% |
| Glycerol | 87% | 2.3 | 5% |
| Bidistilled water | | 25.96 | |
| TEMED | 100% | 0.12 | 0.3% |
| APS | 10% | 0.12 | 0.03% |

Table 26: **Composition of 10x EMSA loading buffer**

| *Substance* | *Concentration stock* [M] | *Volume stock for 1 ml 10x buffer* [µl] | *Concentration 10x buffer* [mM] |
|---|---|---|---|
| Tris-HCl (pH=7.5) | 0.5 | 500 | 250 |
| Glycerol | 87% | 500 | 43.5% |

Table 27: **Composition of 10x Tris Boric acid EDTA (TBE) buffer**

| *Substance* | *Used mass per liter* [g] | *Final concentration* [M] |
|---|---|---|
| Boric acid | 55.03 | 0.89 |
| $Na_2$-EDTA $\cdot$ $2H_2O$ | 7.44 | 0.20 |
| Tris base | 107.81 | 0.89 |
| Bidistilled water | fill up to 1000 ml | |




(a) HyperLadder I [8] (b) HyperLadder IV [9]

(c) pUC Mix Marker 8 [23] (d) RNA Ladder, 0.24 – 9.5 kb [39]

Figure 17: **Sizes and masses of all ladders used**

# 3 Results

In this diploma thesis, I studied functional aspects of a cytosine to adenin substitution that was found 188 bp upstream of the translation start site of the *PLP2* mRNA. The transversion appeared in mental retardation patients, but not in controls with above average IQ.

To find out about an influence of the promoter variation on the expression level of the gene, I performed Northern blots with total RNA from patients exhibiting the $PLP2^{-188(A)}$ genotype and unaffected controls with $PLP2^{-188(C)}$. In parallel, three control panels were screened for the A-variant: one commercially available control DNA collection and one derived from blood donors, both with uncharacterized mental retardation status, and a control collection comprising DNA from individuals with above average IQ.

As I found an obvious decrease of the *PLP2* expression level in the patients, I finally tried to find the trancription factors involved, using the technique of Electrophoretic Mobility Shift Assay *in silico* (with the help of several databases) and *in vitro* (in three different variants of the technique).

## 3.1 Expression Level of the *PLP2* Gene

Expression levels of the *PLP2* gene in lymphoblastoid cell lines from patients in comparison to healthy controls were investigated by Northern blotting. At first, total RNA was isolated from the lymphoblastoid cell lines (see figure 18, for detailed quantitation data see appendix on page 86).

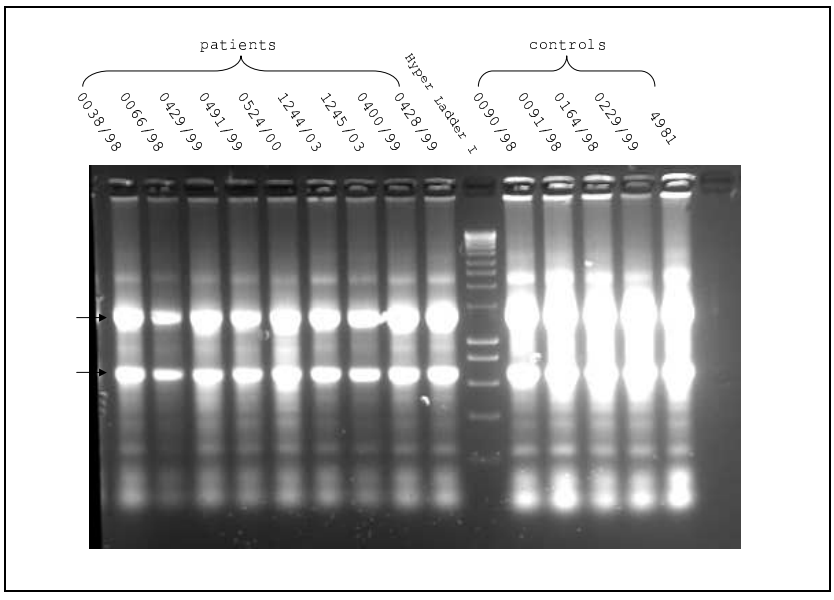
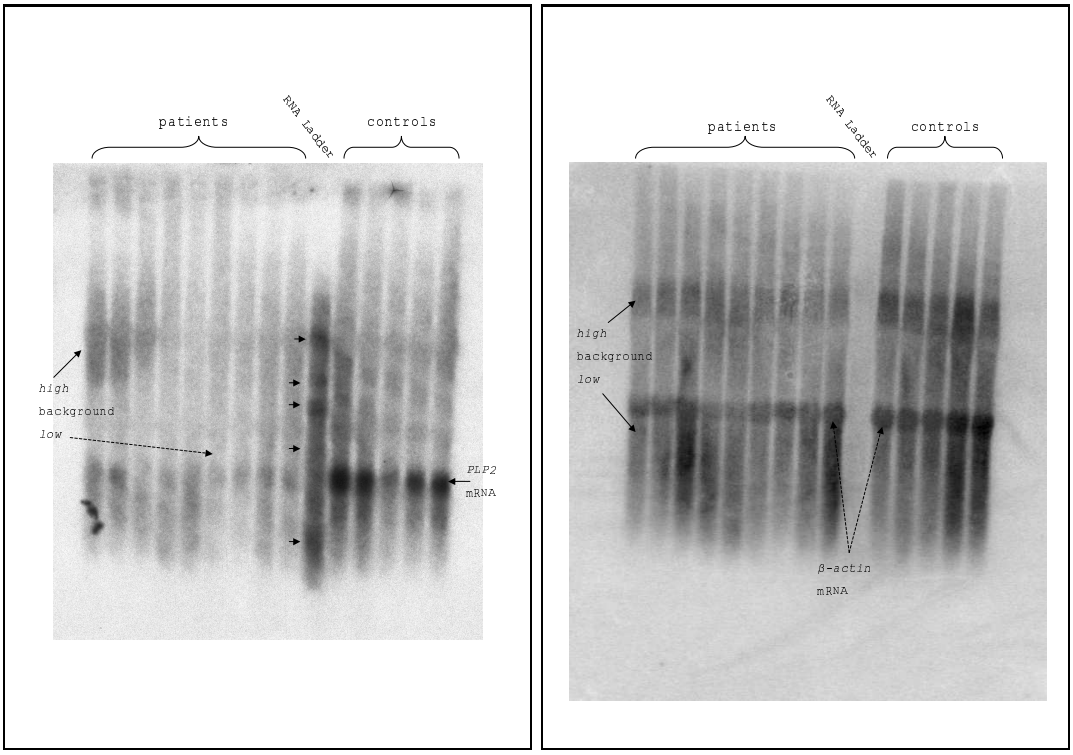



Figure 18: **Non-denaturating agarose gel electrophoresis of isolated total RNA from lymphoblastoid cell lines.** Both ribosomal bands, corresponding to the 18 S and 28 S rRNA (bound to the 5.8 S rRNA), appear with highest intensity (see arrows), whereas all differently sized mRNAs and tRNAs can be found allover the lanes. The stronger RNA signal in the controls was due to uneven UV gel illumination in the documentation system. A DNA ladder was chosen to separate the patient RNA from the control RNA.

All RNAs were subjected to denaturing agarose gel electrophoresis and then transferred to a nylon membrane. In order to quantify *PLP2* expression levels, the blots were first hybridized with a *PLP2*-specific probe, revealing absolute *PLP2* expression levels. For normalization of the obtained expression data, the blots were subsequently treated with a $\beta$-actin-specific probe. $\beta$-actin is a cellular housekeeping gene which is expressed ubiquituously and at constant levels, therefore being suitable as a normalization reference.

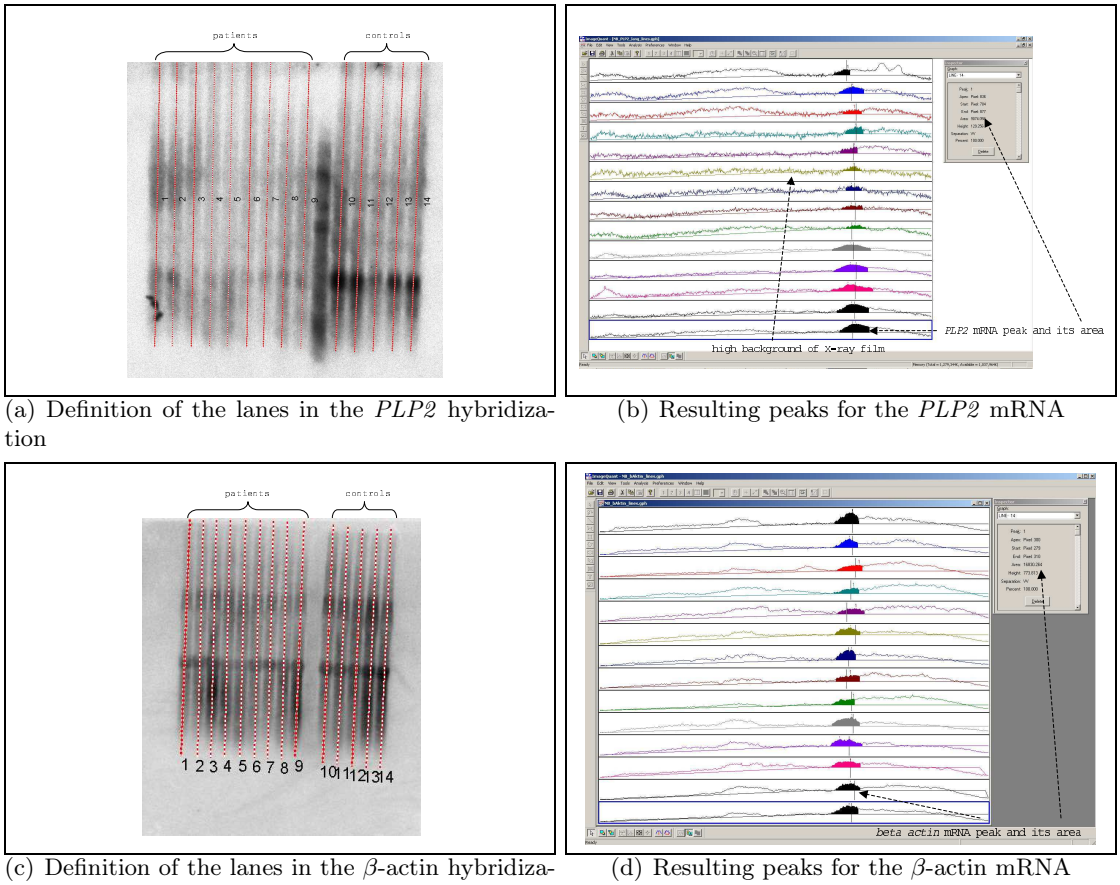



(a) Hybridization of the *PLP2* probe (b) Hybridization of the *β-actin* probe

Figure 19: **Northern hybridization of the *PLP2* probe and the *β-actin* probe.** The *PLP2* mRNA appears as a band of approximately 1 kb size (see arrow in figure 19(a)), slightly below the 1.35 kb band of the RNA ladder, but largely above it's 0.24 kb band (ladder bands denoted by arrowheads). The *PLP2* mRNA size is equal or larger than 945 bases (see NCBI Nucleotide [64] for NM_002668 [5]), depending on the length of the poly-A tail which is obligatory for the 3' end of all eucaryotic mRNAs. In the same manner, one can find the size of the *β-actin* mRNA to be 1793 bases (see [64] for NM_001101 [5]), which also fits to the location of the corresponding band (see figure 19(b) – compare positions of the ladder bands in 19(a)). Areas of strong and weak background are present on both blots (see arrows), but can be distinguished from the bands by being not that well defined, but more shadowy.

Quantitation of the *PLP2* and $\beta$-actin expression was done on a PhosphorImager system, using the provided software. Automated peak recognition did not recognize the peaks precisely, and so, I did it manually by comparing both the peak shape and the appearance of the corresponding band in the blots (see figure 20).

The ratios of the areas under both the *PLP2* and $\beta$-actin peak were calculated for each sample (see appendix on page 87) and subsequently averaged for the patient and control




(a) Definition of the lanes in the *PLP2* hybridization

(b) Resulting peaks for the *PLP2* mRNA

(c) Definition of the lanes in the $\beta$-actin hybridization

(d) Resulting peaks for the $\beta$-actin mRNA

Figure 20: **Determination of the hybridization signal intensities in the *PLP2* and $\beta$-actin hybridization.** In order to quantify the intensity of the radioactive signal of both probes, the thickness and black value of each band had to be measured. After definition of the gel lanes (see figures 20(a) and 20(c)), the ImageQuant software computes the chromatograms shown in figures 20(b) and 20(d). Because of the background noise (see left arrow in 20(b)), which was stronger in the scanned autoradiographs (taken for the *PLP2* hybridization) than in the PhosphorImager images (taken for the $\beta$-actin hybridization), beginning and end of the peaks had to be adjusted manually. Also, the level of the baseline had to be adapted to the noise level surrounding the peak in all lanes. After predefinition of these variables, the software calculated the area under the peaks (see arrows on the right in both figures), the area being the summation of all pixels under the peak times their intensity.

panel, showing that the expression level of the *PLP2* gene is reduced by the transversion to almost one third of the level in $PLP2^{-188(C)}$ (see figure 21).

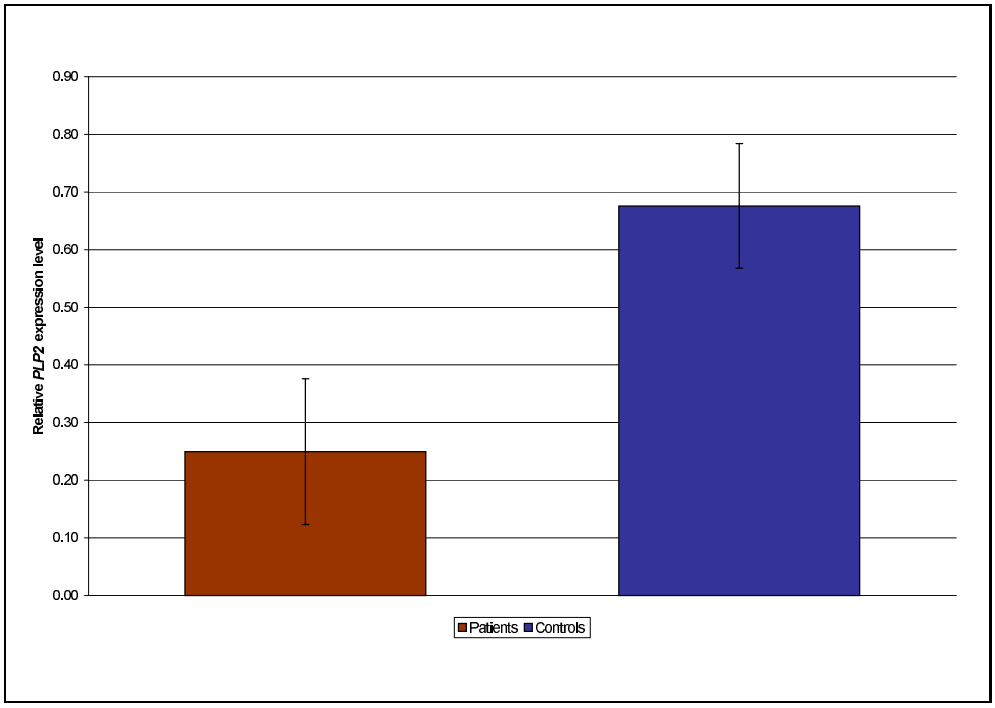



Figure 21: **Normalized expression level of the *PLP2* gene in patients and controls.** By dividing the *PLP2* software calculated peak areas by the $\beta$-actin peak areas, I normalized the *PLP2* expression level for each sample. The mean normalized *PLP2* peak area of both the patient and control sample population was then calculated and is illustrated here. The error bars in the chart denote the standard deviation. I found the normalized *PLP2* expression level to be more than 2.7 fold reduced in the patients in comparison to the controls, pointing towards a functional implication of the mutation in the *PLP2* promoter region. The significance of these data is high, as can be seen from the non-overlapping error bars that represent the standard deviation, but also from the calculated p-value that is as low as 0.000037, meaning that the probability of these data to be a result of pure chance is nearly null.

The significance level of this result is high, as could be determined by a Student's $t$-test analysis [93]. This analysis determines the likelihood of the null hypothesis [91] that the means of two normally distributed populations are equal, meaning that deviations of individual dates from the mean occured by pure chance. Using a two-tailed significance test, which is suitable for samples coming from two different populations, a p-value [92] of $3.7 \cdot 10^{-5}$ was calculated, which is much lower than the 0.05 threshold level for statistical significance. Hence, the null hypothesis that the two groups do not differ is rejected in favor of the alternative hypothesis that the groups do differ. This suggests that the significant reduction of *PLP2* expression is a consequence of the mutation in the promoter of the gene.

## 3.2 Frequency of the *PLP2* Promoter Mutation

To determine the frequency of the *PLP2* promoter mutation in healthy individuals, I analyzed three different control panels. On the one hand, I checked two panels of blood donors with unknown IQ, representing the common population (control pools 1 and 2), and on the other hand, I analyzed individuals with an at least average IQ (control pool 3).

After PCR amplification of a 326 bp fragment containing the site of interest at the $212^{th}$ base, DHPLC analysis was used to prescreen the PCR products for the transversion. Positive or doubtful DHPLC results were verified by sequencing of the corresponding PCR products.

### 3.2.1 PCR Amplification of the Promoter Region

The whole genomic DNA from the control panels was amplified prior to *PLP2* promoter PCR by use of Phi29 DNA polymerase. This procedure amplifies the entire genomic DNA in a representational way, but decreases PCR efficiency. Therefore, I optimized PCR parameters prior to amplification of the entire control panels (see appendix from page 88 on). Nevertheless, I had to revert to using genomic DNA in some cases.

### 3.2.2 Analysis of the PCR Products on the DHPLC System

All PCR products generated were analyzed on the DHPLC system, with exception of the PCR products from control pool 3, for they exhibited a highly diverging concentration and thus had to be submitted directly to sequencing.

For DHPLC analysis, PCR products of two patients have to be mixed (see pipetting schemes on page 92), such that a subsequent thermal denaturation and renaturation step will result not only in double-stranded PCR products that originated from the same patient (homoduplices), but also in PCR products containing reannealed single-strands from both patients. If the sequence of the PCR products of both patients is the same, again only homoduplices will form. But if there is a difference between them, heteroduplices will be generated, these heteroduplices being single-stranded at the site of the mismatch. Due to this imperfectness in base pairing, heteroduplices will elute earlier than homoduplices. So, if there are differences in the sequence of two mixed PCR products, two elution peaks will be seen in the chromatogram, whereas only one is present if both sequences are identical. An example for both cases from my PCR products is shown in figure 22.

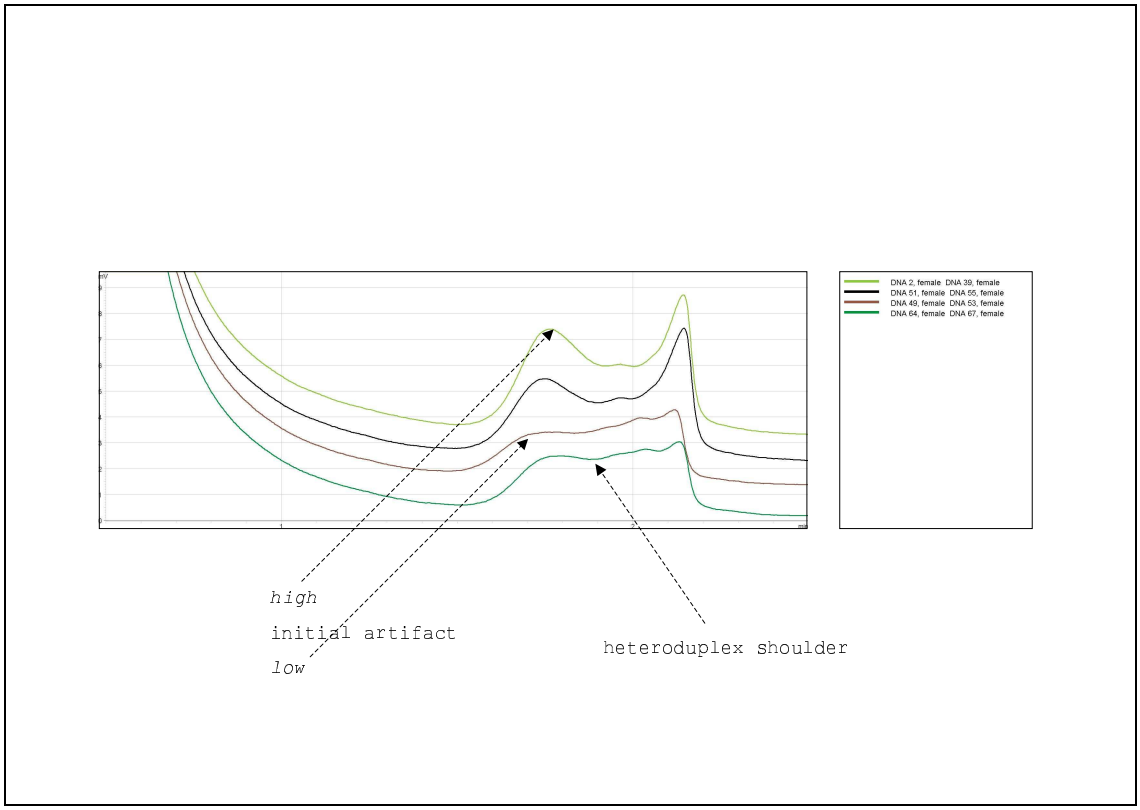
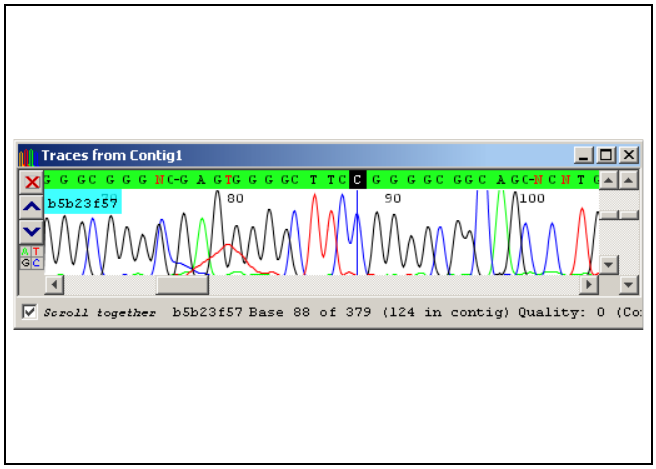



Figure 22: **Evaluation of DHPLC chromatograms.** In theory, a positive DHPLC chromatogram has to consist of two well-separated major peaks, corresponding to the heteroduplex (eluted first) and to the homoduplex, respectively. Both peaks ideally can be recognized to consist of two sub-peaks: one from each of both possible hetero- and homoduplices (see figure 9 on page 25). According to the theory, both negative chromatograms (on top) here also consist of two major peaks. As the first peak is present also in the negative samples (see arrows on the left), it is just an artifact, but gives hints on identifying positive chromatgrams: This artifact is smaller in the positive samples, giving way to a long plateau that is not present in the negative samples (see arrow on the right). Also, the following PCR product elution peak is smaller, resulting in three signs that unambigously indicate a positive sample.

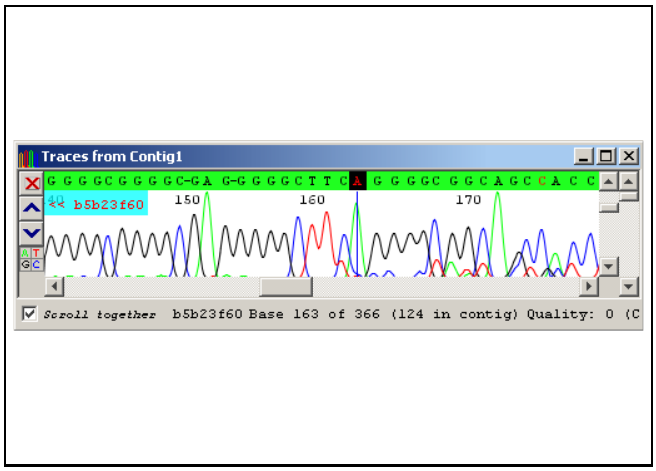
### 3.2.3 Sequencing of Candidate PCR Products

All PCR products that, as a result of the DHPLC analysis, probably contained the A-variant of the *PLP2* promoter region or that failed to be DHPLC analyzed were sequenced. For the females, mutation carriers could be subdivided into homozygotes and heterozygotes (see examples in figure 23).

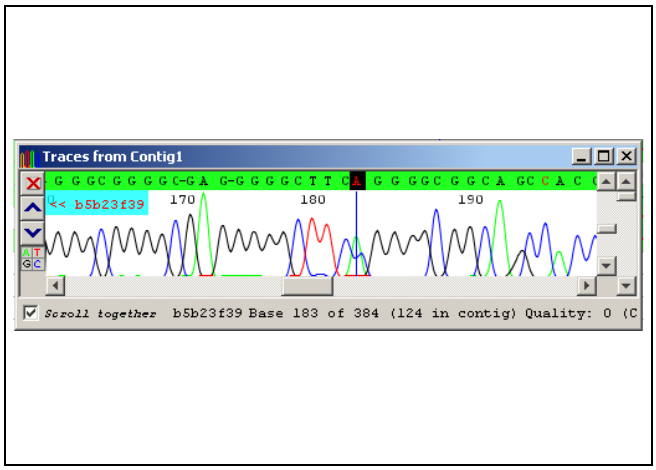
Table 28 on page 67 shows the calculated $PLP2^{-188(A)}$ frequencies in the three different control panels (for detailed results for each panel see appendix on page 96). In those control pools that contained genetic material from people representing the common population (pools 1 and 2), the $PLP2^{-188(A)}$ variant was found with a maximum frequency of 2.1% in the males and 4.3% per X chromosome in the females, whereas a maximum of 0.5% was found for the homozygous females. Contrasting, the $PLP2^{-188(A)}$ variant is not present in the males from the control pool comprising people with at least average IQ (pool 3), but




(a) Sequence of individual homozygous for $PLP2^{-188(C)}$




(b) Sequence of individual homozygous for $PLP2^{-188(A)}$




(c) Sequence of individual heterozygous for $PLP2^{-188(A)}$

Figure 23: **Examples of readouts of $PLP2^{-188}$ sequences.** As the *PLP2* gene is located on the X chromosome, males are always hemizygous and can carry either $PLP2^{-188(C)}$ or $PLP2^{-188(A)}$ (see figures 23(a) and 23(b), respectively). Females can be also heterozygous for both variants (see 23(c)). These heterozygous individuals are recognized by the presence of both the "C" from $PLP2^{-188(C)}$ and the "A" from $PLP2^{-188(A)}$ at same site of the sequence, the two peaks in the chromatogram corresponding to the two differing copies of the gene.

heterozygously in 6.3% of the females. This highest number must be relativized due to the fact that only 16 individuals were screened.

In a previous screening, these frequencies of the A-variant of the *PLP2* promoter were confirmed. Here, the $PLP2^{-188(A)}$ genotype was not found in 140 males with above average IQ, but twice in 35 females being heterozygous for the A-variant, corresponding to 2.8% of the X chromosomes being affected. In controls with unknown mental retardation status, the change was found in 7 out of 266 males (2.6%) and in 7 out of 96 females (3.6% of the X chromosomes).

When taking the patients from the Northern blot analysis, $PLP2^{-188(A)}$ was found in 9 out of 347 males from the corresponding mental retardation families with proven or suspected X-linkage, which results in a frequency of 2.6%. That is exactly the same value as in the male controls with unknown mental retardation status, respresenting the common population.

Taking everything together, this means that $PLP2^{-188(A)}$ is a rare polymorphism, for it is present in the range from 1% to 5% of the common population. However, the absence of $PLP2^{-188(A)}$ in a homozygous state in individuals with at least average IQ argues in favor of an implication of the $PLP2^{-188(A)}$ genotype in mental retardation. The association of the reduced mRNA levels with this polymorphism prompted the question whether the polymorphism is functional. This was analyzed in an electrophoretic mobility shift assay.

Table 28: **Frequencies of the *PLP2* promoter variation in all control panels**

| | *Number of mutated individuals* total (heterozygous/homozygous) | *Total number of individuals* | *Mutation frequency* [%] | *Mutation frequency per X chromosome* [%] |
|---|---|---|---|---|
| *Control pool 1* | | | | |
| males: | 2 | 96 | 2.1 | 2.1 |
| females: | 8 (7/1) | 96 | 8.3 (7.3/1) | 4.2 (3.6/0.5) |
| *Control pool 2* | | | | |
| males: | 0 | 24 | 0 | 0 |
| females: | 2 (2/0) | 23 | 8.7 (8.7/0) | 4.3 (4.3/0) |
| *Control pool 3* | | | | |
| males: | 0 | 27 | 0 | 0 |
| females: | 2 (2/0) | 16 | 12.5 (12.5/0) | 6.3 (6.3/0) |

## 3.3 Electrophoretic Mobility Shift Assay

An Electrophoretic Mobility Shift Assay (EMSA, explained in detail in the methods section) is used to find out if a given DNA sequence can be bound by one or more proteins. Mostly, one will use this feature to analyze regulatory sites in promoter regions of genes and to find the proteins that interact with them.

If a short double-stranded and labeled oligonucleotide is bound by a protein, a protein-DNA complex forms whose electrophoretic mobility is lower than that of the labeled oligonucleotide alone. After electrophoresis, protein-DNA complexes will therefore show a retarded band in comparison to a much farther migrated band from the unbound oligonucleotide. A protein-free control reaction should show only this farther advanced band. When adding increasing amounts of unlabeled oligonucleotide of the same sequence, protein binding to the labeled oligonucleotide must disappear gradually, as the labeled oligonucleotide is being displaced from the protein-DNA complex by unlabeled and therefore invisible oligonucleotide (specific competition). Also, non-specific binding of proteins to the DNA, that can take place e.g. because of the DNA's negative charge, has to be prevented by adding an excess of DNA with unrelated sequence (non-specific competition). With prior knowledge about putatively binding transcription factors (e.g. from an *in silico* analysis in a database), specific antibodies angainst these transcription factors can be selected and added to the binding reaction. If an antibody binds a transcription factor successfully, the protein-DNA complex is becoming larger as compared to the binding of the transcription factor alone. The consequence is that an even more retarded band can be observed (so-called super-shift).

### 3.3.1 *In silico* Binding Studies of the *PLP2* Promoter Element

To get a first idea about the consequences the cytosine to adenine change in the *PLP2* promoter region has for transcription factor binding, I consulted the MatInspector transcription factor database [66]. Predicted binding to 300 bp of sequence from $PLP2^{-188(C)}$ and $PLP2^{-188(A)}$ turned out to be differential (see figure 24). The A-variant was predicted to be unable to bind two transcription factors that $PLP2^{-188(C)}$ can bind: REL (NM_002908 [5]), a member of the Nuclear Factor (NF) family of transcription factors, and ETS1 (NM_005238 [5]), a member of the ETS family of transcription factors (see figure 24). The presence of binding sites for both factors was confirmed by the databases Match [42], rVista [49], ConSite [75], TESS [78], TFBIND [87] and TransFac [94].

MatInspector calculates two variables for the efficiency with which a transcription factor binds its matrix: the core and the matrix similarity. The core similarity denotes the degree of similarity between the usually four consecutive base pairs with highest conservation and the binding matrix, whereas the matrix similarity integrates the overall similarity of the investigated sequence to the usually 20 base pairs of matrix of the particular transcription factor. Both similarities are important for protein binding, and therefore both were included in the $PLP2^{-188(C)}$ and $PLP2^{-188(A)}$ oligonucleotides that were used for

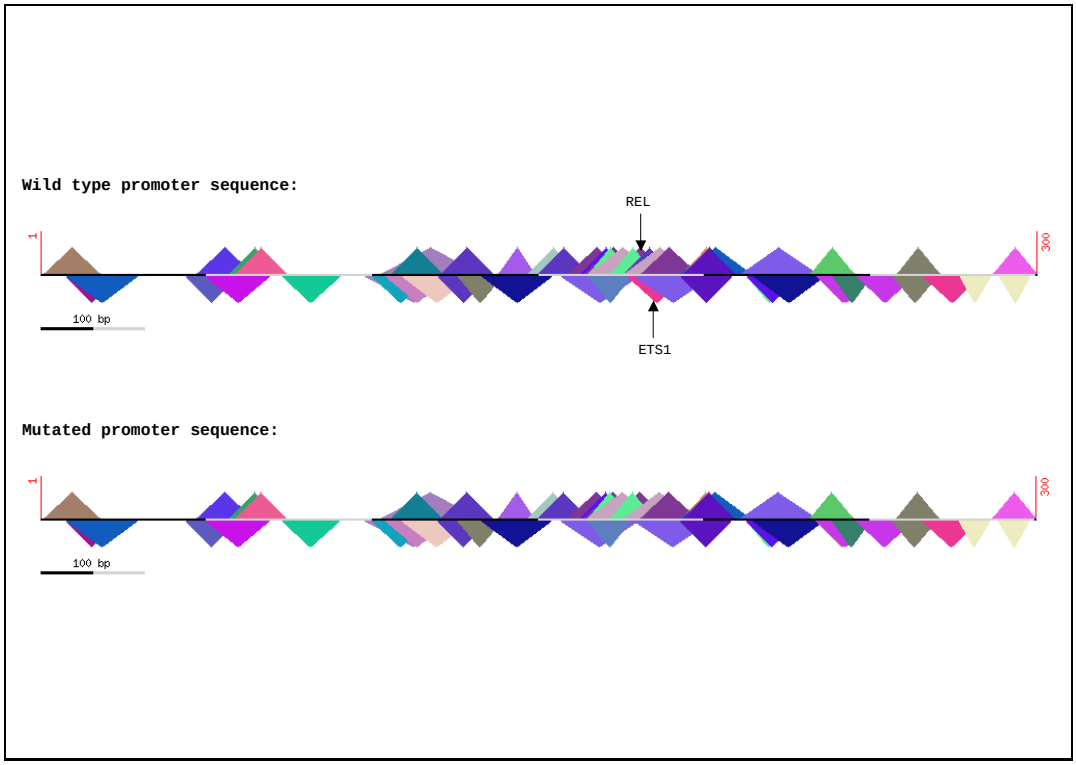



Figure 24: **Putative binding of transcription factors in the promoter region of *PLP2*.** In 300 bp of sequence directly upstream of the *PLP2* transcription start site (located 75 bp upstream of translation start), the MatInspector database [66] predicts a lot of transcription factors to bind (represented by the colored triangels), from which only two show a difference concerning binding of the $PLP2^{-188(C)}$ and the $PLP2^{-188(A)}$ oligonucleotide: the transcription factors REL and ETS1 do not bind to the A-variant of the promoter sequence, whereas they do bind to $PLP2^{-188(C)}$ (see arrows).

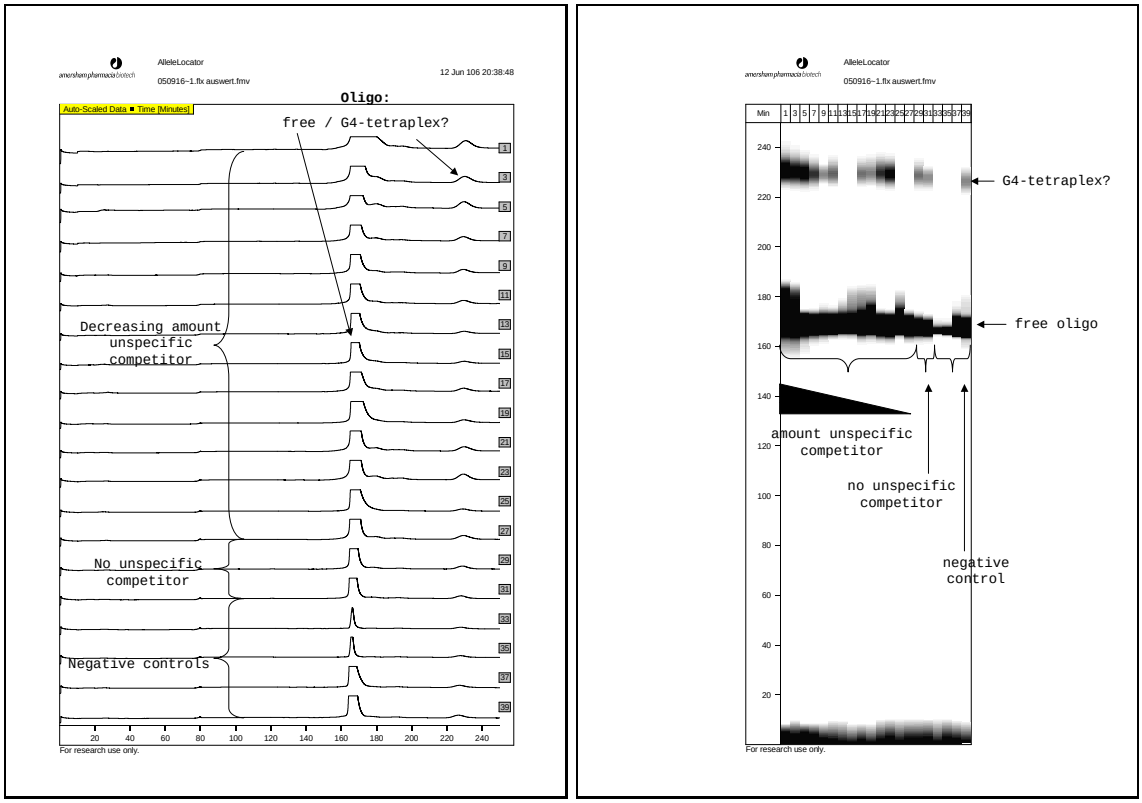
EMSA analysis.

### 3.3.2 EMSA Analysis of the *PLP2* Promoter Mutation

Protein binding was first investigated by means of an fluorescent EMSA, performed in polyacrylamide gels on an A.L.F. DNA Sequencer. As source for transcription factors, extracts of nuclear proteins were prepared from a lymphoblastoid cell line of a patient, carrying the $PLP2^{-188(A)}$ genotype, and a $PLP2^{-188(C)}$ control.

As a first critical question to answer, I wanted to find out how much protein is needed to cause a clearly visible band shift. From online accessible protocols [61], [63], [43] and [83], I learned that the amount of protein to be used in the binding reaction has to be titrated for each analysis, usually lying in a range of 10 ng to 20 µg per reaction. Unfortunately, such a titration did not only result in a band shift in the test reactions, but also in the negative controls that did not contain any protein, but which were otherwise composed identically. Protein contamination of the negative controls could be excluded in repetitions, but the problem remained (see figure 25 for an example).

Thus, I came to the opinion that the retarded band was due to a high molecular




(a) Curve view of EMSA result (b) Software-inferred gel view of 25(a)

Figure 25: **Results from fluorescent EMSA on A.L.F. DNA Sequencer.** To test the amount of unspecific competitor needed in order to successfully suppress unspecific protein-DNA binding, the amounts of poly-dIdC in each reaction was decreased stepwise from 2 µg to 0.0625 µg. 25(a) After three hours, a peak was observed on all lanes, resulting from the labeled free oligonucleotide. Approximately one hour later, a retarded fragment reached the detector. As it was also present in the protein-free negative controls, it could not have been due to the formation of a specific protein-DNA complex, but to multimerization of the oligonucleotides with themselves to so-called G-quartets. 25(b) The amount of competitor did not have any consistent influence on the shifted bands. All reaction volumes were splitted and loaded onto two adjacent lanes, but have differently intense shift bands, as can be seen in lanes 37 and 39, e.g.

weight complex formed by the labeled oligonucleotide alone: As both oligonucleotides contain stretches of four or more guanine nucleotides, they are able to build up so-called G-quartets (see figure 29 on page 77). For the oligonucleotides used in the EMSAs contain three G stretches of at least two nucleotides, they are able to build up a complex and high molecular weight aggregate of oligonucleotides that migrates at a markedly reduced velocity and therefore is able to cause a retarded band in the EMSAs – in the absence of any protein.

The formation of G-quartets might abolish or reduce transcription factor binding, which could also be an explanation for the absence of specific band shifts. However, a very large protein-DNA complex, e.g. resulting from protein-protein interactions, would not be seen on this machine as well, because it would need too much time to reach the detector. To

test both hypotheses, I used a radioactive EMSA system which allows documentation of the entire gel.

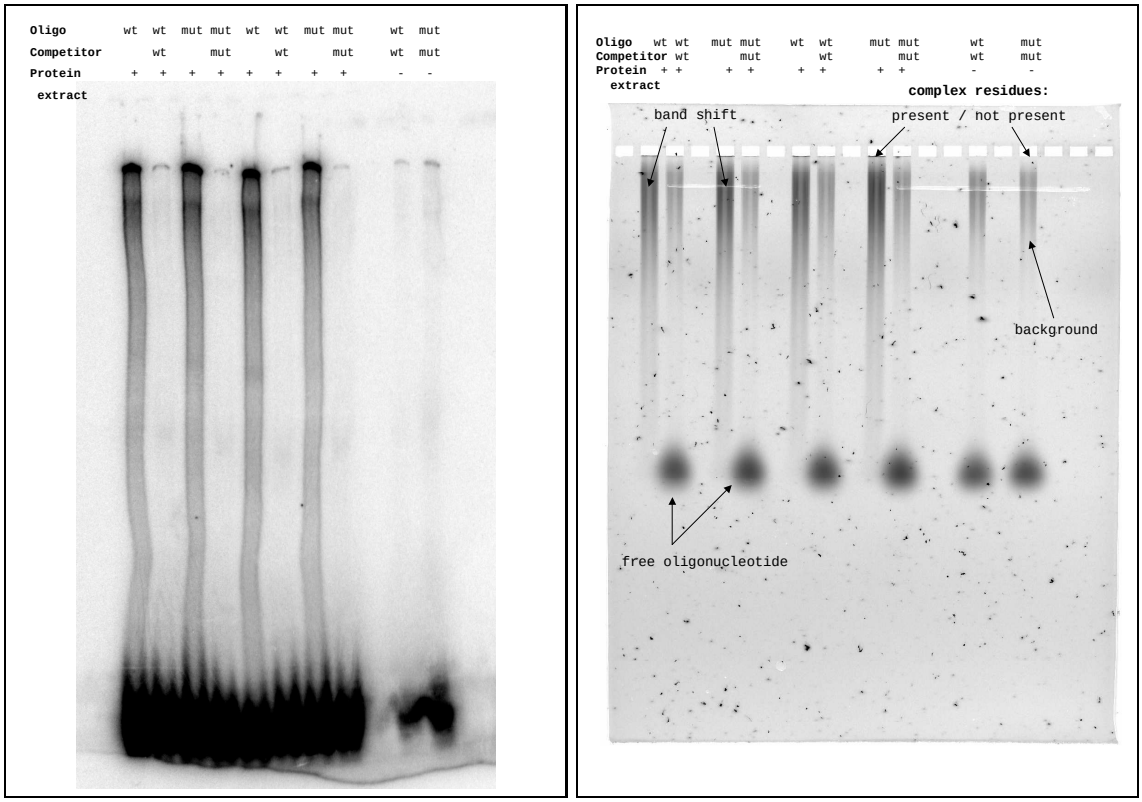
After optimization of EMSA parameters (see page 98) in the radioactive EMSA, a strong signal was detected in the gel pockets (see figure 26), supporting the hypothesis that the protein-DNA complex was too large to enter the gel. As this might have been the consequence of the small pore size that even low-percentage polyacrylamide gel have, an agarose gel EMSA system was used. Agarose gel EMSAs have been reported for analysis of very large complexes, e.g. those that are formed by the eucaryotic transcription factor IID [98].

In fact, this system yielded a much better entrance of the protein-DNA complex in the gel (see figure 26). Especially when looking at the free oligonucleotide, one can clearly see that protein-DNA binding occured: only after addition of specific competitor, free labeled oligonucleotide can be seen, as previously all of it was bound by proteins and shifted up in the gel view after electrophoresis.

In order to investigate putative differences in protein binding capacity of the oligonucleotides representing the $PLP2^{-188(C)}$ and the $PLP2^{-188(A)}$ sequence, specific and reciprocal competitions were done. In specific competition, increasing amounts of unlabeled oligonucleotide are added to the labeled oligonucleotide of identical sequence, whereas in reciprocal competition, increasing amounts of unlabeled oligonucleotide of the other sequence are added. In this approach, differences in the strength of protein binding to both oligonucleotides can be evaluated by the differing amount of unlabeled oligonucleotide needed to achieve the same effect in specific and reciprocal competition. However, no differences were observed, the intensity of the shifted band decreasing at identical rate – irrespective of the specificity of the competitor (see figure 27).

Having found a binding to both the $PLP2^{-188(C)}$ and the $PLP2^{-188(A)}$ oligonucleotide, I tried to figure out a quantitive difference between them. Therefore, I competed the binding to the $PLP2^{-188(C)}$ oligonucleotide not only with an increasing amount of unlabeled $PLP2^{-188(C)}$ oligonucleotide (specific competition), but also with the unlabeled $PLP2^{-188(A)}$ oligonucleotide (reciprocal competition), and *vice versa*. The reciprocal competition will need a higher degree of excess, if the protein binding is specific. Surprisingly, it made no difference which unlabeled oligonucleotide was added: in all cases, the intensity of the shifted band decreased with increasing molar excess of competitor (see figure 27).

After 19 fluorescent EMSAs on the DNA sequencer, 27 with the radioactively labeled oligonucleotides and 9 in the agarose gel, these results could not be improved. Obviously, there is protein binding to both the $PLP2^{-188(C)}$ and the $PLP2^{-188(A)}$ oligonucleotide, but differences explaining the lower *PLP2* expression rates in the A-variant could not be found in these assays. Due to the low migration rate of the protein-DNA complex during electrophoresis, super-shift assays were also inconclusive (not shown).




(a) Radioactive EMSA (b) Fluorescent agarose gel EMSA

Figure 26: **Radioactive polyacrylamide gel EMSA vs. fluorescent agarose gel EMSA.** In figure 26(a), most of the radioactive signal can be found at the very bottom of the gel, representing the free oligonucleotide. In the environment of the gel pockets, a strong radioactive signal can be seen, which disappeared by specific competition and removal of nuclear extract in the binding reaction. However, a signal that is much more intense can be found directly in the gel pockets. This indicates that a very large protein-DNA complex formed and that most of this complex was not able to enter the small meshed polyacrylamide gel, but was retained in the pockets. Use of a much weaker agarose gel could solve this problem (see figure 26(b)): Here, a very large shifted band can be seen, which disappeared to a background level in the specifically competed reactions and the negative controls, again pointing towards the formation of a specific protein-DNA complex. Supportingly, free oligonucleotide was only found in those lanes where either unlabeled specific competitor displaced labeled oligonucleotide from the protein-DNA complexes or where no nuclear extract was present, so that the labeled oligonucleotide could not be bound. In the pockets of the lanes of both the negative controls and the specific competitions, protein-DNA complex residues were not found at all, whereas tiny amounts could be found in the remaining reactions where the complex was much stronger. In contrast to the *in silico* prediction, there was no apparent difference in protein binding capacity between $PLP2^{-188(A)}$ and $PLP2^{-188(C)}$.

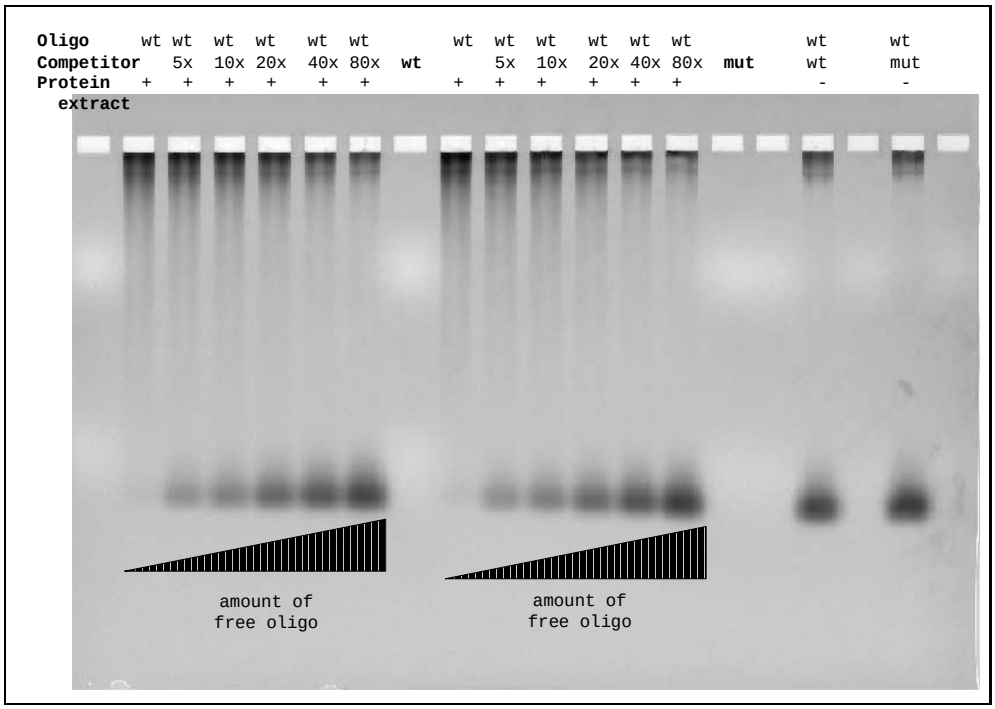



Figure 27: **Protein binding occurs to both the $PLP2^{-188(C)}$ and the $PLP2^{-188(A)}$ oligonucleotide.** With increasing amounts of specific competitor, increasing amounts of labeled $PLP2^{-188(C)}$ oligonucleotide were set free, since most of the protein bound to the excess of unlabeled competitor. The same is true when competing with the unlabeled oligonucleotide with the A-variant sequence. Both types of competition took place with identical efficiency, as one can see no difference in the rate of appearance of free oligonucleotide or disappearance of band shift, respectively.

# 4 Discussion

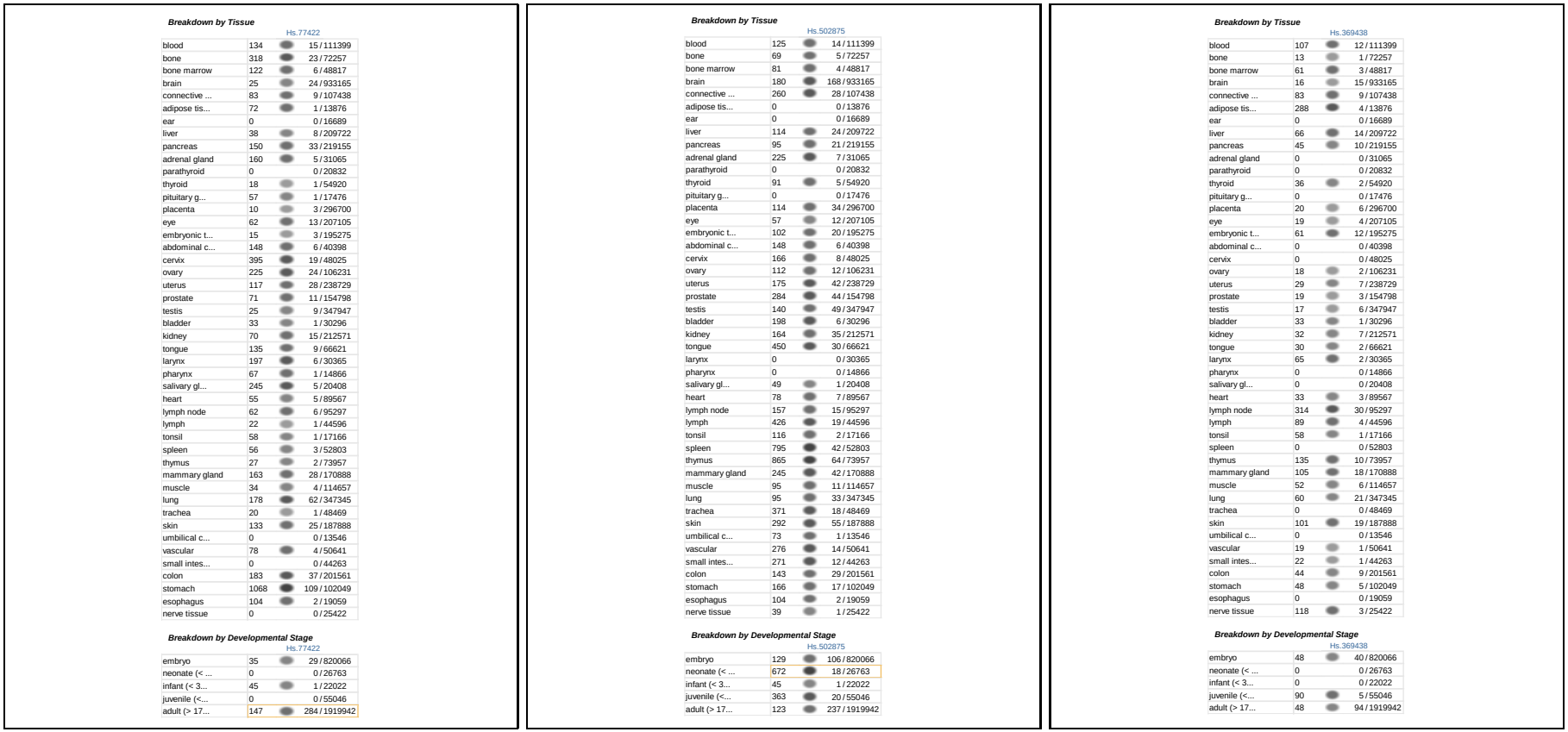
## 4.1 Expression Level of the *PLP2* Gene

The expression level of the *PLP2* gene was studied in lymphoblastoid cell lines. These cell lines result from transfection of blood sample-derived lymphocytes with Epstein-Barr Virus (EBV). EBV transforms human B lymphocytes into lymphoblasts, i.e. tumor cells that are able to proliferate indefinitely in tissue culture [81]. The transformation does not change the gene expression pattern of the lymphocytes [21], with exception of the genes that are involved in immortalization. As lymphocytes show strong similarities to brain tissue concerning gene expression [85] and are readily available in the laboratory, lymphoblastoid cell lines became an accepted model for investigating mental retardation genes.

When analyzing the Northern blots that were done with the RNA isolated from lymphoblastoid cell lines of patients carrying adenine instead of cytosine at position –188 in the *PLP2* promoter region ($PLP2^{-188(A)}$) and with the RNA from the controls with $PLP2^{-188(C)}$, precise peak quantitation necessitated, due to background noise, manual editing of the ImageQuant software output. In this manner, however, coherent peak quantitation data could be obtained (see table 30 in appendix on page 87). *PLP2* expression in patients carrying the promoter variant $PLP2^{-188(A)}$ showed a more than 2.7 fold reduced *PLP2* mRNA levels in direct comparison with the controls. This strongly argues for an involvement of the $PLP2^{-188(A)}$ variant in transcription regulation by affecting the protein binding capacity of a regulatory DNA element. This assumption was substantiated by *in silico* investigations with MatInspector [66] which identified $PLP2^{-188(A)}$ as affecting the core sequence of a protein binding site by abolishing the predicted binding of the transcription factors REL and ETS1.

The transcription factor genes *REL* and *ETS1* are brain-expressed (see figure 28) and good candidates for an implication in regulation of the *PLP2* gene. REL transcription factors are not only involved in immune processes, but also in the control of cell cycle and apoptosis [67]. The PLP2 protein takes part in these two processe, too (see articles [89] and [48]). Hence, *PLP2* might be regulated by REL factors.

For ETS1, similar findings are reported in the literature. Besides its role in induction of angiogenesis [36] and ovarian cancers [82], ETS1 is also shown to be involved in immune responses [7], even to directly interact with proteins from the NF$\kappa$B family [4] REL also belongs to. Moreover, this interaction suggests a participation of ETS1 in apoptosis, which was found in the positive regulation of Fas ligand transcription [41], i.e. a great deal upstream of the location of the PLP2 protein in apoptotic signaling.



**Breakdown by Tissue** — Hs.77422

| | | |
|---|---|---|
| blood | 134 | 15 / 111399 |
| bone | 318 | 23 / 72257 |
| bone marrow | 122 | 6 / 48817 |
| brain | 25 | 24 / 933165 |
| connective ... | 83 | 9 / 107438 |
| adipose tis... | 72 | 1 / 13876 |
| ear | 0 | 0 / 16689 |
| liver | 38 | 8 / 209722 |
| pancreas | 150 | 33 / 219155 |
| adrenal gland | 160 | 5 / 31065 |
| parathyroid | 0 | 0 / 20832 |
| thyroid | 18 | 1 / 54920 |
| pituitary g... | 57 | 1 / 17476 |
| placenta | 10 | 3 / 296700 |
| eye | 62 | 13 / 207105 |
| embryonic t... | 15 | 3 / 195275 |
| abdominal c... | 148 | 6 / 40398 |
| cervix | 395 | 19 / 48025 |
| ovary | 225 | 24 / 106231 |
| uterus | 117 | 28 / 238729 |
| prostate | 71 | 11 / 154798 |
| testis | 25 | 9 / 347947 |
| bladder | 33 | 1 / 30296 |
| kidney | 70 | 15 / 212571 |
| tongue | 135 | 9 / 66621 |
| larynx | 197 | 6 / 30365 |
| pharynx | 67 | 1 / 14866 |
| salivary gl... | 245 | 5 / 20408 |
| heart | 55 | 5 / 89567 |
| lymph node | 62 | 6 / 95297 |
| lymph | 22 | 1 / 44596 |
| tonsil | 58 | 1 / 17166 |
| spleen | 56 | 3 / 52803 |
| thymus | 27 | 2 / 73957 |
| mammary gland | 163 | 28 / 170888 |
| muscle | 34 | 4 / 114657 |
| lung | 178 | 62 / 347345 |
| trachea | 20 | 1 / 48469 |
| skin | 133 | 25 / 187888 |
| umbilical c... | 0 | 0 / 13546 |
| vascular | 78 | 4 / 50641 |
| small intes... | 0 | 0 / 44263 |
| colon | 183 | 37 / 201561 |
| stomach | 1068 | 109 / 102049 |
| esophagus | 104 | 2 / 19059 |
| nerve tissue | 0 | 0 / 25422 |

**Breakdown by Developmental Stage** — Hs.77422

| | | |
|---|---|---|
| embryo | 35 | 29 / 820066 |
| neonate (< ... | 0 | 0 / 26763 |
| infant (< 3... | 45 | 1 / 22022 |
| juvenile (<... | 0 | 0 / 55046 |
| adult (> 17... | 147 | 284 / 1919942 |

(a) *PLP2* expression

**Breakdown by Tissue** — Hs.502875

| | | |
|---|---|---|
| blood | 125 | 14 / 111399 |
| bone | 69 | 5 / 72257 |
| bone marrow | 81 | 4 / 48817 |
| brain | 180 | 168 / 933165 |
| connective ... | 260 | 28 / 107438 |
| adipose tis... | 0 | 0 / 13876 |
| ear | 0 | 0 / 16689 |
| liver | 114 | 24 / 209722 |
| pancreas | 95 | 21 / 219155 |
| adrenal gland | 225 | 7 / 31065 |
| parathyroid | 0 | 0 / 20832 |
| thyroid | 91 | 5 / 54920 |
| pituitary g... | 0 | 0 / 17476 |
| placenta | 114 | 34 / 296700 |
| eye | 57 | 12 / 207105 |
| embryonic t... | 102 | 20 / 195275 |
| abdominal c... | 148 | 6 / 40398 |
| cervix | 166 | 8 / 48025 |
| ovary | 112 | 12 / 106231 |
| uterus | 175 | 42 / 238729 |
| prostate | 284 | 44 / 154798 |
| testis | 140 | 49 / 347947 |
| bladder | 198 | 6 / 30296 |
| kidney | 164 | 35 / 212571 |
| tongue | 450 | 30 / 66621 |
| larynx | 0 | 0 / 30365 |
| pharynx | 0 | 0 / 14866 |
| salivary gl... | 49 | 1 / 20408 |
| heart | 78 | 7 / 89567 |
| lymph node | 157 | 15 / 95297 |
| lymph | 426 | 19 / 44596 |
| tonsil | 116 | 2 / 17166 |
| spleen | 795 | 42 / 52803 |
| thymus | 865 | 64 / 73957 |
| mammary gland | 245 | 42 / 170888 |
| muscle | 95 | 11 / 114657 |
| lung | 95 | 33 / 347345 |
| trachea | 371 | 18 / 48469 |
| skin | 292 | 55 / 187888 |
| umbilical c... | 73 | 1 / 13546 |
| vascular | 276 | 14 / 50641 |
| small intes... | 271 | 12 / 44263 |
| colon | 143 | 29 / 201561 |
| stomach | 166 | 17 / 102049 |
| esophagus | 104 | 2 / 19059 |
| nerve tissue | 39 | 1 / 25422 |

**Breakdown by Developmental Stage** — Hs.502875

| | | |
|---|---|---|
| embryo | 129 | 106 / 820066 |
| neonate (< ... | 672 | 18 / 26763 |
| infant (< 3... | 45 | 1 / 22022 |
| juvenile (<... | 363 | 20 / 55046 |
| adult (> 17... | 123 | 237 / 1919942 |

(b) *REL* expression

**Breakdown by Tissue** — Hs.369438

| | | |
|---|---|---|
| blood | 107 | 12 / 111399 |
| bone | 13 | 1 / 72257 |
| bone marrow | 61 | 3 / 48817 |
| brain | 16 | 15 / 933165 |
| connective ... | 83 | 9 / 107438 |
| adipose tis... | 288 | 4 / 13876 |
| ear | 0 | 0 / 16689 |
| liver | 66 | 14 / 209722 |
| pancreas | 45 | 10 / 219155 |
| adrenal gland | 0 | 0 / 31065 |
| parathyroid | 0 | 0 / 20832 |
| thyroid | 36 | 2 / 54920 |
| pituitary g... | 0 | 0 / 17476 |
| placenta | 20 | 6 / 296700 |
| eye | 19 | 4 / 207105 |
| embryonic t... | 61 | 12 / 195275 |
| abdominal c... | 0 | 0 / 40398 |
| cervix | 0 | 0 / 48025 |
| ovary | 18 | 2 / 106231 |
| uterus | 29 | 7 / 238729 |
| prostate | 19 | 3 / 154798 |
| testis | 17 | 6 / 347947 |
| bladder | 33 | 1 / 30296 |
| kidney | 32 | 7 / 212571 |
| tongue | 30 | 2 / 66621 |
| larynx | 65 | 2 / 30365 |
| pharynx | 0 | 0 / 14866 |
| salivary gl... | 0 | 0 / 20408 |
| heart | 33 | 3 / 89567 |
| lymph node | 314 | 30 / 95297 |
| lymph | 89 | 4 / 44596 |
| tonsil | 58 | 1 / 17166 |
| spleen | 0 | 0 / 52803 |
| thymus | 135 | 10 / 73957 |
| mammary gland | 105 | 18 / 170888 |
| muscle | 52 | 6 / 114657 |
| lung | 60 | 21 / 347345 |
| trachea | 0 | 0 / 48469 |
| skin | 101 | 19 / 187888 |
| umbilical c... | 0 | 0 / 13546 |
| vascular | 19 | 1 / 50641 |
| small intes... | 22 | 1 / 44263 |
| colon | 44 | 9 / 201561 |
| stomach | 48 | 5 / 102049 |
| esophagus | 0 | 0 / 19059 |
| nerve tissue | 118 | 3 / 25422 |

**Breakdown by Developmental Stage** — Hs.369438

| | | |
|---|---|---|
| embryo | 48 | 40 / 820066 |
| neonate (< ... | 0 | 0 / 26763 |
| infant (< 3... | 0 | 0 / 22022 |
| juvenile (<... | 90 | 5 / 55046 |
| adult (> 17... | 48 | 94 / 1919942 |

(c) *ETS1* expression

Figure 28: **Expression level of *PLP2* and the transcription factors *REL* and *ETS1* in different human tissues and at different developmental stages (after UniGene [62] from the NCBI Database [90]).** All three genes are expressed in the brain. The strongest expression is shown for *REL*, followed by *PLP2* and *ETS1* (180, 25 and 16 gene Transcripts Per Million (TPM) total transcripts, as can be calculated from the fraction of the gene's Expressed Sequence Tags (ESTs) in the pool's total ESTs). In the lymphoblastoid cell lines, expression levels of all three genes are more or less the same. The genes are also expressed in the embryo, with increasing strength from *REL* over *ETS1* to *PLP2*.

## 4.2 Frequency of the *PLP2* Promoter Mutation

When investigating the frequency of $PLP2^{-188(A)}$ in patients as compared to controls, the variant was found in 2.6% of the individuals from mental retardation families with proven or suspected X-linkage, in 2.6% of 266 male blood donors with unknown IQ and in 2.1% of 96 males from a commercial control collection. These frequencies earmark $PLP2^{-188(A)}$ as a polymorphism, since it occurs with a frequency of more than 1% in a sexually propagating population [79].

In a study of physical disorders, all people that are physically healthy, especially concerning the trait that is investigated, can be taken as controls. Concerning mental retardation, the disease state of the patients cannot as easily be seen. Particularly people that are affected only to a borderline extent can escape one's notice. I therefore chose to investigate a population with at least average IQ (e.g. university students). In a total of 167 male individuals, the change was completely absent, even though, at a frequency of 2.6%, at least four $PLP2^{-188(A)}$ carriers had to be expected.

Since IQ, as a phenotype, is of quantitative nature and mental retardation is defined as the lower end of the spectrum, this finding suggests that $PLP2^{-188(A)}$ might be a predisposing factor for MR, especially as *PLP2* expression is not completely abolished. Thus, $PLP2^{-188(A)}$ might be generating an environment that – in combination with a different, causative mutation – is able to lower the threshold for the carrier to develop an MR phenotype, or to add to the severity thereof. This is reflected in the concept of genetic modifiers for monogenic diseases, as e.g. stated in [35]:

> "... these single-gene mutations must reside in a permissive genetic background for a disease phenotype to manifest. Segregating background genes can also modify the age of onset, rate of progression or severity of these diseases. These background genes that interact with the disease mutation and that are responsible for the specific phenotypes observed are commonly called genetic modifiers."

Functional implications of polymorphisms in regulatory regions (regulatory Single Nucleotide Polymorphisms, rSNPs) are expected to exist for virtually every gene and to be present in every second gene of an individual [13]. Moreover, it is hypothesized that those sequence variants which influence illness are unlikely to be located in the coding region of a gene, but in its promoter region, and thereby exert their influence via variation of gene expression. Such a functional effect is reported e.g. for TNF-$\alpha$ [2].

## 4.3 Electrophoretic Mobility Shift Assay

Reduced levels of *PLP2* mRNA in lymphoblasts from $PLP2^{-188(A)}$ carriers and *in silico* findings of differential protein binding between $PLP2^{-188(A)}$ and wild type $PLP2^{-188(C)}$ suggested *in vitro* investigations of the putatively affected regulatory element by Electrophoretic Mobility Shift Assays (EMSAs).

In the initially employed approach using fluorescently labeled oligonucleotides on an A.L.F. DNA sequencer, no shift could be observed. This could have been due to the absence of protein binding e.g. because of the formation of guanine quartets by the oligonucleotides. These structures can arise from G repeats of at least four nucleotides, as present two times in the oligonucleotides used for the shift assays. Guanine quartets have a biological function in stabilizing the chromosomal telomers.

G-quartets are four-stranded, right-handed helices, called quadruplexes or tetraplexes. Each G of one strand of the quadruplex establishes hydrogen bonds with two Gs of two neighboring strands (see figure 29A), resulting in a square planar array of four Gs. A central cation stabilizes this structure. The interacting Gs can be localized in four different strands (see figure 29B1), in two strands whose sequence is $G_4N_nG_4$ (figures 29B3 and B4), where N is a spacing stretch composed of non-G repeats, or in only one strand consisting of multiple G repeats (figure 29B2). Once formed, each of the G4 species is quite stable, with measured enthalpies approaching –25 kcal per mole of G-quartet. The association of oligonucleotides to such tetraplexes can effectively inhibit protein binding. However, not all oligonucleotides must be involved in tetraplex formation, and so, protein binding can still occur at a reduced level.

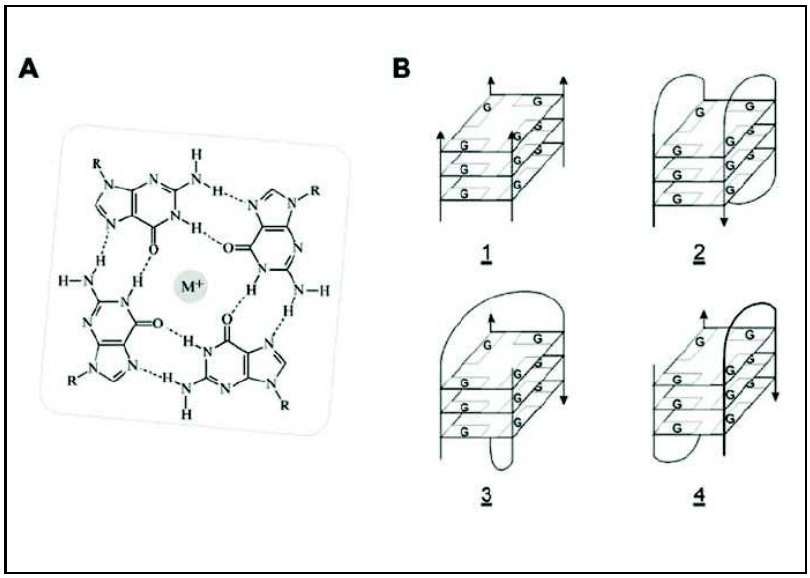



Figure 29: **Sterical structures of DNAs with guanine tetraplexes [88].** A) Four guanines establish Hoogstein hydrogen bonds with each other and are stabilized by a central monovalent cation, such as $Na^+$ or $K^+$. In this way, they give rise to guanine tetraplexes, square planar arrays of four guanines that are also called G-quartets. B) The interacting guanines can be located on four, one or two DNA or RNA strands (see B1, B2, B3 and B4, respectively).

As the A.L.F. system requires a comparatively long gel passage, very large protein-DNA complexes might not be able to reach the detector, either due to their size or due to destabilization as a consequence of the prolonged running time. Therefore, an also polyacrylamide gel based radioactive assay was used to determine whether large complexes were present, which revealed that the majority of the shifted signal was concentrated in the gel pockets.

In order to facilitate entry of the protein-DNA complexes into the gel, I used agarose gels which have much wider pores than polyacrylamide gels. Such kinds of EMSAs are reported in the literature, e.g. when the high molecular weight complexes formed by the essential eucaryotic transcription factor IID were analyzed [98]. In fact, agarose gels could improve complex entrance into the gel, but part of it still remained in the pockets.

As unspecific protein-DNA binding was suppressed by an excess of unspecific competitor poly-dIdC, I still had to demonstrate the specificity of the binding by competing with an excess of specific competitor. Addition of increasing amounts of this unlabeled oligonucleotide released increasing amounts of labeled oligonucleotide, whereas the entire amount of labeled oligonucleotide was bound without competition. This clearly shows that specific protein-DNA binding occurs.

However, I found not only protein-binding to the $PLP2^{-188(C)}$ oligonucleotide, but also to the mutant $PLP2^{-188(A)}$ oligonucleotide, which contradicted the results from the *in silico* analysis. The strength of the protein-binding to both oligonucleotides was the same, which could be demonstrated by competing the $PLP2^{-188(C)}$ oligonucleotide with increasing amounts of the $PLP2^{-188(A)}$ oligonucleotide, and *vice versa*. This can be explained in two ways: Either, the fluorescent assay is not sensitive enough to discern between subtle differences in protein-binding strength, or the A-variant is bound by a different transcription factor complex of similar size.

The latter point can be substantiated by the results of TFBIND [87]. This database predicts that $PLP2^{-188(A)}$, in contrast to $PLP2^{-188(C)}$, contains a binding site for the cAMP Responsive Element Binding Protein 1 (CREB1), a transcription factor involved in neuronal function and development [22] and exhibiting an anti-apoptotic effect [86]. *CREB1* is, as *REL* and *ETS1*, expressed in the brain and at embryonic stage [62]. Therefore, the reduction of *PLP2* expression has to be explained by CREB1 acting as a transcriptional silencer, which is in line with findings reported in the literature [22]. In addition, CREB1 activation is followed by subsequent homodimerization [50] and also by complexation with CREB Binding Protein (CBP) [59], which could result in the observed unresolvability on a polyacrylamide gel.

This leads to the former point: the EMSAs being perhaps not suitable to find a subtle, though present difference in protein-binding to the $PLP2^{-188(C)}$ and $PLP2^{-188(A)}$ oligonucleotides. As most of the protein-DNA complex resided in the gel pockets, differentiation of band intensities was strictly precluded.

## 4.4 A Role for *PLP2* in Mental Retardation

How can the finding that *PLP2* is significantly lower expressed in people affected by mental retardation be interpreted? How can a link from the lower *PLP2* expression be established to the MR phenotype? One out of the seven articles reviewed in the introduction provides an indication:

In human KB epithelial cells, Wang *et al.* [89] found the PLP2 protein to be an interaction partner of the anti-apoptotic protein BAP31 [89]. BAP31 is, as PLP2, located in the membrane of the endoplasmic reticulum [54] and was shown to abolish the oligomerization of BAX and BAK [89], two proteins of the BCL-2 family that are located in the outer mitochondrial membrane [46]. The oligomerization of both proteins is a prerequisite for the formation of a conduit through which cytochrome *c* can leave the mitochondrion [46], which results in subsequent activation of caspases 9 and 3 (see figure 31). Caspases are strong proteases that promote autodigestion of the cell [32]. Hence, extracellular death ligands signal via BCL-2 family members to programmed cell death.

In embryonic brain development, apoptosis plays an important role. At birth, only one third of all neurons that were produced since fertilization are still present, the remaining two thirds having depleted by apoptosis [32]. Findings in mice showed that the presence of pro-apoptotic factors is a prerequisite for viability [47] (see figure 30).

In contrast to that underrepresentation of apoptosis in caspase-9 knockout mice, mental retardation patients suffering from a lower expression of *PLP2* might have a higher extent of apoptosis of neural cells, since PLP2, in cooperation with BAP31, is an anti-apoptotic factor. The consequence could be a lower density of neurons in the brain. A way to demonstrate this could be to compare positron emission tomography images from patients and controls after administration of radioligands binding specifically to distinct receptors of the central nervous system (see [53] for examples). An excellent alternative for displaying soft tissues as the brain, avoiding exposition to radioactive substances and providing resolutions that are at least as good as those from computerized tomography would be magnetic resonance imaging [37].

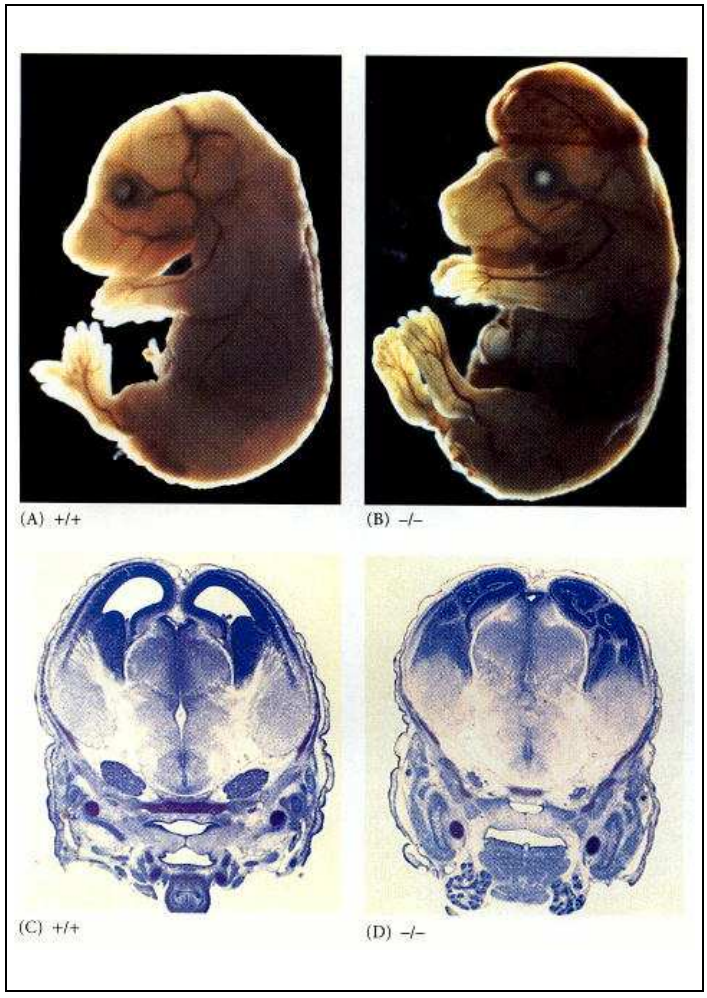



Figure 30: **The role of apoptosis in brain development [47].** A caspase-9 knockout mouse 16 days after implantation (B) has a significantly enlarged brain volume when compared to a wild type mouse of the same age (A). These findings are confirmed by cross-sections through the forebrain. At embryonic day 13.5, the knockout (D) exhibits, when compared to the wild type (C), an almost complete loss of the ventricles and also thickened ventricle walls.

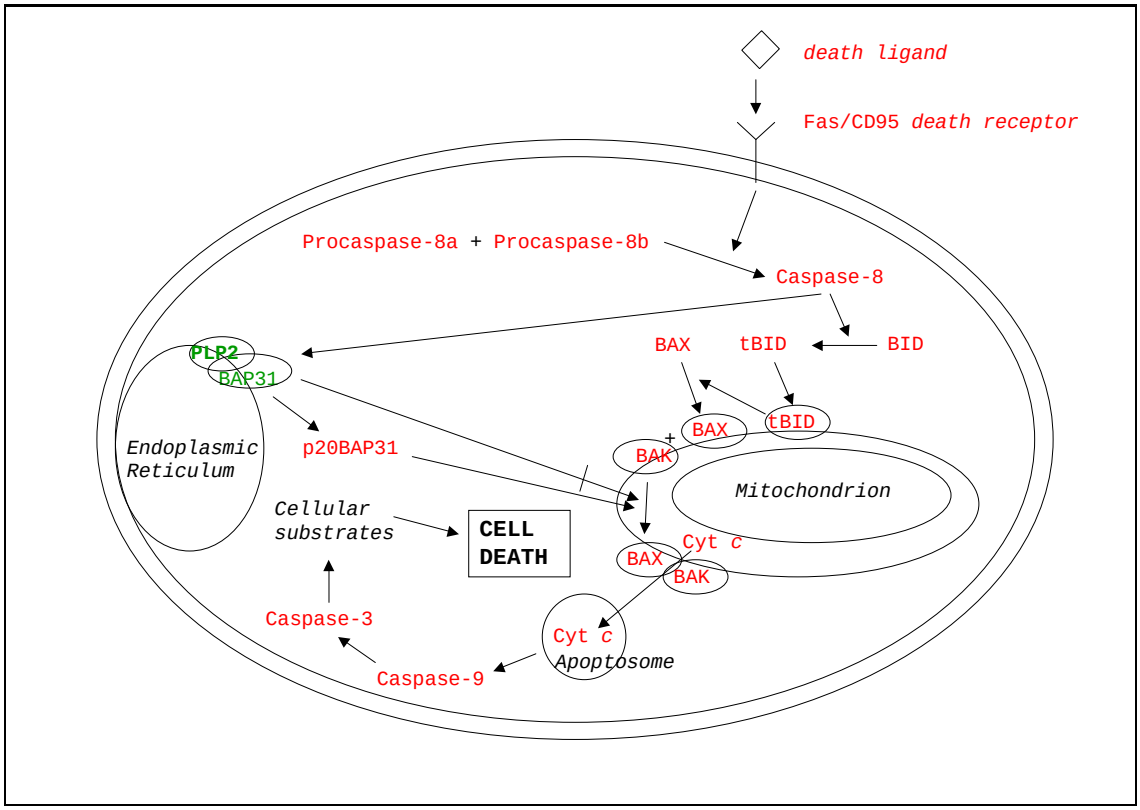



Figure 31: **Overview on apoptotic pathways and the role of the PLP2 protein therein [68].** In this scheme, anti-apoptotic proteins are denoted in green and pro-apoptotic in red. PLP2 acts in caspase-dependent cell death upon Fas/CD95 activation: By binding of death signaling molecules to death receptors of the Fas/CD95 type on the cell surface, caspase-8 holoenzyme is generated from its precursors procaspase-8a and -8b. Caspase-8 cleaves the pro-apoptotic BCL-2 family member BID to tBID, and tBID inserts into the outer mitochondrial membrane. tBID recruits BAX into the mitochondrial membrane and induces the oligomerization of BAX and BAK. The fusion of the BCL-2 members BAX and BAK generates a pore through which cytochrome *c* flows out of the mitochondrion. It becomes an integral constituent of the apoptosome and thus promotes activity of caspase-9. Caspase-9 activates caspase-3, which paves a direct way to cell death. On this track, an external death signal, binding to a surfacial death receptor, leads to the death of the cell. The role for PLP2 in apoptosis comes from the following pathway: Upon activation of caspase-8 via the Fas receptor, BAP31, an integral membrane protein located in the heterooligomeric BAP complex of the endoplasmic reticulum, is cleaved to give rise to p20BAP31. p20BAP31 is an apoptosis promoting protein. Contrastingly, uncleaved BAP31 is an anti-apoptotic factor, inhibiting oligomerization of BAX and BAK in the mitochondrial membrane and thus neutralizing upstream death signals. As an interaction partner of BAP31 [89], PLP2 thus has anti-apoptotic features.

## 4.5 Outlook

### 4.5.1 Expression Level of the *PLP2* Gene

To substantiate the found 2.7 fold reduction of *PLP2* expression in $PLP2^{-188(A)}$ patients in comparison to $PLP2^{-188(C)}$ controls, application of another gene expression tracing technique makes sense. To date, the real-time Reverse Transcriptase Polymerase Chain Reaction (real-time RT-PCR) method is as widely appreciated as Northern blotting.

In this approach, all mRNAs from the total RNA isolate are first transcribed into complementary DNA (cDNA) by help of oligo-dT primers, binding to the polyadenylated 3' end of all eucaryotic mRNAs, and the enzyme Reverse Transcriptase (RT). The RNA strand of the resulting RNA-DNA heteroduplex is then digested by an RNase and replaced by an DNA strand, sythesized by a DNA polymerase that fills up random hexamers. So, each molecule of mRNA will be represented by a double-stranded cDNA.

Quantitation of the amount of cDNAs present in the samples can than be done by means of PCR amplification of the cDNA – and thus the gene under investigation – by primers specific for this cDNA. As amplification occurs absolutely exponentially, tiny differences in the number of starting cDNAs will result in huge differences in the numbers of PCR products from these templates in the end, giving this technique its high sensitivity. Furthermore, very low gene expression levels can be detected by increasing the number of cycles. Quantitation of the amount of cDNA present is done by fluorescent dyes that either bind unspecifically to the DNA and fluoresce only when bound (as SYBR Green), or that bind the cDNA in a sequence-specific manner and fluoresce only when bound (molecular beacons) or when set free after having bound (TaqMan probes) [38].

Generally, determination of gene expression levels via real-time RT-PCR or Northern blotting attacks at an early step from a gene to a protein. Every regulatory effect that occurs after the gene has been transcribed into mRNA cannot be detected, although most regulation of gene expression usually occurs on the transcriptional level [17]. Nevertheless, an endpoint analysis, measuring the amount of protein of interest instead of the amount of RNA belonging to the protein of interest, is a good oppurtunity to confirm results obtained on other experimental pathways.

Such expression analysis on the protein level can be achieved if antibodies against the protein of interest are available. In an Enzyme-Linked Immunosorbent Assay (ELISA), the amount of protein of interest present in a total cell protein isolate can be quantified by first immobilizing the isolate to the surface of plate cavities, which occurs by unspecific binding of the proteins to the plate material. All remaining free cavity surface is blocked by adding irrelevant protein. Then, a primary antibody, raised in a host organism against an epitope of the protein of interest, is added, and unbound antibody is washed away. Finally, a second antibody is added, that antibody recognizing all proteins of the host and being coupled to an enzyme catalyzing the production of a colored product from an uncolored precursor. Thus, the more protein of interest initially was present, the more colored product will be produced in the end. Referencing of the amount of color can be

done with respect to the protein product of a housekeeping gene. As for me, I had to stop expression analysis on the RNA level, because there is no PLP2 antibody yet.

### 4.5.2 Frequency of the *PLP2* Promoter Mutation

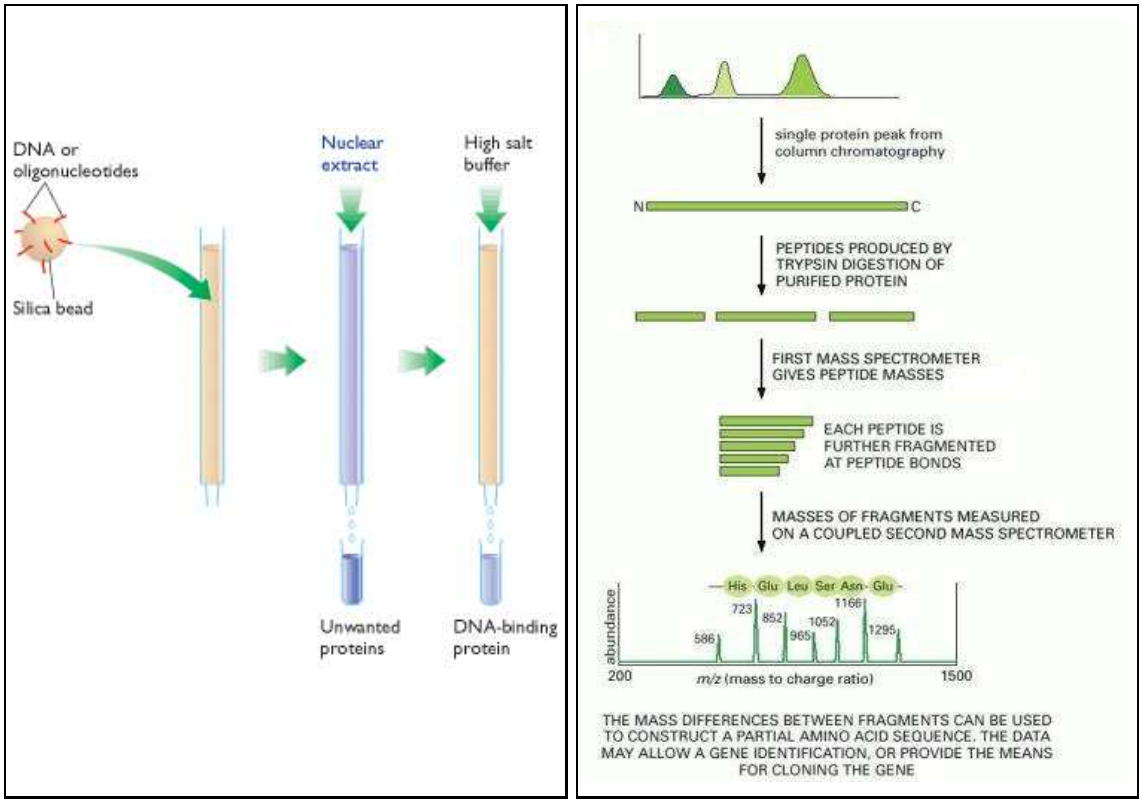
Although the control panels analyzed showed clearly that the A-variant of the *PLP2* promoter is restricted to the common population and to people affected by mental retardation, confirmation of these results by use of much larger control pools is reasonable step. If $PLP2^{-188(A)}$ is found in 26 individuals out of 1000 in total, this finding will outweigh finding the same frequency of 2.6% in 3 individuals out of 115 in total.

### 4.5.3 Electrophoretic Mobility Shift Assay

In the functional investigation of the *PLP2* expression, the observed protein-binding to a *PLP2* promoter fragment argues in favor of a functional relevance of the polymorphism. However, as the complex was too large to be investigated with the EMSAs, differentiation between the respective protein-binding capacities of the $PLP2^{-188(C)}$ and the $PLP2^{-188(A)}$ oligonucleotides was not possible.

In order to circumvent the unavoidable problems of resolving large complexes on gels, comparative reporter gene analysis would be a suitable approach. Additionally, this technique allows for working *in vivo*. After transfection of a human cell line with a vector containing the fusion product of the Green Fluorescent Protein (GFP) coding sequence and the *PLP2* promoter sequence, expression of *GFP* can directly be quantified by the use of fluorescence microscopy. Introducing the cytosine to adenine change in the promoter sequence should comparably reduce the gene's expression level. Thus, binding of transcription factors to the *PLP2* promoter site under investigation is assayed indirectly.

To identify the binding transcription factors, affinity chromatography will be a first step necessary to enrich the transcription factor, and subsequent mass spectrometry might reveal the identity of the protein (see figure 32).




(a) Transcription factor isolation [12] (b) ... and identification [1]

Figure 32: **Principle of identifying transcription factor binding to the *PLP2* promoter element by mass spectrometry.** At first, the transcription factor has to be isolated. (A) Oligonucleotides exhibiting the binding site from the *PLP2* promoter are immobilzed on the surface of silica beads and packed into chromatography columns. Passing a nuclear extract of proteins through the column, the transcription factor of interest will specifically bind the immobilzed oligonucleotide and be retained in the column. Elution with a high salt buffer yields the isolated factor. The isolated protein can then be sequenced by mass spectrometry: (B) The isolated protein is digested into peptides by protease treatment. The masses of these peptides can directly be determined by the spectrometer. Cleavage of single amino acids from the peptides results in mass decrease corresponding to the mass of the cleaved amino acid. In this manner, the amino acid sequence of the peptides is determined step by step. Aligning this sequence with protein databases identifies the transcription factor. If more than one result is found, sequence length has to be increased by using another protease in the first step and assembling the peptides according to their sequence overlaps. If nothing is known, things are slowed down and fundamental molecular biological research has to be done.

## 4.6 Conclusion

The aim of this diploma thesis was to elucidate functional aspects of a cytosine to adenine change that was found in the promoter region of the *PLP2* gene. Concretely, I wanted to answer three questions:

1. Has the base change any influence on the gene expression level?
2. How often does it occur in the common population, and is there a different frequency in people with at least average intelligence quotient?
3. Is the C to A change located in a functional site of the *PLP2* promoter, i.e. is it bound by transcription factors?

The first two questions could be answered without restrictions: Yes, *PLP2* expression is reduced to almost one third in the A-variant as compared to the $PLP2^{-188(C)}$ genotype present in the RefSeq [64], and the frequency of $PLP2^{-188(A)}$ was found to be 2.6% on average in the males of the common population, but zero in people with above average intelligence quotient.

However, the third question was not answered completely: Although I could demonstrate protein binding to the promoter region under investigation, I could not demonstrate any differences between the $PLP2^{-188(C)}$ and $PLP2^{-188(A)}$ genotype. The electrophoretic mobility shift assay used for this purpose was not suitable to resolve those high molecular weight protein-DNA complexes that formed. Therefore, future research is needed to find out about the role of $PLP2^{-188(A)}$ in the mental retardation phenotype.

# Appendix

## Expression Level of the *PLP2* Gene

### RNA Quantitation

After isolation of RNA from the EBV transformed lymphoblastoid cell lines, spectrophotometrical quantitation of the RNA and protein concentration in the samples, necessary for the subsequent blotting procedure, was done (see table 29).

Table 29: **Concentration of all isolated RNAs**

| Cell line | | RNA dilution factor | $OD_{260\,nm}$ | $OD_{280\,nm}$ | $OD_{260\,nm}/OD_{280\,nm}$ | $c_{RNA}$ [µg/ml] | $V_{RNA}$ for 15 µg [µl] | $V_{H20}$ to 16.7 µl [µl] |
|---|---|---|---|---|---|---|---|---|
| 0038/98 | | 100 | 0.464 | 0.299 | 1.55 | 1856 | 8.1 | 8.6 |
| 0066/98 | | 100 | 0.290 | 0.203 | 1.43 | 1160 | 12.9 | 3.8 |
| 0429/99 | | 100 | 0.714 | 0.431 | 1.66 | 2856 | 5.3 | 11.4 |
| 0491/99 | | 100 | 0.609 | 0.379 | 1.61 | 2436 | 6.2 | 10.5 |
| 0524/00 | patients | 100 | 0.658 | 0.408 | 1.61 | 2632 | 5.7 | 11.0 |
| 1244/03 | | 100 | 0.538 | 0.335 | 1.61 | 2152 | 7.0 | 9.7 |
| 1245/03 | | 100 | 0.482 | 0.311 | 1.55 | 1928 | 7.8 | 8.9 |
| 0400/99 | | 100 | 0.609 | 0.384 | 1.59 | 2436 | 6.2 | 10.5 |
| 0428/99 | | 100 | 0.593 | 0.375 | 1.58 | 2372 | 6.3 | 10.4 |
| 90/98 | | 100 | 0.544 | 0.346 | 1.57 | 2176 | 6.9 | 9.8 |
| 91/98 | | 100 | 0.816 | 0.507 | 1.61 | 3264 | 4.6 | 12.1 |
| 164/98 | controls | 100 | 0.548 | 0.346 | 1.58 | 2192 | 6.8 | 9.9 |
| 229/99 | | 100 | 0.504 | 0.318 | 1.58 | 2016 | 7.4 | 9.3 |
| 4981 | | 100 | 0.613 | 0.393 | 1.56 | 2452 | 6.1 | 10.6 |

One point of the quantitation results shall be adressed more in detail here. The quotient of RNA to protein concentration yielded values of about 1.6, indicating a poor purity of the RNA preparation, for quotients of about 1.8 or 2 are considered to represent pure DNA or RNA isolations, respectively. However, one has to bear in mind that base composition influences this quotient, for not all nucleotides absorbe at exactly the same wavelength. Of course, this problem is especially important when measuring short oligonucleotides that do not have even distributions of all four bases. More interestingly in this case here, the absorption maximum of the DNA at 260 nm is much higher than that of the proteins at 280 nm [33], which means that protein-due sample impurities are suppressed by the high DNA/RNA extinction. For this reason, one can much better quantify the contamination of a protein sample by DNA/RNA than the contamination of a DNA sample by proteins – but that is exactly the standard application! Nevertheless, there is no reason for not quantifying the concentration of nucleic acids by the use of the $OD_{260\,nm}$ value, not caring much about the $\frac{OD_{260\,nm}}{OD_{280\,nm}}$ value.

### Detailed Data for *PLP2* Expression Levels

In the results section, only the summarized *PLP2* expression level for both the $PLP2^{-188(A)}$ patient group and the $PLP2^{-188(C)}$ control group is illustrated. Detailed results for each individual and the Student's $t$ test analysis can be found here.

Table 30: **Quantitation of Northern blot bands and Student's $t$ test of normalized *PLP2* expression in patients and controls**

| Sample | Peak area | | Peak area ratio | Mean of peak | Standard deviation |
|---|---|---|---|---|---|
| | PLP2 | beta-actin | PLP2 : beta-actin | area ratio | of peak area ratio |
| 0038/98 | 2387 | 9035 | 0.26 | 0.25 | 0.13 |
| 0066/98 | 3956 | 8056 | 0.49 | | |
| 0429/99 | 1529 | 10856 | 0.14 | | |
| 0491/99 | 1972 | 6053 | 0.33 | | |
| 0524/00 | 2127 | 5771 | 0.37 | | |
| 1244/03 | 938 | 6795 | 0.14 | | |
| 1245/03 | 1333 | 9299 | 0.14 | | |
| 0400/99 | 1842 | 7639 | 0.24 | | |
| 0428/99 | 1337 | 9985 | 0.13 | | |
| 90/98 | 8065 | 13085 | 0.62 | 0.68 | 0.11 |
| 91/98 | 8511 | 10617 | 0.80 | | |
| 164/98 | 6344 | 9681 | 0.66 | | |
| 229/99 | 9033 | 11782 | 0.77 | | |
| 4981 | 9074 | 16830 | 0.54 | | |

| Student's $t$-test assuming equal variances | Peak area ratios | |
|---|---|---|
| | Patients | Controls |
| Mean | 0.249653365 | 0.675826388 |
| Variance | 0.015936917 | 0.011679893 |
| Observations | 9 | 5 |
| Pooled variance | 0.014517909 | |
| Hypothesized mean difference | 0 | |
| Number of degrees of freedom (df) | 12 | |
| $t$ Statistic | -6.341268832 | |
| P(T<=t) one-tail | 1.8561E-05 | |
| $t$ critical one-tail | 1.782286745 | |
| P(T<=t) two-tail | 3.71221E-05 | |
| $t$ critical two-tail | 2.178812792 | |

## Frequency of the *PLP2* Promoter Mutation

### Optimization of PCR Amplification of the Promoter Region

The determination of the frequency of the $PLP2^{-188(A)}$ variant was done by a sequencing approach. Therefore, PCR products, covering the site of the change, had to be generated. As a template, not the genomic DNA of the controls was used, but the product of a whole-genome amplification process (see figure 33). This process uses Phi29 DNA polymerase to generate template copies of up to 70 kb [19]. These copies are long enough to represent the entire template genome in a PCR amplification, since PCR products mostly are no longer than 0.5 kb. The error rate of the enzyme is approximately 100 times lower than that of *Taq* DNA polymerase and lies in the range of $10^{-6} - 10^{-7}$.

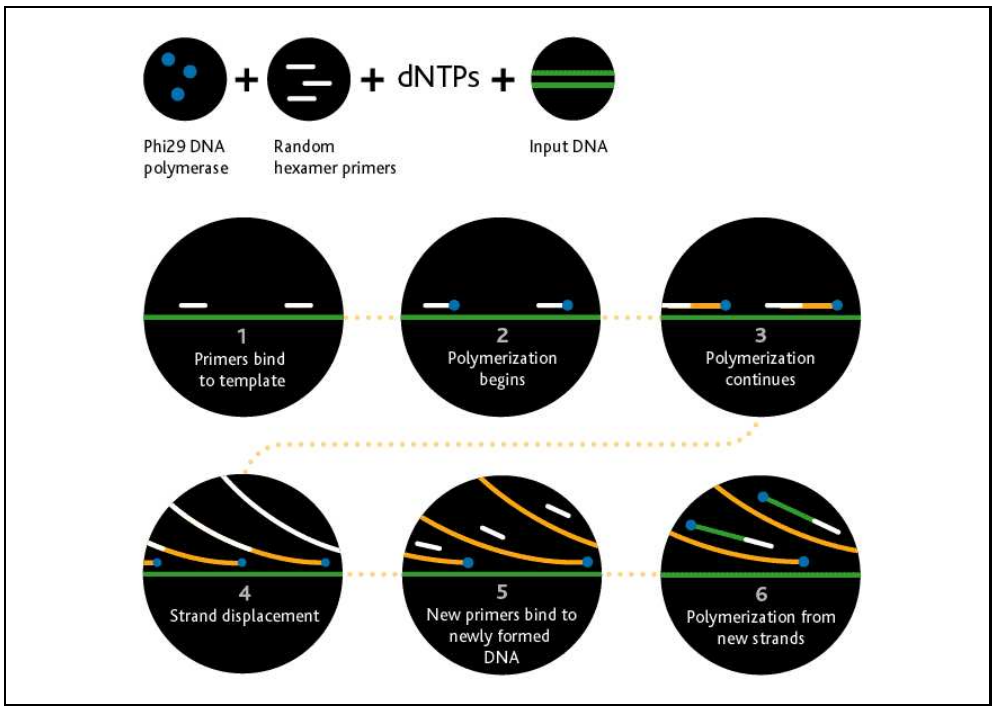



Figure 33: **Whole-genome amplication by Phi29 DNA polymerase [30].** In an isothermal process, Phi29 DNA polymerase elongates random hexamer primers, bound at various sites to the template DNA (1, 2 and 3). If an adjacent polynucleotide stretch is reached, the polymerase displaces it and continues incorporation of nucleotides (4). In this manner, many copies of the original template strand are synthesized (5 and 6), having an average size of 30 – 40 kb. Consequently, for all PCR purposes, representations of the template sequence are present.

However, this genomic amplification process observably reduces the efficiency of PCRs, as can be seen when compairing the PCR success in the non-whole-genome amplified control pool 2 to the pools 1 and 3. Quantitatively, use of 60 ng of genomic DNA yielded, under the same conditions, on average twice as much PCR product as using 200 ng of amplified genomic DNA, resulting in a sixfold reduction of PCR efficiency. The reason for this finding might lie within the gaps between discontinued strands: such gaps prevent successful PCRs, although the mass of template DNA would be high enough to facilitate

amplification when using genomic DNA. Nevertheless, the Phi29 driven amplification of whole genomes makes sense where only low amounts of template are present and elaborate DNA purifications shall be prevented. But if the starting material is of low quality, as in the poor DNA preparations from blood in control pool 3, whole-genome amplification cannot work wonders.

In the very beginning of the determination of the *PLP2* A-variant frequency, I tried to optimize PCR conditions for use with whole-genome amplified DNA to minimize the number of repeats. At first, I took the default cycling conditions suggested by Qiagen [65]:

1. Initial denaturation: 94 °C for 3 min
2. Amplification
   (a) Denaturation: 94 °C for 30 s
   (b) Primer annealing: 55 °C for 30 s
   (c) Primer extension: 72 °C for 1 min
3. Thermal cycling: 29x step 2
4. Final extension: 72 °C for 10 min
5. Program end: Hold 8 °C endlessly

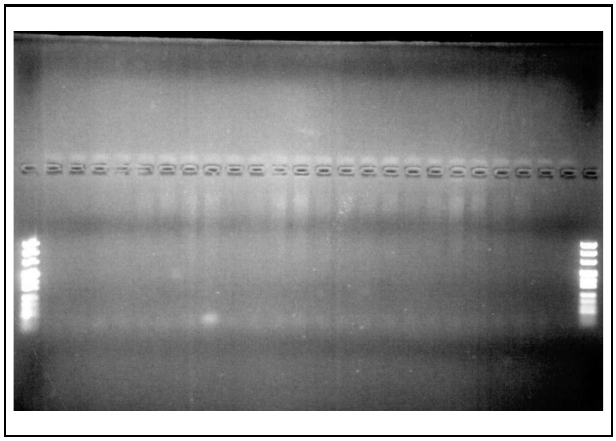
These conditions were used with six differently composed PCR reaction mixes on identical templates. By default, Qiagen's 10x PCR buffer contains magnesium chloride at a concentration of 1.5 mM. Additionally, I prepared two reaction mixes with magnesium chloride concentrations of 3 mM and 5 mM, for PCR amplification of difficult targets might be facilitated by higher concentrations of the $Mg^{2+}$ ion, the ion acting as a cofactor to the DNA polymerase and also stabilizing the structure of the DNA.

For difficult PCRs, Qiagen provides another substance that potentially aids in PCR amplification, the substance called "Q Solution" and consisting of a highly viscous polymer that influences the melting behaviour of the DNA and thereby enhances especially amplification of templates with high GC content [65]. The *PLP2* promoter fragment contains G and C bases to an extent of 71% [55].

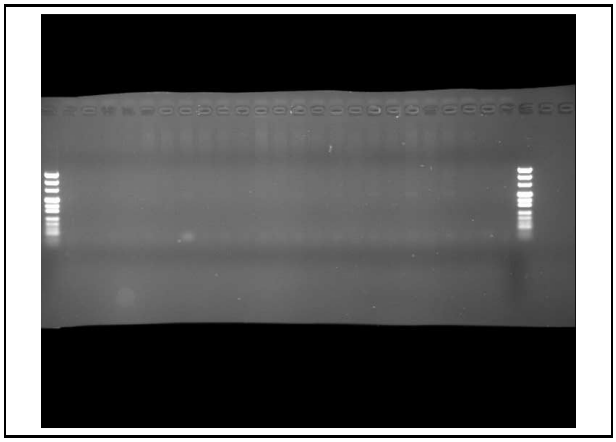
Taking everything together, I finally tested six different reaction mixes: three different $MgCl_2$ concentrations, each with and without Q Solution. As a template for all amplifications, the templiphied DNA of the first eleven females of the ECACC control plate 1 was used. In the end, only the highest concentration of magnesium chloride and the addition of Q Solution yielded PCR products at all (see figure 34). Unfortunately, these products were that weak that a further improvement of the thermal cyling conditions was necessary.

In this next step, I tried alteration of four other parameters:

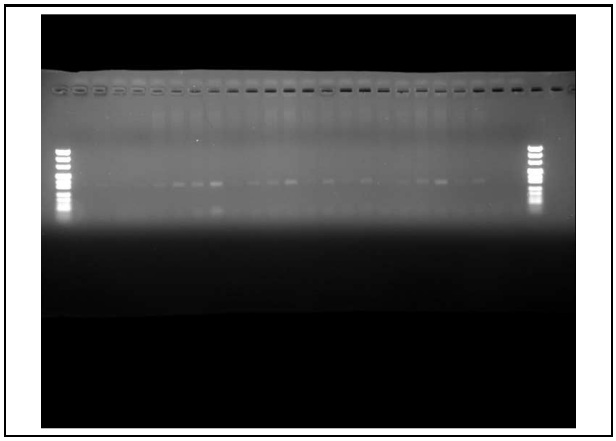
1. **Addition of betain:** Betain can help to decrease the amount of energy that is needed to melt the template's DNA strands [18]. Recommended are concentrations between 0.5 M and 2 M, used was a betain concentration of 1 M.



(a) PCR result with 1.5 mM and 3 mM $MgCl_2$



(b) PCR result with 5 mM $MgCl_2$ and 1.5 mM $MgCl_2$ + Q Solution



(c) PCR result with 3 mM $MgCl_2$ + Q Solution and 5 mM $MgCl_2$ + Q Solution

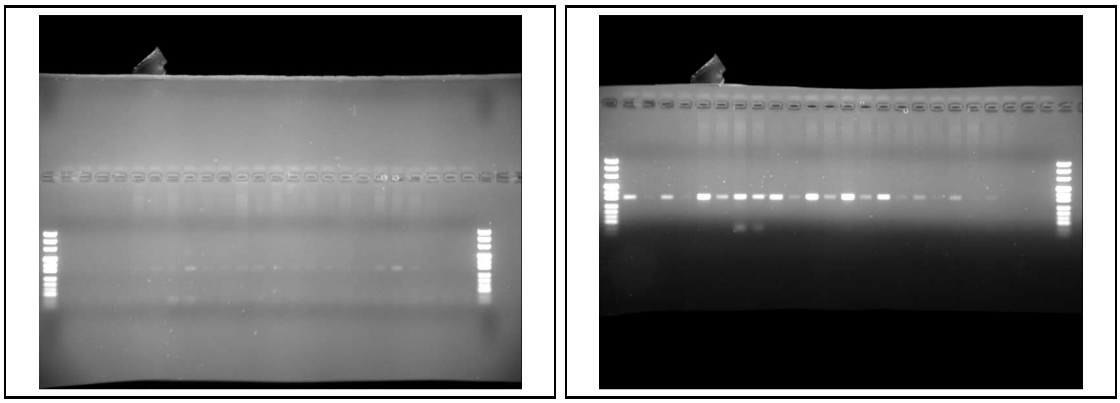
Figure 34: **Result of the first PCR optimization step.** All PCR reactions were loaded to a 1.5% agarose gel using a multichannel pipet, resulting in the presence of coherent samples in every second lane (e.g., DNA of female 1 in lane 1, DNA of female 2 in lane 3 etc.). On the last two lanes, I always loaded the negative controls for each reaction mix, and in the first and last lane, one can find pUC Mix Marker 8. Only the last two reaction mixes show some weak PCR products (see figure 34(c)), with the higher $MgCl_2$ concentration resulting in slightly higher concentrated products. Therefore, a reaction mix, containing 5 mM $MgCl_2$ and Q Solution, was used for further improvement.

2. **Use of less DNA:** As the *Taq* polymerase, DNA also binds $Mg^{2+}$, decreasing the free amount of this ion such that the amount needed by the polymerase might not be present, resulting in a lower activity of the enzyme. To prevent that, only 1 µl of the templiphied template DNA was used, the concentration of the templiphied DNA being around 100 ng/µl.

3. **Prolonging thermal cycling:** The reaction mix with the highest $MgCl_2$ concentration and addition of Q Solution showed some weak PCR products whose amounts might be increased by simply increasing the number of thermal cycles on the template. On the background of the known poor results, I checked 35 instead of previously 30 cycles, ideally increasing the amount of product by a factor of $2^5 = 32$.

4. **Extending the primer annealing time:** This challenging template might require a longer primer annealing time. Instead of 30 s, 60 s was tried.

The output of this testing was, that only the increase of cycle number lead to a better result (see figure 35). That is why I also tried 40, 45 and 50 cycles (not shown). The best result could be obtained with 45 cycles. The background was not increased when increasing the cycle number.



(a) PCR result with 1 M betain and only 100 ng DNA

(b) PCR result with five more cycles and an annealing time of 60 s

Figure 35: **Result of the second PCR optimization step.** Only the additon of five more cycles lead to a notable increase in the amount of PCR product.

Taking everything together, the difficulty of the PCR amplification has not only to be seen in the context of the template's whole-genome amplification, but also with respect to the composition of the amplicon itself. Having a GC content of 71%, melting of the template DNA in the area of the *PLP2* promoter amplicon is largely inhibited. Qiagen's Q Solution influences this melting behaviour in an undefined way, but is mentioned together with dimethyl sulfoxide [65], which implies that it destabilizes the hydrogen bonding of paired DNA bases to improve melting of the double-strand – a prerequisite for primer binding and subsequent elongation. The high amount of $Mg^{2+}$ ions necessary is thought to be due to the high amount of DNA used for amplification of whole-genome amplified DNA.

### Analysis of the PCR Products on the DHPLC System

The PCR products of the *PLP2* promoter amplicon in the control pools 1 und 2, both representing the common population, were analyzed on the Denaturing HPLC (DHPLC) machine. This pre-sequencing mutation screening was a good help to reduce the number of candidate PCR products.

For PCR product mixing, pipetting schemes can be found in the subtables of table 31. The plates for the males from control pool 1 and 2 contain also a positive control. For the males, such a positive control is a mix of two PCR products from which is known that one contains the mutation of interest (corresponding to mixing one affected and one intact X chromosome). Consequently, for the females, a mix of one $PLP2^{-188(\mathrm{A})}$ to three $PLP2^{-188(\mathrm{C})}$ PCR products has to be prepared.

The detailed results from the DHPLC analysis and the sequencing of the corresponding PCR products can be found on page 96 and the following.

| A | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | DNA A1, ♂ | DNA A2, ♂ | DNA A3, ♂ | DNA A4, ♂ | DNA A5, ♂ | DNA A6, ♂ | | | | | | |
| | DNA A7, ♂ | DNA A8, ♂ | DNA A9, ♂ | DNA A10, ♂ | DNA A11, ♂ | DNA A12, ♂ | | | | | | |
| B | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| | DNA B1, ♂ | DNA B2, ♂ | DNA B3, ♂ | DNA B4, ♂ | DNA B5, ♂ | DNA B6, ♂ | | | | | | |
| | DNA B7, ♂ | DNA B8, ♂ | DNA B12, ♂ | DNA B10, ♂ | DNA B11, ♂ | DNA B12, ♂ | | | | | | |
| C | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| | DNA C1, ♂ | DNA C2, ♂ | DNA C3, ♂ | DNA C4, ♂ | DNA C5, ♂ | DNA C6, ♂ | | | | | | |
| | DNA C7, ♂ | DNA C8, ♂ | DNA C9, ♂ | DNA C10, ♂ | DNA C11, ♂ | DNA C12, ♂ | | | | | | |
| D | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 3 | 10 | 11 | 12 |
| | DNA D1, ♂ | DNA D2, ♂ | DNA D3, ♂ | DNA D4, ♂ | DNA D5, ♂ | DNA D6, ♂ | | | | | | |
| | DNA D7, ♂ | DNA D8, ♂ | DNA D9, ♂ | DNA D10, ♂ | DNA D11, ♂ | DNA D12, ♂ | | | | | | |
| E | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| | DNA E1, ♂ | DNA E2, ♂ | DNA E3, ♂ | DNA E4, ♂ | DNA E5, ♂ | DNA E6, ♂ | | | | | | |
| | DNA E7, ♂ | DNA E8, ♂ | DNA E9, ♂ | DNA E10, ♂ | DNA E11, ♂ | DNA E12, ♂ | | | | | | |
| F | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| | DNA F1, ♂ | DNA F2, ♂ | DNA F3, ♂ | DNA F4, ♂ | DNA F5, ♂ | DNA F6, ♂ | | | | | | |
| | DNA F7, ♂ | DNA F8, ♂ | DNA F9, ♂ | DNA F10, ♂ | DNA F11, ♂ | DNA F12, ♂ | | | | | | |
| G | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| | DNA G1, ♂ | DNA G2, ♂ | DNA G3, ♂ | DNA G4, ♂ | DNA G5, ♂ | DNA G6, ♂ | | | | | | |
| | DNA G7, ♂ | DNA G8, ♂ | DNA G9, ♂ | DNA G10, ♂ | DNA G11, ♂ | DNA G12, ♂ | | | | | | |
| H | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| | DNA H1, ♂ | DNA H2, ♂ | DNA H3, ♂ | | | Positive control | | | | | | |
| | DNA H4, ♂ | DNA H5, ♂ | DNA H6, ♂ | | | | | | | | | |

(a) Pooling scheme for the males of control pool 1

| A | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | DNA A1, ♀ | DNA A2, ♀ | DNA A3, ♀ | DNA A4, ♀ | DNA A5, ♀ | DNA A6, ♀ | | | | | | |
| | DNA A7, ♀ | DNA A8, ♀ | DNA A9, ♀ | DNA A10, ♀ | DNA A11, ♀ | DNA A12, ♀ | | | | | | |
| B | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| | DNA B1, ♀ | DNA B2, ♀ | DNA B3, ♀ | DNA B4, ♀ | DNA B5, ♀ | DNA B6, ♀ | | | | | | |
| | DNA B7, ♀ | DNA B8, ♀ | DNA B12, ♀ | DNA B10, ♀ | DNA B11, ♀ | DNA B12, ♀ | | | | | | |
| C | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| | DNA C1, ♀ | DNA C2, ♀ | DNA C3, ♀ | DNA C4, ♀ | DNA C5, ♀ | DNA C6, ♀ | | | | | | |
| | DNA C7, ♀ | DNA C8, ♀ | DNA C9, ♀ | DNA C10, ♀ | DNA C11, ♀ | DNA C12, ♀ | | | | | | |
| D | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 3 | 10 | 11 | 12 |
| | DNA D1, ♀ | DNA D2, ♀ | DNA D3, ♀ | DNA D4, ♀ | DNA D5, ♀ | DNA D6, ♀ | | | | | | |
| | DNA D7, ♀ | DNA D8, ♀ | DNA D9, ♀ | DNA D10, ♀ | DNA D11, ♀ | DNA D12, ♀ | | | | | | |
| E | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| | DNA E1, ♀ | DNA E2, ♀ | DNA E3, ♀ | DNA E4, ♀ | DNA E5, ♀ | DNA E6, ♀ | | | | | | |
| | DNA E7, ♀ | DNA E8, ♀ | DNA E9, ♀ | DNA E10, ♀ | DNA E11, ♀ | DNA E12, ♀ | | | | | | |
| F | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| | DNA F1, ♀ | DNA F2, ♀ | DNA F3, ♀ | DNA F4, ♀ | DNA F5, ♀ | DNA F6, ♀ | | | | | | |
| | DNA F7, ♀ | DNA F8, ♀ | DNA F9, ♀ | DNA F10, ♀ | DNA F11, ♀ | DNA F12, ♀ | | | | | | |
| G | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| | DNA G1, ♀ | DNA G2, ♀ | DNA G3, ♀ | DNA G4, ♀ | DNA G5, ♀ | DNA G6, ♀ | | | | | | |
| | DNA G7, ♀ | DNA G8, ♀ | DNA G9, ♀ | DNA G10, ♀ | DNA G11, ♀ | DNA G12, ♀ | | | | | | |
| H | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| | DNA H1, ♀ | DNA H2, ♀ | DNA H3, ♀ | DNA H4, ♀ | DNA H5, ♀ | DNA H6, ♀ | | | | | | |
| | DNA H7, ♀ | DNA H8, ♀ | DNA H9, ♀ | DNA H10, ♀ | DNA H11, ♀ | DNA H12, ♀ | | | | | | |

(b) Pooling scheme for the females of control pool 1

| A | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Female | DNA 1, ♀ | DNA 49, ♀ | DNA 64, ♀ | DNA 2, ♀ | DNA 51, ♀ | DNA 66, ♀ | | | | | | |
| | DNA 37, ♀ | DNA 53, ♀ | DNA 67, ♀ | DNA 39, ♀ | DNA 55, ♀ | DNA 68, ♀ | | | | | | |
| B | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| Female | DNA 41, ♀ | DNA 57, ♀ | DNA 69, ♀ | DNA 43, ♀ | DNA 58, ♀ | DNA 70, ♀ | | | | | | |
| | DNA 45, ♀ | DNA 60, ♀ | DNA 291, ♀ | DNA 47, ♀ | DNA 62, ♀ | DNA 62, ♀ | | | | | | |
| C | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| Male | DNA 3, ♂ | DNA 11, ♂ | DNA 19, ♂ | DNA 4, ♂ | DNA 12, ♂ | DNA 22, ♂ | | | | | | |
| | DNA 5, ♂ | DNA 13, ♂ | DNA 25, ♂ | DNA 6, ♂ | DNA 14, ♂ | DNA 28, ♂ | | | | | | |
| D | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 3 | 10 | 11 | 12 |
| Male | DNA 7, ♂ | DNA 15, ♂ | DNA 31, ♂ | DNA 8, ♂ | DNA 16, ♂ | DNA 33, ♂ | | | | | | |
| | DNA 9, ♂ | DNA 17, ♂ | DNA 35, ♂ | DNA 10, ♂ | DNA 18, ♂ | DNA 4981, ♂ | | | | | | |
| E | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| | | | | | | | | | | | | |
| | | | | | | | | | | | | |
| F | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| | | | | | | | | | | | | |
| | | | | | | | | | | | | |
| G | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| | | | | | | | | | | | | |
| | | | | | | | | | | | | |
| H | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| | | | | | | | | | | | **Positive control females (20 µl)** Pool 1 males G9+10+11+12 | **Positive control males (20 µl)** Pool 1 males G4+5 |
| | | | | | | | | | | | | |

(c) Pooling scheme for control pool 2

Table 31: **Pooling schemes for PCR products of the control pools 1 and 2**

### Sequencing of the Candidate PCR Products

**Determination of the Concentration of the *PLP2* PCR Products** was done by referencing the intensity of the band of the individual PCR product to the intensity of the bands of the quantitative DNA ladder. The volume of PCR product needed for each sequencing reaction can then be calculated according to section 2.2.3.2 on page 29. The result of these calculations is illustrated in tables 32 and 33.

Table 32: **Concentrations of the *PLP2* PCR products in control pool 1 and 2**

(a) Control pool 1

| Sample | Gender | Amplicon Name | Amplicon Size [bp] | $m_{DNA}$ [ng] | $c_{DNA}$ [ng / 100 bp] | $V_{Gel\,run}$ [µl] | $V_{Seq.}$ [µl] | $V_{Seq.,\,rounded}$ [µl] |
|---|---|---|---|---|---|---|---|---|
| DNA E3 | | 0579-3 R+S | 326 | 30 | 9.20 | 3 | 0.652 | 1 |
| DNA E9 | | 0579-3 R+S | 326 | 30 | 9.20 | 3 | 0.652 | 1 |
| DNA G4 | | 0579-3 R+S | 326 | 30 | 9.20 | 3 | 0.652 | 1 |
| DNA G10 | | 0579-3 R+S | 326 | 30 | 9.20 | 3 | 0.652 | 1 |
| DNA H8 | male | 0579-3 R+S | 326 | 10 | 3.07 | 3 | 1.956 | 2 |
| DNA H9 | | 0579-3 R+S | 326 | 10 | 3.07 | 3 | 1.956 | 2 |
| DNA H10 | | 0579-3 R+S | 326 | 10 | 3.07 | 3 | 1.956 | 2 |
| DNA H11 | | 0579-3 R+S | 326 | 10 | 3.07 | 3 | 1.956 | 2 |
| DNA H12 | | 0579-3 R+S | 326 | 10 | 3.07 | 3 | 1.956 | 2 |
| DNA C2 | | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.391 | 0.5 |
| DNA C8 | | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.391 | 0.5 |
| DNA D1 | | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.391 | 0.5 |
| DNA D7 | | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.391 | 0.5 |
| DNA E3 | | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.391 | 0.5 |
| DNA E9 | | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.391 | 0.5 |
| DNA F4 | | 0579-3 R+S | 326 | 10 | 3.07 | 3 | 1.956 | 2 |
| DNA F10 | female | 0579-3 R+S | 326 | 20 | 6.13 | 3 | 0.978 | 1 |
| DNA G2 | | 0579-3 R+S | 326 | 10 | 3.07 | 3 | 1.956 | 2 |
| DNA G4 | | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.391 | 0.5 |
| DNA G5 | | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.391 | 0.5 |
| DNA G8 | | 0579-3 R+S | 326 | 10 | 3.07 | 3 | 1.956 | 2 |
| DNA G10 | | 0579-3 R+S | 326 | 10 | 3.07 | 3 | 1.956 | 2 |
| DNA G11 | | 0579-3 R+S | 326 | 20 | 6.13 | 3 | 0.978 | 1 |
| DNA H1 | | 0579-3 R+S | 326 | 10 | 3.07 | 3 | 1.956 | 2 |
| DNA H7 | | 0579-3 R+S | 326 | 10 | 3.07 | 3 | 1.956 | 2 |

(b) Control pool 2

| Sample | Amplicon Name | Amplicon Size [bp] | $m_{DNA}$ [ng] | $c_{DNA}$ [ng / 100 bp] | $V_{Gel\,run}$ [µl] | $V_{Seq.}$ [µl] | $V_{Seq.,\,rounded}$ [µl] |
|---|---|---|---|---|---|---|---|
| DNA A1 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA C1 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA A2 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA C2 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA A3 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA C3 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA A4 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA C4 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA A5 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA C5 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA A6 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA C6 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA B1 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA D1 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA B2 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA D2 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA B3 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA D3 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA B4 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA D4 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA B5 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA D5 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA B6 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA D6 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA E1 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA G1 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA E2 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA G2 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA E3 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA G3 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA E4 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA G4 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA E5 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA G5 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA E6 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA G6 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA F1 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA H1 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA F2 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA H2 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA F3 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA H3 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA F4 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA H4 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA F5 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| DNA H5 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |
| Negative control | 0579-3 R+S | 326 | 0 | 0.00 | 3 | | |
| DNA H6 | 0579-3 R+S | 326 | 50 | 15.34 | 3 | 0.3912 | 0.5 |

Table 33: **Concentrations of the *PLP2* PCR products in the males and females of control pool 3**

(a) Males

| Sample | Gender | Exon Name | Exon Size [bp] | $m_{DNA}$ [ng] | $c_{DNA}$ [ng / 100 bp] | $V_{Gellrun}$ [µl] | $V_{Seq.}$ [µl] | $V_{Seq., rounded}$ [µl] |
|---|---|---|---|---|---|---|---|---|
| DNA A1 | | 0579-3 R+S | 326 | 10 | 3.07 | 5 | 3.26 | 3.5 |
| DNA A2 | | 0579-3 R+S | 326 | 50 | 15.34 | 5 | 0.65 | 1 |
| DNA A3 | | 0579-3 R+S | 326 | 50 | 15.34 | 5 | 0.65 | 1 |
| DNA A4 | | 0579-3 R+S | 326 | 50 | 15.34 | 5 | 0.65 | 1 |
| DNA A5 | | 0579-3 R+S | 326 | 50 | 15.34 | 5 | 0.65 | 1 |
| DNA A6 | | 0579-3 R+S | 326 | 10 | 3.07 | 5 | 3.26 | 3.5 |
| DNA A7 | | 0579-3 R+S | 326 | 50 | 15.34 | 5 | 0.65 | 1 |
| DNA A8 | | 0579-3 R+S | 326 | 50 | 15.34 | 5 | 0.65 | 1 |
| DNA A9 | | 0579-3 R+S | 326 | 10 | 3.07 | 5 | 3.26 | 3.5 |
| DNA A10 | | 0579-3 R+S | 326 | 10 | 3.07 | 5 | 3.26 | 3.5 |
| DNA A11 | | 0579-3 R+S | 326 | 10 | 3.07 | 5 | 3.26 | 3.5 |
| DNA A12 | | 0579-3 R+S | 326 | 10 | 3.07 | 5 | 3.26 | 3.5 |
| DNA B1 | | 0579-3 R+S | 326 | 10 | 3.07 | 5 | 3.26 | 3.5 |
| DNA B2 | | 0579-3 R+S | 326 | 50 | 15.34 | 5 | 0.65 | 1 |
| DNA B3 | | 0579-3 R+S | 326 | | 0.00 | 5 | #DIV/0! | |
| DNA B4 | | 0579-3 R+S | 326 | 10 | 3.07 | 5 | 3.26 | 3.5 |
| DNA B5 | | 0579-3 R+S | 326 | 10 | 3.07 | 5 | 3.26 | 3.5 |
| DNA B6 | male | 0579-3 R+S | 326 | 10 | 3.07 | 5 | 3.26 | 3.5 |
| DNA B7 | | 0579-3 R+S | 326 | 10 | 3.07 | 5 | 3.26 | 3.5 |
| DNA B8 | | 0579-3 R+S | 326 | 10 | 3.07 | 5 | 3.26 | 3.5 |
| DNA B9 | | 0579-3 R+S | 326 | 50 | 15.34 | 5 | 0.65 | 1 |
| DNA B10 | | 0579-3 R+S | 326 | 10 | 3.07 | 5 | 3.26 | 3.5 |
| DNA B11 | | 0579-3 R+S | 326 | 10 | 3.07 | 5 | 3.26 | 3.5 |
| DNA B12 | | 0579-3 R+S | 326 | 50 | 15.34 | 5 | 0.65 | 1 |
| DNA C1 | | 0579-3 R+S | 326 | 10 | 3.07 | 5 | 3.26 | 3.5 |
| DNA C2 | | 0579-3 R+S | 326 | | 0.00 | 5 | #DIV/0! | |
| DNA C3 | | 0579-3 R+S | 326 | 50 | 15.34 | 5 | 0.65 | 1 |
| DNA C4 | | 0579-3 R+S | 326 | 50 | 15.34 | 5 | 0.65 | 1 |
| DNA C5 | | 0579-3 R+S | 326 | 30 | 9.20 | 5 | 1.09 | 1.5 |
| DNA C6 | | 0579-3 R+S | 326 | 30 | 9.20 | 5 | 1.09 | 1.5 |
| DNA C7 | | 0579-3 R+S | 326 | 50 | 15.34 | 5 | 0.65 | 1 |
| DNA C8 | | 0579-3 R+S | 326 | 50 | 15.34 | 5 | 0.65 | 1 |
| DNA C9 | | 0579-3 R+S | 326 | 50 | 15.34 | 5 | 0.65 | 1 |
| DNA C10 | | 0579-3 R+S | 326 | 50 | 15.34 | 5 | 0.65 | 1 |
| DNA C11 | | 0579-3 R+S | 326 | 20 | 6.13 | 5 | 1.63 | 2 |

(b) Females

| Sample | Gender | Exon Name | Exon Size [bp] | $m_{DNA}$ [ng] | $c_{DNA}$ [ng / 100 bp] | $V_{Gellrun}$ [µl] | $V_{Seq.}$ [µl] | $V_{Seq., rounded}$ [µl] |
|---|---|---|---|---|---|---|---|---|
| DNA A1 | | 0579-3 R+S | 326 | 50 | 15.34 | 5 | 0.65 | 1 |
| DNA A2 | | 0579-3 R+S | 326 | 30 | 9.20 | 5 | 1.09 | 1.5 |
| DNA A3 | | 0579-3 R+S | 326 | 10 | 3.07 | 5 | 3.26 | 3.5 |
| DNA A4 | | 0579-3 R+S | 326 | 10 | 3.07 | 5 | 3.26 | 3.5 |
| DNA A5 | | 0579-3 R+S | 326 | 50 | 15.34 | 5 | 0.65 | 1 |
| DNA A6 | | 0579-3 R+S | 326 | 20 | 6.13 | 5 | 1.63 | 2 |
| DNA A7 | | 0579-3 R+S | 326 | 10 | 3.07 | 5 | 3.26 | 3.5 |
| DNA A8 | | 0579-3 R+S | 326 | 10 | 3.07 | 5 | 3.26 | 3.5 |
| DNA A9 | | 0579-3 R+S | 326 | 40 | 12.27 | 5 | 0.82 | 1 |
| DNA A10 | | 0579-3 R+S | 326 | 40 | 12.27 | 5 | 0.82 | 1 |
| DNA A11 | | 0579-3 R+S | 326 | 40 | 12.27 | 5 | 0.82 | 1 |
| DNA A12 | | 0579-3 R+S | 326 | 40 | 12.27 | 5 | 0.82 | 1 |
| DNA B1 | | 0579-3 R+S | 326 | 50 | 15.34 | 5 | 0.65 | 1 |
| DNA B2 | | 0579-3 R+S | 326 | 50 | 15.34 | 5 | 0.65 | 1 |
| DNA B3 | | 0579-3 R+S | 326 | 50 | 15.34 | 5 | 0.65 | 1 |
| DNA B4 | | 0579-3 R+S | 326 | 50 | 15.34 | 5 | 0.65 | 1 |
| DNA B5 | | 0579-3 R+S | 326 | 50 | 15.34 | 5 | 0.65 | 1 |
| DNA B6 | | 0579-3 R+S | 326 | 20 | 6.13 | 5 | 1.63 | 2 |
| DNA B7 | | 0579-3 R+S | 326 | 10 | 3.07 | 5 | 3.26 | 3.5 |
| DNA B8 | | 0579-3 R+S | 326 | 10 | 3.07 | 5 | 3.26 | 3.5 |
| DNA B9 | | 0579-3 R+S | 326 | 10 | 3.07 | 5 | 3.26 | 3.5 |
| DNA B10 | | 0579-3 R+S | 326 | 50 | 15.34 | 5 | 0.65 | 1 |
| Negative control | | 0579-3 R+S | 326 | 0 | 0.00 | 5 | 0.00 | 0 |
| DNA B12 | | 0579-3 R+S | 326 | 10 | 3.07 | 5 | 3.26 | 3.5 |
| DNA C1 | | 0579-3 R+S | 326 | 10 | 3.07 | 5 | 3.26 | 3.5 |
| DNA C2 | female | 0579-3 R+S | 326 | 10 | 3.07 | 5 | 3.26 | 3.5 |
| DNA C3 | | 0579-3 R+S | 326 | 50 | 15.34 | 5 | 0.65 | 1 |
| DNA C4 | | 0579-3 R+S | 326 | 10 | 3.07 | 5 | 3.26 | 3.5 |
| DNA C5 | | 0579-3 R+S | 326 | 10 | 3.07 | 5 | 3.26 | 3.5 |
| DNA C6 | | 0579-3 R+S | 326 | 10 | 3.07 | 5 | 3.26 | 3.5 |
| DNA C7 | | 0579-3 R+S | 326 | 10 | 3.07 | 5 | 3.26 | 3.5 |
| DNA C8 | | 0579-3 R+S | 326 | 10 | 3.07 | 5 | 3.26 | 3.5 |
| DNA C9 | | 0579-3 R+S | 326 | 10 | 3.07 | 5 | 3.26 | 3.5 |
| DNA C10 | | 0579-3 R+S | 326 | 30 | 9.20 | 5 | 1.09 | 1.5 |
| DNA C11 | | 0579-3 R+S | 326 | 10 | 3.07 | 5 | 3.26 | 3.5 |
| DNA C12 | | 0579-3 R+S | 326 | 20 | 6.13 | 5 | 1.63 | 2 |
| DNA D1 | | 0579-3 R+S | 326 | 10 | 3.07 | 5 | 3.26 | 3.5 |
| DNA D2 | | 0579-3 R+S | 326 | 50 | 15.34 | 5 | 0.65 | 1 |
| DNA D3 | | 0579-3 R+S | 326 | 10 | 3.07 | 5 | 3.26 | 3.5 |
| DNA D4 | | 0579-3 R+S | 326 | 10 | 3.07 | 5 | 3.26 | 3.5 |
| DNA D5 | | 0579-3 R+S | 326 | 10 | 3.07 | 5 | 3.26 | 3.5 |
| DNA D6 | | 0579-3 R+S | 326 | 40 | 12.27 | 5 | 0.82 | 1 |
| DNA D7 | | 0579-3 R+S | 326 | 40 | 12.27 | 5 | 0.82 | 1 |
| DNA D8 | | 0579-3 R+S | 326 | 10 | 3.07 | 5 | 3.26 | 3.5 |
| DNA D9 | | 0579-3 R+S | 326 | 40 | 12.27 | 5 | 0.82 | 1 |
| DNA D10 | | 0579-3 R+S | 326 | 10 | 3.07 | 5 | 3.26 | 3.5 |
| DNA D11 | | 0579-3 R+S | 326 | 30 | 9.20 | 5 | 1.09 | 1.5 |
| DNA D12 | | 0579-3 R+S | 326 | 20 | 6.13 | 5 | 1.63 | 2 |
| DNA E1 | | 0579-3 R+S | 326 | 40 | 12.27 | 5 | 0.82 | 1 |
| DNA E2 | | 0579-3 R+S | 326 | 30 | 9.20 | 5 | 1.09 | 1.5 |
| DNA E3 | | 0579-3 R+S | 326 | 50 | 15.34 | 5 | 0.65 | 1 |
| DNA E4 | | 0579-3 R+S | 326 | 10 | 3.07 | 5 | 3.26 | 3.5 |

**Detailed Sequencing Results** are illustrated in the following tables. If a heteroduplex was detected in DHPLC analysis, both corresponding PCR products were sequenced. All empty tabular cells denote the non-finding of heteroduplices. In case that the amount of PCR was heterogeneous, no DHPLC analysis was done and all PCR products were sequenced. As males are hemizygous for the X-linked *PLP2* gene, they cannot be homozygous (hom) nor heterozygous (het) for the promoter variation, whereas females can be. In some samples, sequencing failed (s.f.) even after three trials.

Table 34: **Detailed sequencing result for the males of control pool 1.**

| | *1* | *2* | *3* | *4* | *5* | *6* | *7* | *8* | *9* | *10* | *11* | *12* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *A* | | | | | | | | | | | | |
| *B* | | | | | | | | | | | | |
| *C* | | | | | | | | | | | | |
| *D* | | | | | | | | | | | | |
| *E* | | | wt | | | | | | mut | | | |
| *F* | | | | | | | | | | | | |
| *G* | | | | wt | | | | | | mut | | |
| *H* | | | | | | | | wt | wt | wt | wt | wt |

Table 35: **Detailed sequencing result for the females of control pool 1.**

| | *1* | *2* | *3* | *4* | *5* | *6* | *7* | *8* | *9* | *10* | *11* | *12* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *A* | | | | | | | | | | | | |
| *B* | | | | | | | | | | | | |
| *C* | | het | | | | | | wt | | | | |
| *D* | wt | | | | | | het | | | | | |
| *E* | | | wt | | | | | | het | | | |
| *F* | | | | wt | | | | | | het | | |
| *G* | | het | | het | wt | | | wt | | wt | het | |
| *H* | hom | | | | | | wt | | | | | |

Table 36: **Detailed sequencing result for control pool 2.** The first three columns contain female controls, the last three the males.

| | *1* | *2* | *3* | *4* | *5* | *6* |
|---|---|---|---|---|---|---|
| *A* | | wt | het | | | |
| *B* | | | | | | |
| *C* | | het | wt | | | |
| *D* | | | | | | |
| *E* | | | | | | |
| *F* | | | | | | |
| *G* | | | | | | |
| *H* | | | | | | |

Table 37: **Detailed sequencing result for the males of control pool 3.**

| | *1* | *2* | *3* | *4* | *5* | *6* | *7* | *8* | *9* | *10* | *11* | *12* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *A* | wt | wt | wt | wt | wt | wt | wt | wt | wt | s.f. | s.f. | s.f. |
| *B* | wt | wt | – | wt | wt | s.f. | wt | wt | wt | wt | s.f. | wt |
| *C* | s.f. | – | wt | wt | wt | wt | wt | wt | wt | wt | wt | s.f. |

Table 38: **Detailed sequencing result for the females of control pool 3.**

| | *1* | *2* | *3* | *4* | *5* | *6* | *7* | *8* | *9* | *10* | *11* | *12* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *D* | wt | s.f. | wt | s.f. | wt | wt | s.f. | s.f. | wt | wt | wt | s.f. |
| *E* | wt | het | wt | wt | wt | s.f. | s.f. | s.f. | wt | het | – | s.f. |
| *F* | s.f. | s.f. | s.f. | s.f. | s.f. | s.f. | s.f. | wt | s.f. | s.f. | s.f. | s.f. |
| *G* | s.f. | s.f. | s.f. | s.f. | s.f. | s.f. | s.f. | s.f. | s.f. | s.f. | s.f. | s.f. |
| *H* | wt | s.f. | s.f. | s.f. | – | – | – | – | – | – | – | – |

In the sequencing step, poor-quality PCR products also resulted in poor-quality sequences, as seen primarily in control pool 3, where lots of repetitions were necessary to get a reliable sequence readout. Normally, sequencing of a PCR product that was yielded in low amount is no problem, as its copy number is still much higher than that of the template genome. Hence, amplification of a PCR product by use of primers that were used also in the generation of these PCR products is an easy task, but not if annealing of the primers is inhibited by a high template GC content, precluding thermal double-strand melting into single-strands. Addition of the denaturing agent betain helped here as the denaturing "Q Solution" helped in the previous PCR amplification.

## Electrophoretic Mobility Shift Assay

### A.L.F. DNA Sequencer System Test

At first, I performed my fluorescent EMSAs on Pharmacia's A.L.F. DNA sequencing machine. As this machine is usually used for sequencing, I had to connect an external thermostat which bypassed the internal heating water circulation of the machine, for temperatures above physiological values, which are needed for sequencing, will denature any native protein structure and therefore prevent successful EMSAs. As the machine was not in use since a long time, I had to check if especially the laser emitting subunit was still working. Therefore, I loaded a fluorescein-coupled oligonucleotide to several lanes of a native polyacrylamide gel and waited until the oligonucleotide passed the laser beam. The laser excited the oligonucleotides, and the oligonucleotides themselves emitted light that could be detected by a panel of detectors, one detector being located behind each gel lane. Fortunately, excitation and detection worked very well (see figure 36).

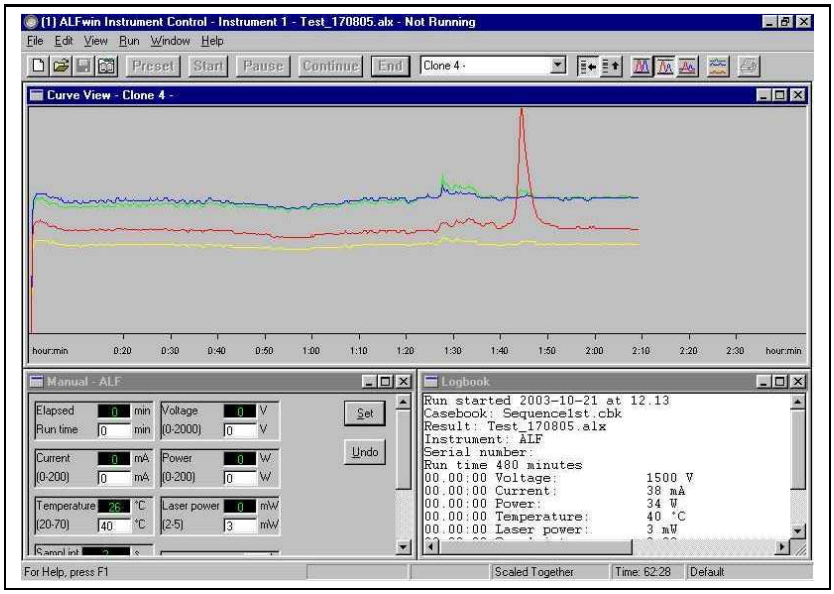



Figure 36: **Testing of the A.L.F. DNA sequencing machine that should be used for fluorescent EMSAs.** With a 6% native polyacrylamide gel, developed at 1500 V and 38 mA, the test oligonucleotide was successfully detected after approximately 110 minutes (see peak in the red line). The ALFWin Instrument Control software groups four subsequent lanes into one clone, for four lanes are necessary to resolve all chain termination reactions of one sequencing reaction, unless one is using a more modern machine that can detect several fluorophores at the same time. For an EMSA, this feature is not necessary.

### Optimization of EMSA Parameters

When trying to optimize EMSAs, the focus has to be on improving the protein DNA binding reaction. Here, one can point out the following critical points: the ionic strength and pH of the binding buffer, the presence of nonionic detergents, glycerol or carrier proteins (e.g. BSA), the presence/absence of divalent cations (e.g. $Mg^{2+}$ or $Zn^{2+}$), the concentra-

tion and type of competitor DNA and the temperature and time of the binding reaction [61]. A binding reaction consists of the following components:

- 1x binding buffer
- 2 µg poly-dIdC (unspecific competitor)
- *x fold specific competitor*
- 3 µg nuclear protein
- *2.5 µg antibody*
- 20 ng labeled oligo
- Fill up to 50 µl with Aqua ad iniectabilia

For optimization, I started with a binding reaction composed as simply as possible. Neither competitors nor antibodies were used. From the remaining components, the binding buffer and the nuclear proteins were considered more closely.

When reading the EMSA protocols that can be found summarized on [58] and in different research articles, various binding buffers can be found. They can be separated into two groups: binding buffers based on the buffering agent tris (tris(hydroxymethyl)aminomethane), and and buffers using HEPES (2-[4-(2-Hydroxyethyl)-1-piperazinyl]-ethanesulfonic acid), both at pH values around 8. Based on that, additions of specific compounds are done. These additives appeared not in all buffers, and their concentrations were also different. From these individual buffers, I built up an "consensus" tris and HEPES binding buffer (see table 39). Both buffers were tested in EMSAs in parallel, and the HEPES buffer showed the strongest protein-DNA bands.

Table 39: **Composition of binding buffers used in research articles or published by work groups or companies on the internet.** All buffer receipts are calculated to be five times concentrated. *Phenylmethylsulfonyl fluoride

| Tris buffers: | *Pierce* [60] | *Cepko Lab* [83] | *Hahn Lab* [84] | *Promega* [63] | *Consensus* |
|---|---|---|---|---|---|
| Tris-HCl | 50 mM (pH=7.5) | 50 mM (pH=7.5) | 100 mM (pH=8) | 50 mM (pH=7.5) | 50 mM (pH=7.5) |
| KCl | 250 mM | 250 mM | 300 mM | | 250 mM |
| $MgCl_2$ | | | 25 mM | 5 mM | 25 mM |
| NaCl | | | | 250 mM | |
| DTT | 5 mM | 2.5 mM | | 2.5 mM | 2.5 mM |
| Glycerol | | 50% | 20% | 20% | 20% |
| BSA | | | 500 µg/ml | | |
| EDTA | | | | 2.5 mM | |
| **HEPES buffers:** | *Meisel Lab* [72] | *Fujifilm* [28] | *Oxford University* [24] | *Washington University* [43] | *Consensus* |
| HEPES | 100 mM | 80 mM (pH=7.9) | 100 mM (pH=7.6) | 100 mM (pH=8.0) | 100 mM (pH=7.9) |
| KCl | 250 mM | 200 mM | 250 mM | 250 mM | 250 mM |
| $MgCl_2$ | | | 5 mM | 5 mM | 10 mM |
| DTT | 5 mM | 20 mM | 2.5 mM | 2.5 mM | 10 mM |
| Glycerol | 50% | 40% | | 20% | 30% |
| EDTA | 5 mM | 0.8 mM (pH=8.0) | | 0.25 mM | 1 mM |
| PMSF* | | 2 mM | | | |
| Ficoll | | | 20% | | |

As another work group in the institute works also with a HEPES binding buffer in EMSAs, I tried it and found good protein binding. Therefore, this binding buffer was used by default. It contains all substances of the "consensus" HEPES buffer at comparable concentrations, but misses DTT, KCl and EDTA (see table 40).

Table 40: **Composition of the default binding buffer**

| *Substance* | *Concentration* |
|---|---|
| HEPES | 100 mM (pH=7.5) |
| $MgCl_2$ | 12.5 mM |
| Glycerol | 25% |
| Nonidet P40 | 0.05% |

Having found the best EMSA binding buffer, the amount of unspecific competitor poly-dIdC necessary for complete suppression of unspecific protein-DNA binding was investigated. In [77], amounts of poly-dIdC from 0 – 1.6 µg were tested. From 0.2 µg of poly-dIdC on, no unspecific bands were observed. To be sure that this holds also for my experimentations, I decided to use 2 µg of this substance. In addition, the intensity of the specific shift was only slightly decreased by increasing amounts of poly-dIdC. Generally, the DNA-like polymer poly-dIdC should be preferred as unspecific competitor compared to real double-stranded DNA of unrelated sequence, for poly-dIdC is similar enough to DNA to bind all unspecifically binding proteins, but different enough to provide a binding site for the specific protein of interest [29].

For the amount of antibody for the super-shift assays, I directly followed the recommendations of the antibody suppliers. Very general hints on the EMSAs were obtained from Promega [63], Pierce [61] and Maniatis' Lab Book [74].

## References

[1] Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. *Molecular Biology of the Cell*, 4th ed. Garland Science, 2002.

[2] Allen, R. D. Polymorphism of the human tnf-alpha promoter–random variation or functional diversity? *Molecular Immunology 36* (1999), 1017 – 1027.

[3] American Association on Mental Retardation. Aamr home page. `http://www.aamr.org/`. Visited 2006, April.

[4] Bassuk, A. G., Anandappa, R. T., and Leiden, J. M. Physical interactions between ets and nf-kappab/nfat proteins play an important role in their cooperative activation of the human immunodeficiency virus enhancer in t cells. *Journal of Virology 71* (1997), 3563 – 3573.

[5] Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., and Wheeler, D. L. Genbank. *Nucleic Acids Research 33* (2005), D34 – D38. Database Issue.

[6] Berg, J. M., Tymoczko, J. L., Stryer, L., and Clarke, N. D. *Biochemistry*, 5th ed. W.H. Freeman and Company, 2002.

[7] Bhat, N. K., Thompson, C. B., Lindsten, T., June, C. H., Fujiwara, S., Koizumi, S., Fisher, R. J., and Papas, T. S. Reciprocal expression of human ets1 and ets2 genes during t-cell activation: regulatory role for the protooncogene ets1. *Proceedings of the National Academy of Sciences of the U.S.A. 87* (1990), 3723 – 3727.

[8] Bioline. Hyperladder i. `http://www.bioline.com/h_prod_detail.asp?user_prodname=Hyperladder+I`. Visited 2006, March.

[9] Bioline. Hyperladder iv. `http://www.bioline.com/h_prod_detail.asp?user_prodname=Hyperladder+IV`. Visited 2006, March.

[10] Bonhoeffer, T., and Yuste, R. Spine motility. phenomenology, mechanisms, and function. *Neuron 35* (2002), 1019 – 1027.

[11] Breitwieser, G. E., McLenithan, J. C., Cortese, J. F., Shields, J. M., Oliva, M. M., Majewski, J. L., Machamer, C. E., and Yang, V. W. Colonic epithelium-enriched protein a4 is a proteolipid that exhibits ion channel characteristics. *American Journal of Physiology 272* (1997), C957 – C965.

[12] Brown, T. A. *Genomes*, 2nd ed. BIOS Scientific Publishers, 2002.

[13] Buckland, P. R. The importance and identification of regulatory polymorphisms and their mechanisms of action. *Biochimica et Biophysica Acta 1762* (2006), 17 – 28.

[14] Carl Zeiss CellScience, Ltd. Fluorophore data. `http://www.microscopy.bio-rad.com/fluorescence/fluorophoradata.htm#E-G`. Visited 2006, March.

[15] Chen, W., Erdogan, F., Ropers, H. H., Lenzner, S., and Ullmann, R. Cghpro – a comprehensive data analysis tool for array cgh. *BioMed Central Bioinformatics 6* (2005), 85 – 91.

[16] Church, G. M., and Gilbert, W. Genomic sequencing. *Proceeding of the National Academy of Sciences 81* (1984), 1991 – 1995.

[17] Cooper, G. M. *The Cell – A Molecular Approach*, 2nd ed. Sinauer Associates, 2000.

[18] Corporation, I. Pcr & rt-pcr, pcr optimization. `http://www.invitrogen.com/content.cfm?pageid=10642&fuseaction=appendix.browse&category=PCR_RTPCR`. Visited 2006, March.

[19] Deadman, R., and Jones, K. Dna purification through amplification: Use of phi29 dna polymerase to prepare dna for genomic analyses. *Amersham Biosciences Life Science News 18* (2004), 14 – 15.

[20] Euro-MRX Consortium. The euro-mrx consortium: ground-breaking research into mental retardation. `http://www.euromrx.com/`. Visited 2006, April.

[21] European Collection of Cell Cultures (ECACC). Ask ecacc ... focus on epstein-barr virus lymphocyte transformation. `http://www.ecacc.org.uk/`. Visited 2006, May.

[22] Euskirchen, G., Royce, T. E., Bertone, P., Martone, R., Rinn, J. L., Nelson, F. K., Sayward, F., Luscombe, N. M., Miller, P., Gerstein, M., Weissman, S., and Snyder, M. Creb binds to multiple loci on human chromosome 22. *Molecular and Cellular Biology 24* (2004), 3804 – 3814.

[23] Fermentas Life Sciences. Conventional phage and plasmid dna markers, 8-1118 bp. `http://www.fermentas.com/catalog/electrophoresis/convphagedamarkers.htm`. Visited 2006, March.

[24] Figueiredo, M. S., and Brownlee, G. G. cis-acting elements and transcription factors involved in the promoter activity of the human factor viii gene. *Journal of Biological Chemistry 270* (1995), 11828 – 11838.

[25] Fisher, S. E., Ciccodicola, A., Tanaka, K., Curci, A., Desicato, S., D'urso, M., and Craig, I. W. Sequence-based exon prediction around the synaptophysin locus reveals a gene-rich area containing novel genes in human proximal xp. *Genomics 45* (1997), 340 – 347.

[26] Fujifilm Corporation. Application note no. 2: Dna detection with sybr green i, 1998.

[27] Fujifilm Corporation. Application note no. 3: Dna detection with etbr, 1998.

[28] Fujifilm Corporation. Application note no. 21: Gel shift assay using fluorescent probe, 2003.

[29] GE Healthcare. Poly(di-dc)·poly(di-dc). http://www1.amershambiosciences.com/aptrix/upp01077.nsf/Content/Products?OpenDocument&parentid=41313&moduleid=41316&zone=. Visited 2006, May.

[30] GE Healthcare – formerly Amersham Biosciences. How genomiphi works. http://www4.amershambiosciences.com/APTRIX/upp01077.nsf/Content/phi29_dna_polymerase_whole_genome~genomiphi_work. Visited 2006, May.

[31] GE Healthcare – formerly Amersham Biosciences. The standard in gel, blot, and macroarray analysis. http://www1.amershambiosciences.com/aptrix/upp01077.nsf/Content/gel_blot_storm?OpenDocument&featlinks=storm. Visited 2006, March.

[32] Gilbert, S. F. *Developmental Biology*, 6th ed. Sinauer Associates, 2000.

[33] Glasel, J. A. Validity of nucleic acid purities monitored by 260 nm/280 nm absorbance ratios. *BioTechniques 18* (1995), 62 – 63.

[34] Griffiths, A. J. F., Miller, J. H., Suzuki, D. T., Lewontin, R. C., and Gelbart, W. M. *An Introduction to Genetic Analysis*, 7th ed. W. H. Freeman, 2000.

[35] Haider, N. B., Ikeda, A., Naggert, J. K., and Nishina, P. Genetic modifiers of vision and hearing. *Human Molecular Genetics 11* (2002), 1195 – 1206.

[36] Hashiya, N., Jo, N., Aoki, M., Matsumoto, K., Nakamura, T., Sato, Y., Ogata, N., Ogihara, T., Kaneda, Y., and Morishita, R. In vivo evidence of angiogenesis induced by transcription factor ets-1: Ets-1 is located upstream of angiogenesis cascade. *Circulation 109* (2004), 3035 – 3041.

[37] Health Services Research Information Program. Health services/technology assessment text (hstat). http://www.ncbi.nlm.nih.gov/books/bv.fcgi?rid=hstat6.section.1596. Visited 2006, June.

[38] Institut national de la santé et de la recherche médicale (INSERM). *Mapping Protein/DNA Interactions by Cross-Linking*, 1st ed. INSERM, 2001.

[39] Invitrogen Corporation. 0.24-9.5 kb rna ladder. https://catalog.invitrogen.com/index.cfm?fuseaction=viewCatalog.viewProductDetails&productDescription=15499&CMP=LEC-GCMSSEARCH&HQS=rna%20ladder. Visited 2006, March.

[40] Karolchik, D., Baertsch, R., Diekhans, M., Furey, T. S., Hinrichs, A., Lu, Y. T., Roskin, K. M., Schwartz, M., Sugnet, C. W., Thomas, D. J., Weber, R. J., Haussler, D., and Kent, W. J. The ucsc genome browser database. *Nucleic Acids Research 31* (2003), 51 – 54.

[41] Kavurma, M. M., Bobryshev, Y., and Khachigian, L. M. Ets-1 positively regulates fas ligand transcription via cooperative interactions with sp1. *The Journal of Biological Chemistry 277* (2002), 36244 – 36252.

[42] Kel, A. E., Gossling, E., Reuter, I., Cheremushkin, E., Kel-Margoulis, O. V., and Wingender, E. Match: A tool for searching transcription factor binding sites in dna sequences. *Nucleic Acids Research 31* (2003), 3576 – 3579.

[43] Kenoyer, A. Aimee's non-radioactive emsa protocol. `http://www.protocol-online.org/prot/Detailed/3500.html`. Visited 2006, March.

[44] Kent, W. J. Blat – the blast-like alignment tool. *Genome Research 12* (2002), 656 – 664.

[45] Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M., and Haussler, D. The human genome browser at ucsc. *Genome Research 12* (2002), 996 – 1006.

[46] Korsmeyer, S. J., Wei, M. C., Saito, M., Weiler, S., Oh, K. J., and Schlesinger, P. H. Pro-apoptotic cascade activates bid, which oligomerizes bak or bax into pores that result in the release of cytochrome c. *Cell Death and Differentiation 7* (2000), 1166 – 1173.

[47] Kuida, K., Haydar, T. F., Kuan, C. Y., Gu, Y., Taya, C., Karasuyama, H., Su, M. S., Rakic, P., and Flavell, R. A. Reduced apoptosis and cytochrome c-mediated caspase activation in mice lacking caspase 9. *Cell 94* (1998), 325 – 337.

[48] Lee, S. M., Shin, H., Jang, S. W., Shim, J. J., Song, I. S., Son, K. N., Hwang, J., Shin, Y. H., Kim, H. H., Lee, C. K., Ko, J., Na, D. S., Kwon, B. S., and Kim, J. Plp2/a4 interacts with ccr1 and stimulates migration of ccr1-expressing hos cells. *Biochemical and Biophysical Research Communications 324* (2004), 768 – 772.

[49] Loots, G. G., Ovcharenko, I., Pachter, L., Dubchak, I., and Rubin, E. M. rvista for comparative sequence-based discovery of functional transcription factor binding sites. *Genome Research 12* (2002), 832 – 839.

[50] Maglott, D., Ostell, J., Pruitt, K. D., and Tatusova, T. Entrez gene: gene-centered information at ncbi. *Nucleic Acids Research 33* (2005), D54 – D58. Database Issue.

[51] Max Planck Institute for Molecular Genetics, Home Page on Mental Retardation. Mental retardation. `http://www.molgen.mpg.de/~abt_rop/mr/welcome.html`. Visited 2006, April.

[52] McBride, S. M., Choi, C. H., Wang, Y., Liebelt, D., Braunstein, E., Ferreiro, D., Sehgal, A., Siwicki, K. K., Dockendorff, T. C., Nguyen, H. T., McDonald, T. V., and Jongens, T. A. Pharmacological rescue of synaptic plasticity, courtship behavior, and mushroom body defects in a drosophila model of fragile x syndrome. *Neuron 45* (2005), 753 – 764.

[53] National Institutes of Health. Molecular imaging and contrast agent database (micad). `http://www.ncbi.nlm.nih.gov/books/bv.fcgi?rid=micad`. Visited 2006, June.

[54] Ng, F. W., Nguyen, M., Kwan, T., Branton, P. E., Nicholson, D. W., Cromlish, J. A., and Shore, G. C. p28 bap31, a bcl-2/bcl-xl- and procaspase-8-associated protein in the endoplasmic reticulum. *The Journal of Cell Biology 139* (1997), 327 – 338.

[55] Northwestern University. Oligonucleotide properties calculator. `http://www.basic.northwestern.edu/biotools/oligocalc.html`. Visited 2006, March and April.

[56] Oliva, M. M., Cortese, J. F., and Yang, V. W. Promoter regulation of a differentially expressed gene in the human colonic epithelial cell lines ht29-18 and ht29-18-c1. *Gene 159* (1995), 151 – 157.

[57] Oliva, M. M., Wu, T. C., and Yang, V. W. Isolation and characterization of a differentiation-dependent gene in the human colonic cell line ht29-18. *Archives of Biochemistry and Biophysics 302* (1993), 183 – 192.

[58] Online, P. Protocol online – your lab's reference book. `http://www.protocol-online.org/`. Visited 2006, March.

[59] Parker, D., Ferreri, K., Nakajima, T., LaMorte, V. J., Evans, R., Koerber, S. C., Hoeger, C., and Montminy, M. R. Phosphorylation of creb at ser-133 induces complex formation with creb-binding protein via a direct mechanism. *Molecular and Cellular Biology 16* (1996), 694 – 703.

[60] Pierce Biotechnology. Lightshift chemiluminescent emsa kit, 2005.

[61] Pierce Biotechnology, Inc. Introduction to the emsa (gel shift) technique. `http://www.piercenet.com/Products/Browse.cfm?fldID=99A16E15-3470-4FD1-9A3C-0BBABBBE45F9`. Visited 2006, March.

[62] Pontius, J. U., Wagner, L., and Schuler, G. D. *The NCBI Handbook*. National Center for Biotechnology Information (NCBI), 2003, ch. 21: UniGene: A Unified View of the Transcriptome, pp. 1 – 12.

[63] Promega Corporation. Gel shift assay system. `http://www.promega.com/tbs/tb110/tb110.pdf`. Visited 2006, March.

[64] Pruitt, K. D., Tatusova, T., and Maglott, D. R. Ncbi reference sequence (refseq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Research 33* (2005), D501 – D504. Database Issue.

[65] Qiagen. *Taq* pcr handbook. `http://www1.qiagen.com/literature/handbooks/PDF/PCRAndReverseTranscription/KitsAndEnzymes/PCR_Taq/1035063_Qiag_HB.pdf`. Visited 2006, March.

[66] Quandt, K., Frech, K., Karas, H., Wingender, E., and Werner, T. Matind and matinspector – new fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acids Research 23* (1995), 4878 – 4884.

[67] Rayet, B., and Gelinas, C. Aberrant rel/nfkb genes and activity in human cancer. *Oncogene 18* (1999), 6938 – 6947.

[68] Reed, J. C., and Huang, Z. *Apoptosis pathways and drug targets*. Nature Reviews Drug Discovery / Molecular Cell Biology, 2004.

[69] Ropers, H. H. X-linked mental retardation: many genes for a complex disorder. *Current Opinion in Genetics and Development 16* (2006), 260 – 269.

[70] Ropers, H. H., and Hamel, B. C. J. X-linked mental retardation. *Nature Reviews Genetics 6* (2005), 46 – 57.

[71] Ropers, H. H., Hoeltzenbein, M., Kalscheuer, V., Yntema, H., Hamel, B., Fryns, J. P., Chelly, J., Partington, M., Gecz, J., and C, C. M. Nonsyndromic x-linked mental retardation: where are the missing mutations? *TRENDS in Genetics 19* (2003), 316 – 320.

[72] Ruscher, K., Reuter, M., Kupper, D., Trendelenburg, G., Dirnagl, U., and Meisel, A. A fluorescence based non-radioactive electrophoretic mobility shift assay. *Journal of Biotechnology 78* (2000), 163 – 170.

[73] Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A., and Arnheim, N. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science 230* (1985), 1350 – 1354.

[74] Sambrook, J., Fritsch, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual*, 2nd ed. Cold Spring Harbor Laboratory Press, 1989.

[75] Sandelin, A., Wasserman, W. W., and Lenhard, B. Consite: web-based prediction of regulatory elements using cross-species comparison. *Nucleic Acids Research 32* (2004), W249 – W252. Web Server Issue.

[76] Sanger, F., Nicklen, S., and Coulson, A. R. Dna sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the U.S.A 74* (1977), 5463 – 5467.

[77] Schöler, H. R. Lab hint – use of poly [d(i-c)] in the electrophoretic mobility shift assay (emsa). Boehringer Mannheim. Molecular Biology, 10 – 11.

[78] Schug, J., and Overton, G. C. Tess: Transcription element search software on the www. `http://www.cbil.upenn.edu/cgi-bin/tess/tess`. Visited 2006, March.

[79] Strachan, T., and Read, A. P. *Human Molecular Genetics*, 2nd ed. BIOS Scientific Publishers, 1999.

[80] Strausberg, R. L., Feingold, E. A., Grouse, L. H., Derge, J. G., Klausner, R. D., Collins, F. S., Wagner, L., Shenmen, C. M., Schuler, G. D., Altschul, S. F., Zeeberg, B., Buetow, K. H., Schaefer, C. F., Bhat, N. K., Hopkins, R. F., Jordan, H., Moore, T., Max, S. I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A. A., Rubin, G. M., Hong, L., Stapleton, M., Soares, M. B., Bonaldo, M. F., Casavant, T. L., Scheetz, T. E., Brownstein, M. J., Usdin, T. B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S. S., Loquellano, N. A., Peters, G. J., Abramson, R. D., Mullahy, S. J., Bosak, S. A., McEwan, P. J., McKernan, K. J., Malek, J. A., Gunaratne, P. H., Richards, S., Worley, K. C., Hale, S., Garcia, A. M., Gay, L. J., Hulyk, S. W., Villalon, D. K., Muzny, D. M., Sodergren, E. J., Lu, X., Gibbs, R. A., Fahey, J., Helton, E., Ketteman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A. C., Shevchenko, Y., Bouffard, G. G., Blakesley, R. W., Touchman, J. W., Green, E. D., Dickson, M. C., Rodriguez, A. C., Grimwood, J., Schmutz, J., Myers, R. M., Butterfield, Y. S., Krzywinski, M. I., Skalska, U., Schnerch, D. E. S. A., Schein, J. E., Jones, S. J., Marra, M. A., and Mammalian Gene Collection Program Team. Generation and initial analysis of more than 15,000 full-length human and mouse cdna sequences. *Proceedings of the National Academy of Sciences of the U.S.A. 99* (2002), 16899 – 16903.

[81] Sugden, B., Yates, J., and Mark, W. Transforming functions associated with epstein-barr virus. *The Journal of Investigative Dermatology 83* (1984), 82s – 87s. Supplement 1.

[82] Takai, N., Miyazaki, T., Nishida, M., Nasu, K., and Miyakawa, I. c-ets1 is a promising marker in epithelial ovarian cancer. *International Journal of Molecular Medicine 9* (2002), 287 – 292.

[83] The Cepko Lab, Harvard Medical School. Emsa using ds oligonucleotides. `http://axon.med.harvard.edu/~cepko/protocol/mike/E1.html`. Visited 2006, March.

[84] The Hahn Laboratory, Fred Hutchinson Cancer Research Center. Gel mobility shift assay conditions -mg/edta in gel and buffer. `http://www.fhcrc.org/science/labs/hahn/methods/biochem_meth/gel_shift.html`. Visited 2006, May.

[85] Toda, T., and Sugimoto, M. Proteome analysis of epstein-barr virus-transformed b-lymphoblasts and the proteome database. *Journal of Chromatography. B, Analytical Technologies in the Biomedical and Life Sciences 787* (2003), 197 – 206.

[86] Trevisan, R., Daprai, L., Paloschi, L., Vajente, N., Chieco-Bianchi, L., and Saggioro, D. Antiapoptotic effect of human t-cell leukemia virus type 1 tax protein correlates with its creb transcriptional activity. *Experimental Cell Research 312* (2006), 1390 – 1400.

[87] Tsunoda, T., and Takagi, T. Estimating transcription factor bindability on dna. *Bioinformatics (Oxford, England) 15* (1999), 622 – 630.

[88] van Dyke, M. W. *Do DNA Triple Helices or Quadruplexes Have a Role in Transcription?* Eurekah Bioscience Collection. Landes Bioscience, 2004.

[89] Wang, B., Nguyen, M., Breckenridge, D. G., Stojanovic, M., Clemons, P. A., Kuppig, S., and Shore, G. C. Uncleaved bap31 in association with a4 protein at the endoplasmic reticulum is an inhibitor of fas-initiated release of cytochrome c from mitochondria. *Journal of Biological Chemistry 278* (2003), 14461 – 14468.

[90] Wheeler, D. L., Church, D. M., Federhen, S., Lash, A. E., Madden, T. L., Pontius, J. U., Schuler, G. D., Schriml, L. M., Sequeira, E., Tatusova, T. A., and Wagner, L. Database resources of the national center for biotechnology. *Nucleic Acids Research 31* (2003), 28 – 33.

[91] WIKIPEDIA, The Free Encyclopedia. Null hypothesis. `http://en.wikipedia.org/wiki/Null_hypothesis`. Visited 2006, May.

[92] WIKIPEDIA, The Free Encyclopedia. P-value. `http://en.wikipedia.org/wiki/P-value`. Visited 2006, May.

[93] WIKIPEDIA, The Free Encyclopedia. Student's t-test. `http://en.wikipedia.org/wiki/Student%27s_t-test`. Visited 2006, May.

[94] Wingender, E., Chen, X., Fricke, E., Geffers, R., Hehl, R., Liebich, I., Krull, M., Matys, V., Michael, H., Ohnhaeuser, R., Pruess, M., Schacherer, F., Thiele, S., and Urbach, S. The transfac system on gene expression regulation. *Nucleic Acids Research 29* (2001), 281 – 283.

[95] World Health Organization. Mental health. `http://www.who.int/topics/mental_health/en/`. Visited 2006, April.

[96] Xiao, W., and Oefner, P. J. Denaturing high-performance liquid chromatography: A review. *Human Mutation 17* (2001), 439 – 474.

[97] Yuste, R., and Bonhoeffer, T. Genesis of dendritic spines: insights from ultrastructural and imaging studies. *Nature Reviews Neuroscience 5* (2004), 24 – 34.

[98] Zerby, D., and Lieberman, P. M. Functional analysis of tfiid-activator interaction by magnesium-agarose gel electrophoresis. *Methods (San Diego, California) 12* (1997), 217 – 223.

## Acknowledgements

I wish to thank the following people:

- *Doctors Andreas W. Kuß and Lars Riff Jensen:* for help and advice concerning experimental procedures and preparation of in-house presentations, for extensive discussion of the experimental results, and for carefully reading the thesis manuscript
- *Technicians Marion Amende-Acar, Bettina Lipkowitz and Marianne Schlicht:* for lots of experimental hints and jokes that both made laboratory life much easier
- *PhD students Mahdi Mohammad Motazacker, Masoud Garshasbi, Wei Chen, Lia Abbasi-Moheb and Joanna Walczak as well as student assistants Katharina Albers, Annika Quast and Franziska Welzel:* for being very friendly and humorous colleagues
- *Prof. Dr. H.-Hilger Ropers:* for giving me the opportunity to do my diploma thesis in his department
- *Prof. Dr. Roland Lauster:* for establishing a well organized course of studies that is up to date regarding the scientific background in the lectures as well as the laboratory methods in the practical courses – and for being a helpful contact person for his students in the truest sense of the word
- *My parents:* for supporting my long-term education culminating in this work
- *My girlfriend Katharina:* for enlightening my life

## Eidesstattliche Erklärung

Die selbständige und eigenständige Anfertigung versichert an Eides statt

Ort, Datum Unterschrift