

# ***Cis and trans-acting elements in somatic hypermutation***

Dissertation

der Fakultät für Biologie  
der Ludwig-Maximilian-Universität München

vorgelegt von

**Maik Klasen**

München, den 06.10.2004

Diese Dissertation wurde im Labor von Prof. Dr. Matthias Wabl an der University of California in San Francisco, Department of Microbiology and Immunology, durchgeführt.

Folgende Veröffentlichungen sind aus dieser Arbeit hervorgegangen:

Klasen, M. The  $E\mu$  intronic enhancer in somatic hypermutation. 2002. FUTURA **17**: 92-94

Klasen, M. and Wabl, M. Silent point mutation in DsRed resulting in enhanced relative fluorescence intensity. 2004. BioTechniques **36** (2): 236-238

Klasen, M., Spillmann, F., and Wabl, M. Somatic hypermutation and mismatch repair in non-B cells. 2004. Submitted

Klasen, M., Spillmann, F., Lorens, J.B., and Wabl, M. Retroviral vectors to monitor somatic hypermutation. 2004. Submitted

Tag der mündlichen Prüfung: 09.12.2004

Dekan: Professor Ute Harms

1. Gutachter: Professor Elisabeth Weiss

2. Gutachter: Professor Fredericke Eckardt-Schupp

Sondergutachter: Professor Matthias Wabl

# Table of Contents

## Abbreviations

<b>Chapter 1: Introduction</b>	<b>1</b>
1.1 <i>Cis</i> -acting elements in somatic hypermutation	3
1.2 <i>Trans</i> -acting elements in somatic hypermutation	4
1.2.1 Activation-induced cytidine deaminase (AID)	4
1.2.2 Mismatch repair in somatic hypermutation	9
1.3 Mutation rates measurements	13
1.4. Specific Aims	15
1.4.1. Specific aim 1: Develop a retroviral based point mutation indicator system	15
1.4.2. Specific aim 2: Characterize the IgH major intronic enhancer and other enhancer elements, as <i>cis</i> -elements in hypermutation	16
1.4.3. Specific aim 3: Mismatch repair in somatic hypermutation in non-B cells	16
<b>Chapter 2: Materials and Methods</b>	<b>17</b>
2.1. Materials	17
2.1.1. Instruments	17
2.1.2. Enzymes and reaction kits	17
2.1.3. Antibodies and enzyme conjugated anti-antibodies	18
2.1.4. Bacterial strains	19
2.1.5. Plasmid and retroviral vector construction	19
2.1.6. Oligonucleotides	24
2.2. Cell culture based methods	26
2.2.1. Mammalian cell lines and cell culture conditions	26
2.2.2. Cryostorage of cell lines	27

2.2.3. Rapid production of retroviruses in Phoenix packaging cells	28
2.2.4. Spin infection of adherent cells	28
2.2.5. Spin infection of suspension cells	29
2.2.6. Fluorescent adsorbent cell sorting (FACS)	30
2.2.7. Analytical flow cytometry	30
2.3. Molecularbiological Methods	31
2.3.1. Strain maintenance and strain conservation	31
2.3.2. Making chemically competent <i>E.coli</i> cells (RbCl method by Hanahan)	31
2.3.3. Heath shock transformation of competent <i>E.coli</i> cells	31
2.3.4. Purification of plasmid DNA	32
2.3.5. Isolation of chromosomal DNA	32
2.3.6. Isolation of polyA <sup>+</sup> mRNA	32
2.3.7. Preparative and analytical gel electrophoresis	32
2.3.8. Concentration measurements and purity estimations of plasmid DNA	33
2.3.9. Endonuclear restriction digest of DNA	33
2.3.10. DNA-Ligation	33
2.3.11. PCR amplification of DNA fragments	34
2.3.12. First strand cDNA synthesis	35
2.4. Biochemical Methods	35
2.4.1. Protein concentration measurements after Warburg and Christian	35
2.4.2. Protein measurements after Bradford	35
2.4.3. SDS polyacrylamide gelelectrophoresis after Laemmli	35
2.4.4. Silver staining of SDS polyacrylamide gels modified after Blum	36
2.4.5. Coomassie staining of SDS polyacrylamide gels by Wilson	37
2.4.6. Western blot modified after Towbin	37
<b>Chapter 3: Results</b>	<b>39</b>
<b>3.1. Specific Aim 1: Develop a retroviral based point mutation indicator system</b>	
3.1.1. Expression of activation induced cytidine deaminase (AID) in 18-81 cells	39
3.1.2. High virus titer production	40
3.1.3. Retroviral backbone	42
3.1.5. GFP family members	44

3.1.6. A silent point mutation in DsRed resulting in enhanced relative fluorescence intensity	47
3.1.7. Alternative red fluorescence proteins	51

### **3.2. Specific aim 2: Characterize the IgH major intronic enhancer and other enhancer elements, as *cis*-elements in hypermutation**

3.2.1. CRU5-Red537-IRES-GFP(X) constructs	52
3.2.2. CRU5-DsRed537(X)-IRES-GFP vectors	56
3.2.3. <i>Cis</i> -acting elements	59

### **3.3. Specific aim 3: Somatic hypermutation and mismatch repair in non-B cells**

3.3.1. AID expression in NIH/3T3 cells	63
3.3.2. AID mediated hypermutation in NIH/3T3 cells	64
3.3.3. Hypermutation in the presence and absence of mismatch repair	66

## **Chapter 4: Discussion and conclusion**

### **4.1. Specific Aim 1: Develop a retroviral based point mutation indicator system**

4.1.1. High virus titer production	70
4.1.2. CRU5 based retroviral reporter constructs	71
4.1.3. Fluorescence marker genes versus antibiotic selection markers	73
4.1.4. DsRed	73
4.1.5. A silent point mutation in DsRed resulting in enhanced relative fluorescence intensity	74

### **4.2. Specific Aim 2: Characterize the IgH major intronic enhancer and other enhancer elements, as *cis*-elements in hypermutation**

4.2.1. <i>Cis</i> -acting elements in somatic hypermutation	75
---	----

### **4.3. Specific Aim 3: Somatic hypermutation and mismatch repair in non-B cells**

4.3.1. Hypermutation rate in the 293T cell line	77
4.3.2. Conclusion	79

<b>Chapter 5: Summary</b>	81
<b>Chapter 6: References</b>	83
<b>Chapter 7: Acknowledgements</b>	96

## Abbreviations

AID	activation induced cytidine deaminase
AP-site	apurinic or apyrimidinic site in a DNA strand
BER	base excision repair
C	immunoglobulin constant region
CSR	class switch recombination
D	diversity gene segment
$E\mu$	large heavy chain Ig intronic enhancer
H	heavy chain
HIGM2	hyper-IgM syndrome 2
HNPCC	hereditary nonpolyposis colorectal cancer
Ig	immunoglobulin
J	joining gene segment
$\kappa 3'$	kappa chain 3' enhancer
$\kappa i$	kappa chain intronic enhancer
L	light chain
MAR	matrix attachment region
MMR	mismatch repair
mutator	mutator enzyme complex
RAG	recombination-activating gene
RPA	replication protein A
SHM	somatic hypermutation
UNG	uracil DNA glycosylase
V	variable gene segment

## 1. Introduction

This thesis focuses on the genetic elements and molecular mechanisms involved in somatic hypermutation of the immunoglobulin (Ig) locus. On the basis of theoretical considerations, it was first suggested by Lederberg in 1959, that somatic hypermutation is one of the generators of antibody diversity [Lederberg, 1959]. The first concrete experimental evidence for this was provided much later by Weigert *et al.* in 1970, before the primary and secondary repertoires were distinguished in molecular terms [Weigert *et al.*, 1970]. When a B cell is activated by antigen contact, it can home in the germinal centers of the follicles of lymphoid organs (i.e. spleen, tonsils, Peyer's patches). There the gene segments encoding the variable region (V-region) of the B cell receptor may be mutated at a high rate [Wabl *et al.*, 1999]. Some of these mutations may change the affinity of the receptor, and B cells with a higher affinity to the antigen are selected and proliferate. Thus, hypermutation contributes to the so-called "affinity maturation" of activated B-cells and creates the secondary repertoire.

The primary antibody repertoire is generated during early B-lymphocyte development by the rearrangement of the V, D, and J gene segments at the Ig heavy (H) chain locus, and the V, J rearrangement of the light (L) chain gene segments [Tonegawa, 1983]. This process is mediated by the products of the lymphoid specific recombination-activating genes 1 (RAG-1) and 2 (RAG-2) [Hesslein and Schatz, 2001]. A naïve (pre-immune) B-cell repertoire is thus generated by various diversifiers: (i) combination of different H and L chains; (ii) combination of different gene segments; (iii) and junctional diversity at the ligation sites of the gene segments.



However, the primary B-cell repertoire generates B cell receptors and antibodies with rather low affinity to antigen structures. High affinity antibodies are generated mainly in the secondary repertoire through hypermutation, after a B cell has been activated, but before they differentiate into plasma or memory cells. In this process, point mutations are introduced at high rates at the Ig locus. The mutation rate is a million fold higher than the estimated spontaneous mutation rate of  $10^{-9}$  to  $10^{-10}$  per bp per cell generation, at other loci of the genome [McKean *et al.*, 1884; Allen *et al.*, 1987; Berek *et al.*, 1991]. The mutations that are created by hypermutation are largely restricted to nucleotide substitutions rather than deletions and insertions. A comparison of hypermutated DNA sequences from an Ig database revealed sites of even higher intrinsic mutability than others; these are called hot spots, with the consensus sequence motif RGYW (R = A or G; Y = C or T; W = A or T) [Rogozin and Kolchanov, 1992; Betz *et al.*, 1993; Jolly *et al.*, 1996; Rogozin *et al.*, 1996]. It was reasoned that hypermutation at the Ig genes is a paradigm of a site-specific, stage-specific, and lineage-specific "mutator system" that generates point mutations [Wabl *et al.*, 1985]. Indeed, at the Ig locus the mutator action is restricted to a 2 kb stretch encompassing the rearranged VDJ or VJ sequences of the endogenous Ig genes [Neuberger and Milstein, 1995]. The 5' boundary of this stretch is well defined by the transcriptional start site of the Ig promoter; the 3' boundary near the Ig intronic enhancer is more loosely defined [Gearhart *et al.*, 1981; Gearhart and Bogenhagen, 1983]. Surprisingly, the V-region is not absolutely necessary for triggering hypermutation as shown by Milstein, Neuberger and colleagues in transgenic constructs where they substituted the Ig V-regions by reporter genes, like *gpt* or *neo<sup>r</sup>* [Yelamos *et al.*, 1995]. Still, the RGYW motif is a mutational hot spot in these reporter genes. Later,

the hot spot motif was confirmed by actual experimentation [Bachl *et al.*, 1997]. From the beginning, and still to date, the questions the field has been trying to answer are: What are the *cis*-acting and *trans*-acting elements, and what kind of molecular mechanisms are active in somatic hypermutation?

### 1.1. *Cis*-acting elements in somatic hypermutation

There are many reviews that cover the *cis*-acting genetic elements involved in hypermutation [Klein *et al.*, 1998; Kong *et al.*, 1998; Neuberger *et al.*, 1998; Storb *et al.*, 1998; Winter and Gearhart, 1998]. It is well known that transcription is necessary for driving hypermutation. In an inducible system, Bachl *et al.* [Bachl *et al.*, 2001] showed that hypermutation rates correlate with transcription rates, but that transcription does not have to be driven by the Ig promoter [Bachl *et al.*, 1998; Bachl *et al.*, 2001]. Apart from the promoter, for full hypermutation either of the large Ig intronic enhancers ( $E\mu$  or  $E_{\text{kappa}}$  enhancer) and an additional 3' enhancer were suggested to be sufficient to target the mutator [Betz *et al.*, 1994; Bachl *et al.*, 1998]. Again, the 3' enhancer has not necessarily to be the 3' Ig enhancer; but without it hypermutation occurs at a lower rate [Bachl and Olsson, 1999]. At the beginning of this study, it seemed that the  $E\mu$  enhancer, in the right distance and orientation to the gene to be mutated, is the only Ig-locus specific sequence required to recruit the *trans*-acting factors for hypermutation [Bachl *et al.*, 1998]. This was shown by Bachl and Olsson [Bachl and Olsson, 1999] in experiments, where the large intronic enhancer was placed next to a green fluorescence protein (GFP) to confer hypermutability, without any other Ig sequence present nearby.

The experiments described above are consistent with the notion that hypermutation of Ig genes is the prime example for site directed mutagenesis. However, because several non-Ig genes (i.e. *BCL-6*, *B29* and *mb1*, *CD95*) have been reported to hypermutate in activated B-cells [Shen *et al.*, 1998; Gordon *et al.*, 2003; Muschen *et al.*, 2002], the site-specificity of the mutator action has been challenged. On the other hand, selected house keeping genes in activated B cells show no indication of increased mutability [Shen *et al.*, 2000]. To fit these results into the general view of site-directed hypermutation in Ig genes, it was hypothesized that the non-Ig loci would contain unidentified Ig-like *cis*-acting regulatory sequences, which would allow the targeting of the mutator complex. However, non-Ig enhancers have not been tested extensively enough to exclude their activity of targeting the mutator, nor have the intronic Ig enhancer, or any other enhancer, been trimmed to the minimal sequence sufficient for hypermutation.

## **1.2. *Trans*-acting factors in somatic hypermutation**

### **1.2.1. Activation-induced cytidine deaminase (AID)**

At the beginning of this study, not much was known about the *trans*-acting factors and the molecular mechanisms involved in hypermutation, which gave rise for several intriguing hypotheses [e.g. Manser, 1990; Rogerson *et al.*, 1991; Gearhart, 1983]. It was the discovery of the activation-induced cytidine deaminase (AID) and its absolute requirement for somatic hypermutation, which triggered the recent progress in the field [Muramatsu *et al.*, 1999; Muramatsu *et al.*, 2000]. AID is exclusively expressed in activated B cells of the germinal center; and mice deficient in AID are completely

defective in somatic hypermutation [Muramatsu *et al.*, 2000]. Patients diagnosed with an autosomal recessive form of the hyper-IgM syndrome 2 (HIGM2) revealed that AID inactivating mutations cause the lack of somatic hypermutation [Revy *et al.*, 2000; Quartier *et al.*, 2004]. Further, ectopic expression of AID in hybridoma [Martin *et al.*, 2002], fibroblasts [Yoshikawa *et al.*, 2002], and *E.coli* [Petersen-Mahrt *et al.*, 2002] induces hypermutation, suggesting that AID is the only B cell specific *trans*-acting factor required for the process. Biochemically, AID deaminates deoxycytidine to deoxyuridine exclusively in single-stranded (ss), but not in double-stranded DNA; nor in RNA, or DNA-RNA hybrids [Bransteitter *et al.*, 2003; Pham *et al.*, 2003, Dickerson *et al.*, 2003]. In a cell free *in vitro* assay, semi-purified AID targeted dC's in the short WRC motif (W = A or T, R = A or G) preferentially on the non-template DNA strand [Pham *et al.*, 2003; Chaudhuri *et al.*, 2003; Ramiro *et al.*, 2003; Sohail *et al.*, 2003]. However, *in vivo* data show no preference between the transcribed and the non-transcribed strand in mice [Boursier *et al.*, 2004; Milstein and Neuberger, 1998], and even earlier results show that purine residues on the coding strand seem to be preferred over pyrimidines as targets for mutations [Neuberger and Milstein, 1995].

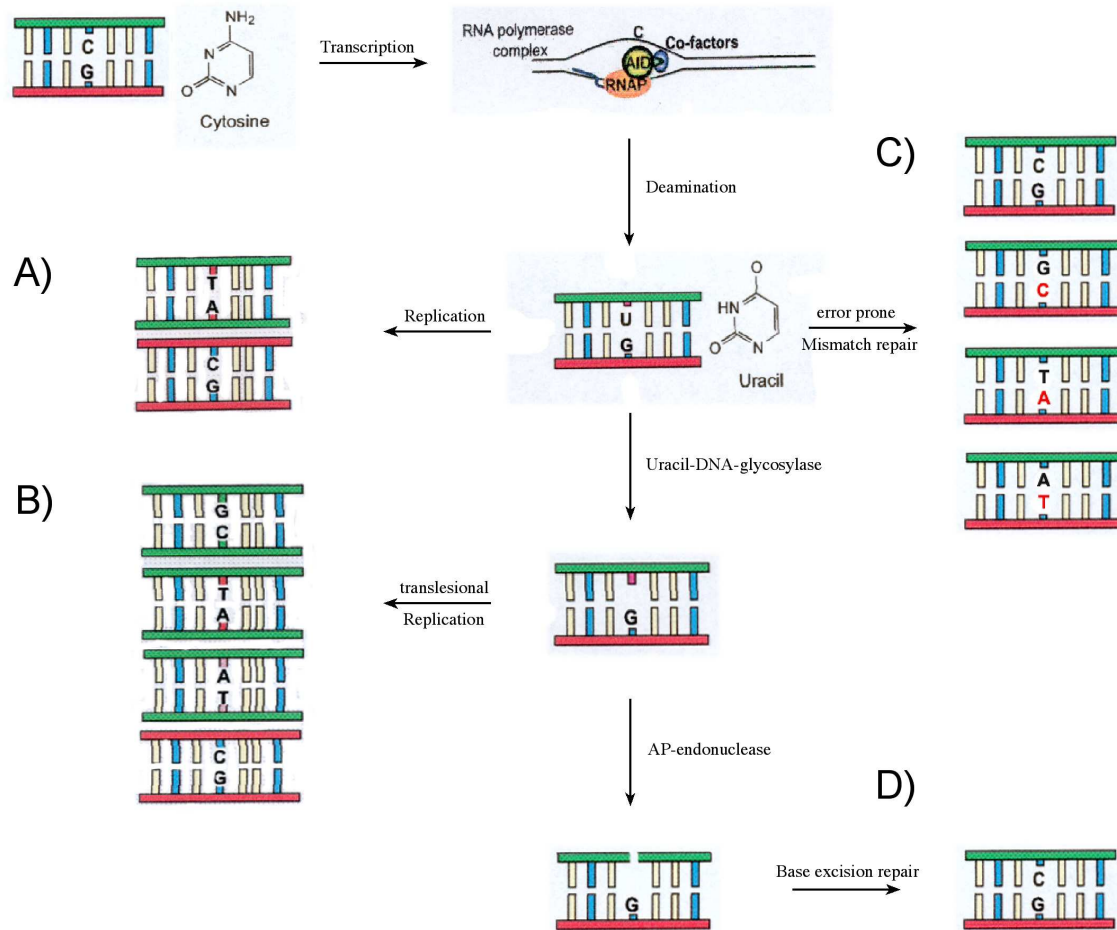
Amino acid (aa) sequence analysis of AID revealed an N-terminal bipartite nuclear targeting sequence (aa position 9 - 26), but also a nuclear export signal at the C-terminus (aa position 183 - 198) of the protein [<http://au.expasy.org/prosite/>]. Mutations in the N-terminus of AID inactivate somatic hypermutation [Shinkura *et al.*, 2004], whereas mutations in the C-terminus of AID selectively inactivate Ig-class switch recombination but retain somatic hypermutation [Barreto *et al.*, 2003; Ta *et al.*, 2003, McBride *et al.*, 2004].

Class switch recombination is the second mechanism mediated by AID expression in activated B-cells. This recombination leads to switching in protein expression from IgM to the other Ig isoforms—IgG, IgE or IgA. AID inactivating mutations were pinpointed as cause for the name-giving characteristics of the hyper-IgM syndrome 2 in patients: the lack of other immunoglobulin isoforms [Quartier *et al.*, 2004]. Because induction of class switch recombination and somatic hypermutation can be assigned to separate domains of AID, it was hypothesized that specific co-factors might interact with different domains of AID in order to distinguish the two molecular mechanisms [Shinkura *et al.*, 2004]. Such co-factors could also account for specific targeting of AID to the V-region during hypermutation or to the switch regions during switch recombination, respectively. Secondary modification of AID might be required for specific interaction with co-factors. The amino acid sequence of AID shows multiple phosphorylation sites at its C-terminus [<http://au.expasy.org/prosite/>], and a phosphorylated form of AID has been reported [Chaudhuri *et al.*, 2004]. In this paper, it was shown that the non-phosphorylated AID has no activity in RGYW targeted hypermutation.

Most of the ectopically and endogenously expressed AID is localized in the cytoplasm [Ito *et al.*, 2004; Rada *et al.*, 2002a]. This fact, together with the high DNA sequence homology of AID to the apolipoprotein mRNA-editing enzyme catalytic polypeptide 1 (APOBEC-1) led to the notion that AID might edit mRNA rather than ssDNA [Kinoshita and Honjo, 2001; Honjo, 2002]. Additionally, the requirement for *de novo* protein synthesis in the presence of active AID in order to induce somatic hypermutation and class switch recombination suggested the view that AID does target mRNA [Doi *et al.*, 2003]. The targeted mRNA would encode the inactive form of the mutator enzyme or

switch recombinase respectively, which would be "activated" by an AID mediated editing process. But recent accumulation of biochemical data and the finding that AID is actively exported from the nucleus via the exportin protein CRM-1 pathway [McBride *et al.*, 2004] led the authors of the mRNA-editing model to abandon it [Honjo *et al.*, 2004].

With AID activating somatic hypermutation by deaminating dC to dU, several downstream processes have been proposed (Fig. 1). In case the G-U mismatch would remain through DNA replication, a C  $\rightarrow$  T mutation (or G  $\rightarrow$  A mutation on the opposite DNA strand) would be produced. This is a mutation pattern which is frequently found during ectopic AID expression in cell lines and *E.coli* [Yoshikawa *et al.*, 2002; Sohail *et al.*, 2003]. Downstream of the AID action, the uracil DNA glycosylase (UNG) plays a major role in somatic hypermutation by removing the uracil bases. Inhibition of UNG activity in DT40 cells, and deletion of UNG in AID expressing *E.coli*, results in an increase of GC base pair targeted transitions [Di Noia and Neuberger, 2002; Petersen-Mahrt *et al.*, 2002]. Furthermore, UNG deficiency in mice lead to a perturbed hypermutation spectrum [Rada *et al.*, 2002b]. However, UNG action leaves an abasic site (AP site = apurinic or apyrimidinic site), which could have various outcomes depending on the race between replication and the base excision repair (BER) pathway. In the BER pathway, the AP-site would be targeted by the AP-endonuclease, which creates a single strand lesion. The lesion would be repaired by the BER pathway, but without creating mutations (Fig. 1D). On the other hand, the AP-site could be filled with any base during translesional replication (Fig. 1B). Another possibility includes the mismatch repair pathway (MMR), and its involvement in somatic hypermutation has been demonstrated in several knock out models (Fig. 1C).



**Figure 1.** Model of somatic hypermutation. During RNA polymerase-mediated transcription (RNAP, orange oval), AID (yellow oval) and co-factors (blue oval) deaminate cytosines (C) on the single-stranded DNA of the transcription bubble to generate uracil (U). The result of the reaction is a G-U mismatch, which can be resolved as follows: **A)** G-U mismatches remain unaltered, enter DNA replication and result in C  $\rightarrow$  T mutation. **B)** UNG removes uracil to generate an abasic site, which can be filled with any base during replication. **C)** Mismatch repair (MMR) proteins bind to the G-U base mispairs or G-abasic mismatches and recruits additional MMR proteins, error prone polymerases and B cell specific subversion factors to generate and fix mutations. **D)** UNG removes uracil and the abasic site is processed by base excision repair in an error free manner. (Model and figure modified after Li *et al.*, 2004 and Neuberger *et al.*, 2003)

### 1.2.2. Mismatch repair in somatic hypermutation

Misincorporated bases introduced during replication of DNA are repaired by the postreplicative mismatch repair (MMR) system [Kolodner, 1996]. The eukaryotic MMR system is mediated by heterodimeric complexes of its bacterial MutL (hMLHs) and MutS (hMSHs) homologues [Jiricny, 1998; Modrich, 1991]. Depending on the characteristics of the mismatch, two different complexes initiate the repair process. The hMSH2-hMSH6 (MutS $\beta$ ) complex recognizes single base mismatches, insertions and deletions [Marsischky *et al.*, 1998], whereas the hMSH2-hMSH3 (MutS $\alpha$ ) complex detects two to four base pair insertions and deletions [Sia *et al.*, 1997; Alani, 1996; Umar, *et al.*, 1998; Wei *et al.*, 2002]. Since the recognition of the different classes of mismatches is not stringent, both MMR complexes share overlapping base mismatch patterns [Genschel *et al.*, 1998]. The subsequent removal of the base mismatch is promoted by the MutL $\alpha$  complex, formed by hMLH1 and hPMS2, which interacts with hMSH2 of the MutS $\alpha$  complex [Plotz *et al.*, 2003]. Although there is reasonable knowledge of the initiation of the MMR system, less is known about the subsequent excision of the mismatches and about the proteins involved in the downstream processes. It was suggested that the 3' - 5' exonuclease 1 (EXO1) is required for MMR in that it presumably stabilizes the MMR complexes by interacting with hMSH2 and hMLH1 [Genschel *et al.*, 2002; Schmutte *et al.*, 2001; Amin *et al.*, 2001]. Additionally, FEN1 (RAD27), DNA polymerases  $\beta$  and  $\delta$ , and the DNA replication factors RPA (replication protein A), PCNA (proliferating cell nuclear antigen) and RFC (replication factor C) might be involved [Kolodner, 1996; Tischkoff *et al.*, 1997; Amin *et al.*, 2001; Buermeyer *et al.*, 1999]. However, the general mismatch repair scheme proposes, that after the mismatch has been detected by one of



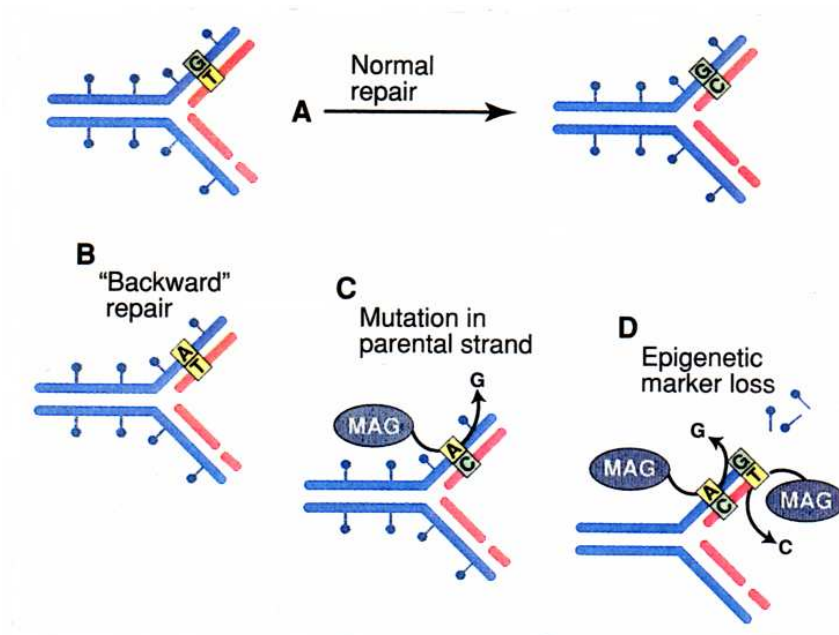
the various MMR complexes, single stranded nicks would be introduced into the DNA, which subsequently would be nibbled back and resynthesized by the DNA polymerases mentioned above.

Loss of MMR leads to a mutator phenotype in both prokaryotic and eukaryotic cells. In humans, loss of MMR ultimately results in microsatellite instability, which is frequently found in tumors of patients diagnosed with hereditary nonpolyposis colorectal cancer (HNPCC) [Bronner *et al.*, 1994]. Cancer predisposition is linked to germline mutations in MMR genes and most mutations associated with HNPCC were found in hMLH1, hMSH2, and hMSH6 genes [Peltomaki and Vasen, 1997]. Mutations in hPMS2 were reported sporadically. The majority of all HNPCC causing mutations were found in hMLH1 (61%), making the analysis of pathogenic hMLH1 variants a proprietary goal [Trojan *et al.*, 2002]. Interestingly, loss of EXO1, FEN1 and PCNA function results in a mutator phenotype, but does not seem to significantly contribute to HNPCC formation [Kolodner, 1996].

In contrast to the increased levels of spontaneous mutations in MMR deficient cell lines and mice, somatic hypermutation frequencies decrease in activated B cells of MLH1 or PMS2 deficient mice [Cascalho *et al.*, 1998; Phung *et al.*, 1999]. Mutation frequencies do not seem to be decreased in MSH2 deficient mice [Phung *et al.*, 1998]. It has been suggested that AID would introduce the G-U mismatch, which would be detected by the mismatch repair system. Ultimately this would lead to several single stranded nicks in the DNA backbone, which could result, if positioned in close proximity, in double strand breaks. The findings of double and single strand breaks in V(D)J regions during somatic hypermutation support this theory [Papavisiliou and Schatz, 2000; Bross *et al.*, 2000;

Kong and Maizels, 2001]. After introducing the nick into the DNA segment, it would be nibbled back, resynthesized, and repaired as postulated for repair of damaged DNA. Error prone polymerases, such as polymerase  $\beta$ ,  $\delta$ , and  $\epsilon$  would then introduce mutations during the resynthesis step [Diaz *et al.*, 2001; Poltoratsky *et al.*, 2001; Rogozin *et al.*, 2001; Zeng *et al.*, 2001]. However, there is a difference between "normal" DNA repair and the resynthesis of a hypermutated DNA: the mutations need to get fixed, not eliminated. We suggest a subversion factor, which would interact with DNA repair factors, such as *pms2*, *mlh1*, or *msh2*. A general model for how this would occur has been suggested by Cascalho *et al.* [Cascalho *et al.*, 1998] (Fig. 2). Here, the newly synthesized DNA strand rather than the pre-existing one, is used as a repair template and the mismatches would be corrected in the reverse (Fig. 2B). In another suggested mechanism, epigenetic marks, which distinguish parental from newly synthesized strands, are removed and mismatch repair would choose randomly between the two strands for its template (Fig. 2D). In this case, a base substitution stands a 50:50 chance of being immortalized. Yet another suggestion involves homologous DNA repair [Papavasiliou and Schatz, 2000].

However, contradicting mutation rates have been calculated in MMR knock out mice models, and this may depend on the approach used to measure mutation rates in those models [Frey *et al.*, 1998; Jacobs *et al.*, 1998; Rada *et al.*, 1998]. In these studies, a comparison of DNA sequences with wild type sequences or germline sequences build the basis for conclusions. Frequently, MMR deficient mice were injected with a particular antigen and antigen specific B-cells were collected after several rounds of antigen boosts [Frey *et al.*, 1998; Phung *et al.*, 1999; Kim *et al.*, 1999]. Alternatively, memory B-cells from peripheral blood samples of non-immunized mice were analyzed.



**Figure 2.** General model of fixing mismatches during somatic hypermutation. **A)** Normal postreplicative mismatch repair (MMR) in which the parental DNA strand serves as template to eliminate the mutation. **B)** Backward repair during somatic hypermutation: The newly synthesized DNA strand serves as template and MMR fixes the mutation in the parental strand. **C)** Mutation activation genes (MAG) introduce mutations in the parental DNA strand during somatic hypermutation, which would remain undetected and would be fixed in the newly synthesized DNA strand. **D)** Epigenetic marker loss causes undifferentiated integration of mutations by mutation activation genes (MAG) in the parental and newly synthesized DNA strand. MMR fixes mutations with a 50:50 chance in one of the DNA strands depending on which strand would serve as repair template. (modified after Shannon and Weigert, 1998)

This approach should reduce the selection pressure towards the antigen and the antibody specificities analyzed should represent their "natural" antigen environment [Cascalho *et al.*, 1998]. Subsequently, selected V-regions were amplified by PCR and the products were sequenced. Frey *et al.* [Frey *et al.*, 1998] and others [Jacobs *et al.*, 1998; Phung *et al.*,

1998] found no difference in mutation frequencies between MSH2-deficient mice and wild type mice when they analyzed splenic germinal center cells. But when the same groups analyzed Peyer's patch germinal center B cells from the same mice, they observed a significant reduction in mutation frequencies [Frey *et al.*, 1998; Phung *et al.*, 1998]. It can thus be concluded that in order to study the influence of mismatch repair proteins in somatic hypermutation, a system without selection pressure by antigen would be desirable.

In addition to the altered mutation frequencies, mutational patterns seem to differ depending on the MMR protein that is knocked out. MSH2 or MSH6 knock out mice show the majority of mutations in GC base pairs embedded in the WRCY motif, leading to G → A or C → T transitions [Phung *et al.*, 1998; Rada *et al.*, 1998]. Whereas mice deficient in MLH1 or PMS2 do not exhibit dramatic alterations in their mutation spectra, but still reveal decreased mutation rates [Cascalho *et al.*, 1998; Phung *et al.*, 1999].

### **1.3. Mutation rates measurements**

As already mentioned, most of the MMR data in somatic hypermutation were accumulated in knock out mouse models where selection pressure towards the antigen will influence the read out of mutation frequencies and mutation patterns. However, generally one has to consider what classic population genetics taught about mutation frequencies [Wabl *et al.*, 1987]. Factors like mutation, selection, migration, drift and recombination influence the frequency of an allele in a population. Especially for *in vivo* experiments the "selection" factor influences the read out, since clonal selection is the

basic fact of the immune system. Immunizing mice drives the clonal selection process and it seems obvious that mutation frequencies and mutation patterns depend on the nature of the antigen, the application form and entrance way of the antigen, the B cell type analyzed and the tissue environment B cells were taken from. Further, the distinct definition between "mutation frequency" and "mutation rate" needs to remain clear. A high mutation frequency does not necessarily result in a high mutation rate as shown by Wabl *et al.* [Wabl *et al.*, 1987].

Taking these issues into consideration, the *in vitro* compartmentalization test offers the opportunity to avoid the influence of selection, migration, and drift in cell culture experiments [von Borstel *et al.*, 1971]. The original compartmentalization test ( $P_0$  method) by Luria and Delbrück depends on a majority of clones that do not contain any mutants [Luria and Delbrück, 1943]. Because one needs to estimate the average number of mutational events per clone, high numbers of individual clones need to be included per experiment—a time-consuming undertaking [Wabl *et al.*, 1987]. An experimental solution to the so-called Luria-Delbrück fluctuation problem in somatic hypermutation was first achieved by Bachl *et al.* [Bachl *et al.*, 1999]. Here a hypermutating cell line was stably transfected with an indicator plasmid, which contained all *cis*-acting elements to target the mutator. Mutator action was detected by the reversion of an amber stop codon embedded in the RGYW hot spot motif for somatic hypermutation in a green fluorescence protein gene (GFP). Based on one clone, a sizable cell population was grown up and then purged from pre-existing revertants by fluorescence activated cell sorter (FACS). Sorted cells were taken back into culture and expanded for some time. As a result, the number of revertants increased linearly with time. Even though this method

could avoid the Luria-Delbrück fluctuation effect, it still had to deal with certain limitations of positional effects of the transfected reporter.

#### **1.4. Specific Aims**

For this thesis, I had three specific aims: (1) Develop a retroviral based point mutation indicator system. (2) Characterize the IgH major intronic enhancer and other enhancer elements, as *cis*-elements in hypermutation. And (3) Mismatch repair in somatic hypermutation in non-B cells.

##### **1.4.1. Specific Aim 1: Develop a retroviral based point mutation indicator system**

To measure mutation rates fast and reliably, I set out to develop a retroviral indicator system for hypermutation that can be used in cell lines and would monitor mutation rates and transcription rates simultaneously. Furthermore, mutator activity should be easily detected by flow cytometry. Because retroviruses are able to infect high numbers of cells, and because the retroviral constructs integrate at different sites of the host cell genome, a position effect is less likely than for low numbers of stable transfectants. Further, working with high numbers of infected cells avoids time and material consuming compartmentalization tests, which are necessary when events follow a Poisson distribution; instead, a normal (Gauss) distribution can be expected. Finally, working with bulk cultures avoids time consuming single cell subcloning.

#### **1.4.2. Specific Aim 2: Characterize the IgH major intronic enhancer and other enhancer elements, as cis-elements in hypermutation**

With a mutation indicator system in hand that allows to simultaneously monitor transcription rates and mutation rates, the plan was to study the role of *cis*-acting elements in targeting somatic hypermutation. Because it is known that hypermutation rates correlate with transcription rates, I set out to study the ability of various immunoglobulin and other enhancer elements to target the mutator complex to the indicator gene; and their function as transcriptional enhancers. This would be done by placing enhancer elements at various locations within the reporter construct.

#### **1.4.3. Specific Aim 3: Mismatch repair in somatic hypermutation in non-B cells**

MMR deficiency leads to a substantial reduction of the mutation rate in the V-region of hypermutating B cells. Consequently, we hypothesized that the mismatch repair system is co-opted in hypermutating B cells, and that it actively fixes base mismatches in V regions rather than repairing them. Since fixing mismatches needs to be restricted to activated B cells, it was suggested that one or more B cell specific factors would be involved in the process. However, in non-B cells, the loss of MMR should increase the mutation rate, or, if it is very high, not influence it at all. I set out to test these assumptions in a kidney cell line, which allows regulated expression of the mismatch repair protein MLH1. Mutation frequencies would be measured by using retroviral point mutation indicator vectors in the presence and absence of MMR.

## 2. Materials and Methods

### 2.1. Materials

#### 2.1.1. Instruments

Centrifuge: Sorvall RC-5C, rotor: GSA, SS34 [Sorvall, Wilmington, CA]  
TechnoSpin R, rotor: microplate carrier [Sorvall, Wilmington, CA]

PCR-machines: PTC-100 Peltier Thermal Cycler [MJ Research, Watertown, MA]  
PTC-200 Peltier Thermal Cycler [MJ Research, Watertown, MA]

Gel-electrophoreses: Hoefer Scientific instruments [Hoefer Sci., San Francisco, CA]

Photo documentation: UV-lamp: Prep I, Fotodyne [Fotodyne, Foster City, CA]  
Camera: Fotodyne and Polaroid film

Spectrophotometer: Spectronic 601 [Bausch and Lomb, San Clemente, CA]

Culture-shaker: Innova 4000 [New Brunswick scientific, Princeton, NJ]  
C24 [New Brunswick scientific, Princeton, NJ]

Incubator: VWR 1530 [VWR Scientific Inc., Hayward, CA]  
Isotemp 500 [Fisher Scientific, San Francisco, CA]  
IS-61 [American Scientific, La Jola, CA]

Cell culture-incubator: B5060-EK-CO<sub>2</sub> [Haeraeus, Diagnostic Product Corporation,  
San Francisco, CA]

pH-meter: pH-Meter [Beckmann, Diagnostic Product Corporation, San  
Francisco, CA]

FACS MoFlow<sup>®</sup> [Cytomation, Fort Collins, CL]  
FACS Vantage<sup>®</sup> [Becton Dickinson, Fremont, CA]

#### 2.1.2. Enzymes and reaction kits

All restriction enzymes were obtained from New England Biolabs [New England Biolabs, Beverly, MA]. Taq- and Pfu-DNA polymerases were obtained from Invitrogen [Invitrogen, Carlsbad, CA].



The following reaction kits were used for this thesis:

Kit type	Maker
Plasmid isolation kits:	
QIAprep <sup>®</sup> spin Miniprep kit	Qiagen, Valencia, CA
QIAfilter <sup>®</sup> Plasmid Maxi kit	
QIAEX II <sup>®</sup> gel extraction kit	Qiagen, Valencia, CA
Oligotex <sup>®</sup> mRNA Mini kit	Qiagen, Valencia, CA
DNeasy Tissue kit	Qiagen, Valencia, CA
SuperScript <sup>®</sup> III RNase H <sup>-</sup> RT	Invitrogen, Carlsbad, CA
QuickChange <sup>®</sup> Site-Directed Mutagenesis	Stratagene, La Jolla, CA

### 2.1.3. Antibodies and enzyme conjugated anti-antibodies

AID of human or mouse origin was detected with a polyclonal rabbit anti-AID antibody, which was a kind gift of F. W. Alt [Chaudhuri *et al.*, 2003]. AID antibodies were detected in western blots with goat anti-rabbit IgG (H+L)-HRP (human and mouse adsorbed) antibodies [Southern Biotechnology Associates Inc., Birmingham, AL]. To detect human MLH1, monoclonal mouse anti-human MLH1 antibody [BD Pharmingen, San Jose, CA] was used, followed by HRP-conjugated goat anti-mouse IgG (H+L) polyclonal antibody [Southern Biotechnology Associates Inc., Birmingham, AL]. A polyclonal rabbit anti-mouse actin serum [Oncogene Research Products, San Diego, CA] was used to detect mouse actin in loading controls. The anti-actin Ab was detected with a goat anti-rabbit IgG Ab coupled to HRP [Southern Biotechnology Associates Inc., Birmingham, AL].

#### 2.1.4. Bacterial strains

In cloning strategies and for plasmid storage, *E.coli* TOP10 cells [Invitrogen, Carlsbad, CA] were used. As a recipient for recombinant DNA modified by site directed mutagenesis with the Qiaquick Change<sup>®</sup> kit, Epicurian coli<sup>®</sup> XL 1-Blue supercompetent cells were used [Stratagene, La Jolla, CA; Bullock *et al.*, 1987].

<i>E.coli</i> strain	Genotype
Top10	F <sup>-</sup> mcrA $\Delta$ (mrr-hsdRMS-mcrBC) $\Delta$ 80lacZ $\Delta$ M15 $\Delta$ lacX74 recA1 deoR araD139 $\Delta$ (ara-leu)7697 galU galK rpsL (Str <sup>R</sup> ) endA1 nupG
Epicurian coli <sup>®</sup> XL 1-Blue	RecA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F' proAB lac <sup>q</sup> Z $\Delta$ M15 Tn10 (Tet <sup>r</sup> )]

#### 2.1.5. Plasmid and retroviral vector construction

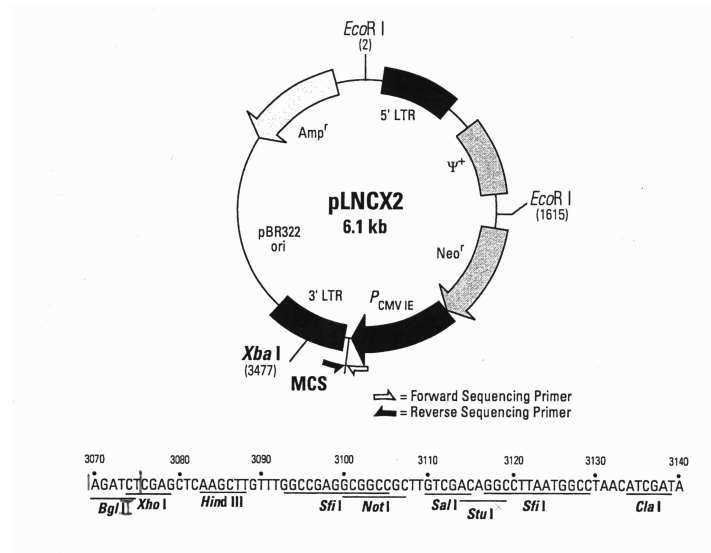
To create pLNCX2-eGFP, the GFP gene of pEGFP-N1 [Clontech, San Jose, CA] was excised from the *Xho*I and *Not*I restriction sites and placed into the according restriction sites of pLNCX2 [Clontech, San Jose, CA; figure 4]. The coding sequence of activation induced cytidine deaminase (AID) was reverse transcribed from mRNA of the murine 18-81 pre B-cell line, which expresses AID constitutively [Wabl *et al.*, 1985; Bachl *et al.*, 2001]. For specific amplification of AID from cDNA the primer set AID-fwd and AID-back was used [2.1.5]. The PCR product was cloned into the pCR2.1-TOPO<sup>®</sup> plasmid [Stratagen, La Jolla, CA] before it was transferred into the *Hind*III and *Not*I site of the retroviral pLNCX2 vector to create pLNCX2-mAID.

All mutation indicator plasmids were based on the retroviral CRU5-IRES-eGFP vector (102.IG) [J.B. Lorens, unpublished, figure 4], which is derived from the p96-dsG plasmid [Lorens *et al.*, 2000].

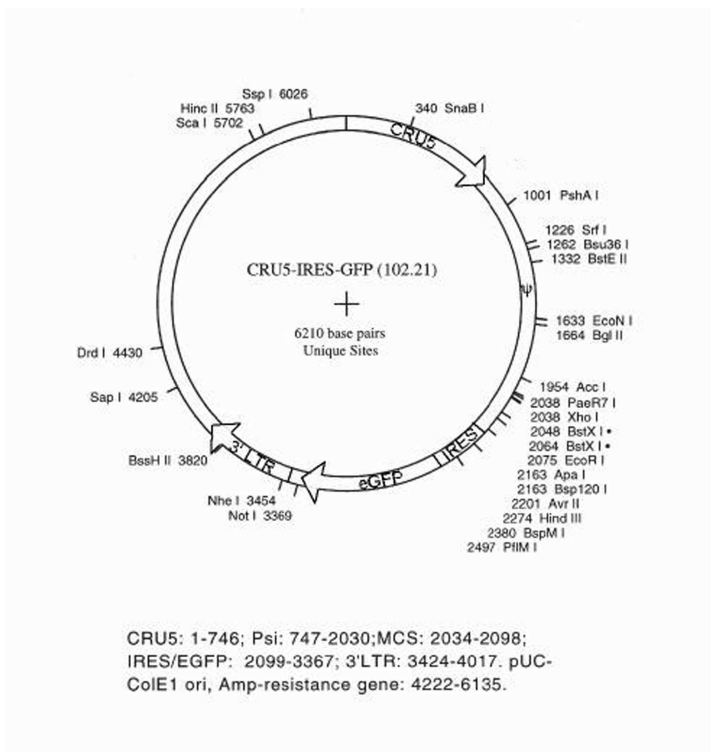
**CRU5-YFP-IRES-GFP vectors.** To create the CRU5-YFP (102.Y) single color positive control, the eYFP coding sequence of the pEYFP plasmid [Clontech, San Jose, CA] replaced the IRES-GFP fragment between the *XhoI* and *NotI* restriction sites in 102.IG. The 102.Y-IG two-color positive control vector was created by first amplifying the eYFP gene from pEYFP using the YFP-*XhoI* and the YFP-*BstXI* primer [2.1.5]. The PCR product was introduced into the 102.IG vector via the *XhoI* and *BstXI* restriction sites. To prepare the point mutation indicator gene, an amber stop-codon was introduced at base pair 321 into the eYFP coding sequence by site directed mutagenesis using the QuickChange<sup>®</sup> Site-Directed Mutagenesis kit with the YFP-stop oligomere [2.1.5]. Plasmid pEYFP served as template for the reaction and the created sequence was designated Y(X). Y(X) was PCR amplified using the YFP-*XhoI* and YFP-*BstXI* primer set and introduced into the *XhoI* and *BstXI* restriction sites of 102.IG to create the 102.Y(X)-IG vector.

**CRU5-DsRed-IRES-GFP constructs.** In our cloning strategy to make point mutation indicator plasmids, we mutated the serine codon UCC at aa position 179 to a UCA codon which also encoded for serine, in the DsRed2 gene [Klasen and Wabl, 2004]. This silent point mutation destroyed the *BstXI* restriction site in the DsRed2 gene. To obtain this mutant, pDsRed2-1 served as PCR template for site directed PCR mutagenesis using the DsRed-mut 537 oligomer [2.1.5] and the QuickChange<sup>®</sup> Site-Directed Mutagenesis kit.

A.



B.



**Figure 3.** Plasmid maps of **(A)** pLNCX2 [Clontech, San Jose, CA] and **(B)** 102.21 (CRU5-IRES-eGFP) [J.B. Lorens, Rigel Inc., San Francisco, CA]. CRU5 = composite minimal CMV and MoMuLV 5'LTR promoter fusion, □ = (extended) packaging signal, IRES = internal ribosome binding site, eGFP = enhanced green fluorescence protein, CMV IE = full size immediate early CMV promoter, Neo<sup>r</sup> = neomycin resistance gene, LTR = long terminal repeat

Because the procedure replaced the cytidine by an adenosine at nucleotide position 537, we designated the mutated gene DsRed537. The sequence of the mutated gene was verified before and after cloning into mammalian vectors. The mutated and original DsRed genes were cloned into the *XhoI* and *NotI* restriction sites of the CRU5-IRES-GFP, replacing the IRES-GFP fragment. This way plasmids CRU5-DsRed2 (102.R) and CRU5-DsRed537 (102.Rm) were created. The same strategy was used to mutate and clone the DsRed-Express1 gene based on the pDsRed-Express1 vector [BD Biosciences Clontech, Palo Alto, CA] to create CRU5-DsRed-Express1 (102. REx) and CRU5-Red-Express537 (102. RExm) [Klasen and Wabl, 2004]. Replacing the IRES-GFP fragment between the *XhoI* and *NotI* site of 102.IG with the HcRed coding sequence from pHcRed1-N 1/1 [Clontech, San Jose, CA] created the plasmid CRU5-HcRed (102.Hc). However, the DsRed537 coding sequence was PCR amplified using the Red-*XhoI* 5' primer and the Red-*BstXI* 3' primer [2.1.5]. The PCR fragment contained a 5' *XhoI* and a 3' *BstXI* restriction site and was placed upstream of the IRES element into 102.IG to create 102.Rm-IG (CRU5-Red537-IRES-eGFP). An amber stop-codon was introduced into the DsRed537 coding sequence at nucleotide position 519 by site directed PCR mutagenesis using the Red-stop oligomere [2.1.5]. The resulting gene was designated Rm(X), which was PCR amplified using the primer pair Red-*XhoI* 5' and Red-*BstXI* 3'. The PCR product replaced the DsRed537 sequence in 102.Rm-IG to create the 102.Rm(X)-IG vector.

The 1 kb E $\mu$  intronic enhancer was PCR amplified from the p $\mu$  plasmid [Grosschedl *et al.*, 1984] using the oligomeres Emu - *BstXI* 5' and Emu - *BstXI* 3' [2.1.5]. The 721 bp fragment of the kappa intronic enhancer ( $\kappa$ i) was PCR amplified from the pAkV- $\kappa$ i vector

[C.L. Wang, personal communications; Klix *et al.*, 1998] using the oligomers  $\square$ i - *Bst*XI 5' and  $\square$ i - *Bst*XI 3'. The 236 bp SV40 enhancer was amplified from the pGL3-enhancer plasmid [Stratagene, La Jolla, CA] using the oligomers SV40 COR 5' and SV40 COR 3' [2.1.5]. The 808 bp comprising kappa 3' enhancer ( $\square$ 3') [Max *et al.*, 1981] was PCR amplified from chromosomal DNA of the 18-81 cell line using the primers Kappa *Bst*XI 5' and Kappa *Bst*XI 3'. All enhancer PCR fragments contained a *Bst*XI restriction site on the 5' and 3' end and were placed into the *Bst*XI site of the 102.Rm-IG or 102.Rm(X)-IG vector, respectively. To create the 102.Rm(X)-IG- $E\mu$  vector the  $E\mu$  enhancer was amplified with  $E\mu$ -*Not*I forw and  $E\mu$ -*Not*I back oligomers [2.1.5] and placed into the *Not*I restriction site of the 102.Rm(X)-IG plasmid.

When the GFP gene was used as point mutation indicator gene in the CRU5-IRES-GFP vector, the amber stop-codon was introduced at base pair 321 into the GFP gene. The 102.IG plasmid served as template for the site directed mutagenesis reaction using the YFP-stop oligomere to introduce the stop-codon, which generated the 102.IG(X) vector. Using the cloning strategy previously mentioned, the DsRed537 gene was introduced into 102.IG(X) between the *Xho*I and *Bst*XI restriction site to create 102.Rm-IG(X). The  $E\mu$  enhancer was introduced into the 102.Rm-IG(X) construct as described above via the *Not*I site positioned downstream of the GFP gene. This engineering step resulted in the 102.Rm-IG(X)- $E\mu$  plasmid.

The vectors pGFP-Ipuro (CRU5-GFP-IRES-Puromycin<sup>r</sup>), pGFP\*-Ipuro and pGFP\*-Emu-Ipuro were kind gifts of C. L. Wang [Wang *et al.*, 2004]. To create the vectors 102.Rm-IN (CRU5-DsRed537-IRES-Neomycin<sup>r</sup>) and 102.Rm(X)-IN, the IRES-GFP fragment of

the corresponding 102.Rm-IG vectors was replaced with the IRES-Neomycin fragment between the *EcoRI* and *NheI* restriction sites of the pBMN-Z-I-Neo vector [kindly provided by Gary P. Nolan, Stanford University School of Medicine, Stanford, CA. Plasmid engineering by F.J.X. Spillmann, personal communications]. All constructs were verified by DNA sequencing.

### 2.1.6. Oligonucleotides

#### Cloning primers

AID-fwd

5'- CGC CCA TGG TGG TGG GCC ACC ATG GAC AGC CTT CTG ATG -3'

AID-back

5'- ATG GTT CCA ATT TAA TGG TCA AAA TCC CAA CAT ACG AAA TGC -3'

YFP-*XhoI* primer

Ov (XhoI site) annealing region  
5'- TTA AG (CTC GAG) GCT CTA GAA CTA GTG GAT CCC CGC G -3'

YFP-*BstXI* primer

Ov (BstXI site b) annealing region  
5'- TA (CCA ATT TAA TGG) TTA CTT GTA CAG CTC GTC CAT GCC -3'

YFP-stop oligomer

5'- TCG GCG CGG GTC TGC TAG TTG CCG TCG TCC TT -3'

DsRed-mut 537 oligomere

5'- TGG TGG AGT TCA AGT CAA TCT ACA TGG CCA AGA AGC C -3'

Red-*XhoI* 5' primer

Ov (XhoI) annealing region  
5'- AA (CTC GAG) CGC CAC CAT GGC CTC CTC CG -3'

Red-*BstXI* 3' primer

Ov (BstXI site b) annealing region  
5'- AA (CCA ATT TAA TGG) CTA CAG GAA CAG GTG GTG GCG G -3'

Red-stop oligomer

5'- GAC GGC GGC CAC TAG CTG GTG GAG TTC AAG -3'

Emu - *Bst*XI 5'

Ov (*Bst*XI site b) annealing region

5'- TT (CCA TTA AAT TGG) TCT AGA GAG GTC TGG TGG AG -3'

Emu - *Bst*XI 3'

Ov (*Bst*XI site b) annealing region

5'- TA (CCA ATT TAA TGG) TCT AGA TAA TTG CAT TCA TTT AAA AAA A -3'

□i - *Bst*XI 5'

Ov (*Bst*XI site b) annealing region

5'- TT ACG (CCA TTA AAT TGG) AGC TCA AAC CAG CTT AGG C -3'

□i - *Bst*XI 3'

Ov (*Bst*XI site b) annealing region

5'- ATG GTT (CCA ATT TAA TGG) CTA GAA CGT GTC TGG GCC -3'

SV40 COR 5'

Ov (*Bst*XI site b) annealing region

5'- TT ACG (CCA TTA AAT TGG) CGA TGG AGC GGA GAA TG -3'

SV40 COR 3'

Ov (*Bst*XI site b) annealing region

5'- TG GTT (CCA ATT TAA TGG) GCT GTG GAA TGT GTG TC -3'

Kappa *Bst*XI 5'

Ov (*Bst*XI site b) annealing region

5'- TT ACG (CCA TTA AAT TGG) AGC TTA ATG TAT ATA ATC TTT TAG -3'

Kappa *Bst*XI 3'

Ov (*Bst*XI site b) annealing region

5'- ATG GTT (CCA ATT TAA TGG) GGT GTC CAG AAA TAT TCT G -3'

*Eμ* - *Not*I forw.

Ov (*Not*I) annealing region

5'- AAT (GC GGCC GC) TCT AGA GAG GTC TGG TGG AG -3'

*Eμ* - *Not*I back.

Ov (*Not*I) annealing region

5'- AAT (GC GGCC GC) TCT AGA TAA TTG CAT TCA TTT AAA AAA A -3'



## Sequencing primers

102. MCS-5'

5'- GAT CTC GAG GGA TCC ACC ACC ATG G -3'

IRES-back

5'- CGC TCT AGA ACT AGT GG -3'

pLNCX2 seq5'

5'- AGC TCG TTT AGT GAA CCG TCA GAT C -3'

pLNCX2 seq3'

5'- ACC TAC AGG TGG GGT CTT TCA TTC CC -3'

## 2.2. Cell culture based methods

### 2.2.1. Mammalian cell lines and cell culture conditions

**Murine 18-81 lymphocyte cell line [Wabl *et al.*, 1984].** The 18-81 pre B-cell line continuously hypermutates in culture. Because its light chain locus is not rearranged and a pre B-cell receptor is not expressed on the cell surface, the cell line is considered to be a pre B-cell line. To obtain this cell line, murine pre B-cells were transformed with the v-abelson virus. In this thesis the 18-81.10-17.93 sublone was used and cultured in RPMI 1680 medium supplemented with 10% (v/v) fetal bovine serum (FBS), 2 mM L-glutamine, 1 mM sodium pyruvate, 50 IU/mL penicillin-streptomycin, and 50 mM  $\beta$ -mercaptoethanol. At a temperature of 37°C and 5% CO<sub>2</sub> cells divided three times in 24 hours.

**Murine 70Z/3 pre B-cell line.** The 70Z/3 cell line is a hypermutation inactive pre B-cell line and was cultured under the same conditions like the 18-81 line.

**Phoenix eco- and amphotropic cell line [Swift *et al.*, 1999].** The Phoenix E and Phoenix A cell line was a kind gift of Gary P. Nolan [Stanford University School of Medicine, Department of Pharmacology, Stanford, CA]. Both cell lines are based on the human kidney cell line 293T and carry the MuMoLV packaging genes [<http://www.stanford.edu/group/nolan/>]. The Phoenix cell line was cultured in DMEM medium containing 10% (v/v) FBS (heat inactivated), 4 mM L-glutamine and 50 IU/mL penicillin-streptomycin at 10% CO<sub>2</sub> and 37°C.

**Murine NIH/3T3 fibroblast cell line ATCC CRL-1658.** The NIH/3T3 cell line was obtained from ATCC [[www.atcc.org](http://www.atcc.org)] and was cultured like Phoenix cells.

**Human embryonic 293T-Tet-Off-hMLH1 kidney cell line.** The human 293T-Tet-Off-hMLH1 cell line was maintained in culture as characterized by Trojan *et al.* [Trojan *et al.*, 2002]. Briefly, cells were cultured at 37°C and 10% CO<sub>2</sub> in DMEM medium supplemented with 4 mM L-glutamine, 100 U/ml penicillin-streptomycin, 10% Tet-system approved FCS [Clontech, San Jose, CA], 300 µg/ml hygromycin B [Roche Applied Science, Indianapolis, IN], and 100 µg/ml Zeocin [Invitrogen, Carlsbad, CA]. Tet-response element controlled hMLH1 expression was inhibited with 50 ng/ml doxycyclin [BD Biosciences Clontech, Palo Alto, CA].

### **2.2.2. Cryostorage of cell lines**

18-81, 70Z/3, and NIH/3T3 cells were frozen in their complete culture medium, which was supplemented with 5% (v/v) DMSO.  $2 \times 10^7$  cells were frozen in 800 µl freezing medium in a cryotube [Nunc, Rochester, NY]. Phoenix cells were frozen in 90% FCS

(heat inactivated) containing 10% (v/v) DMSO.  $5 \times 10^6$  cells were frozen in a volume of 800  $\mu$ l freezing medium. All cell lines were stored in liquid Nitrogen Cryotank at -140°F.

### **2.2.3. Rapid production of retroviruses in Phoenix cells [Swift *et al.*, 1999].**

This method is applicable to both the ecotropic and the amphotropic Phoenix packaging cell line.

**Preparation and transfection of Phoenix packaging cells.** 24 hours prior transfection,  $2 \times 10^5$  Phoenix cells were seeded into one 6 cm well of a Falcon multiwell plate [Becton Dickinson, New Jersey, NY]. Transfections were performed with 3  $\mu$ l FuGENE<sup>®</sup> 6 reagent [Roche Applied Science, Indianapolis, IN] and 1  $\mu$ g retroviral plasmid DNA (Qiagen maxiprep quality) according to the manufacture's manual. Depending on the target cell line, 24 hours after transfection the medium was changed to the medium conditions of the cell line to transduce. On average, Phoenix cells produced a virus titer of  $5 \times 10^5$  infectious virus particles per ml (IVP/ml) after 48 hours of incubation at 37°C and 5% or 10% CO<sub>2</sub> [Pear *et al.*, 1993]. Optional cells were transferred to a 32°C incubator, which slowed down cell division without significantly decreasing virus production. This temperature has been suggested to increase the stability of retroviral particles [Kotani *et al.*, 1994].

### **2.2.4. Spin infection of adherent cells**

Twelve to sixteen hours prior infection, adherent cells (e.g. NIH/3T3 cells) were trypsinized and  $5 \times 10^5$  cells were transferred into one 6 cm well of a Falcon multiwell plate [Becton Dickinson, New Jersey, NY]. To infect target cells, the culture medium of the target cells was removed and the retrovirus containing supernatant from one 6 cm

well of transiently transfected Phoenix E cells was added by filtering it through a 0.45  $\mu\text{m}$  surfactant free cellulose acetate membrane filter [Nalgene, New York, NY]. 4  $\mu\text{g}/\text{ml}$  Polybrene $\square$  (hexadimethrine bromide) [Sigma, St. Louis, MO] were added to the supernatant and the 6 well plates were sealed with parafilm. Plates were centrifuged at 1200 x g for 60 min at RT on a microplate carrier [Sorvall, Wilmington, CA]. After centrifugation, parafilm was removed and cells were placed in a 37 $\square$ C incubator over night. The next morning, the culture supernatant was replaced by fresh complete medium according to the culture conditions of the target cell line.

#### **2.2.5. Spin infection of suspension cells**

To infect  $1 \times 10^6$  exponentially growing target cells (e.g. 18-81), they were transferred into one 6 cm well of a Falcon multiwell plate [Becton Dickinson, New Jersey, NY]. Cells were centrifuged on a microplate carrier [Sorvall, Wilmington, CA] at 1200 x g for 5 min at RT. After centrifugation, the supernatant of the target cells was carefully removed. The retroviral supernatant from three samples was pooled and filtered through a 0.45  $\mu\text{m}$  surfactant free cellulose acetate membrane filter [Nalgene, New York, NY]. The retroviral supernatant was added to the target cells and supplemented with 4  $\mu\text{g}/\text{ml}$  Polybrene $\square$  (hexadimethrine bromide) [Sigma, St. Louis, MO]. The plates were sealed with parafilm and centrifuged on a microplate carrier at 1200 x g for 60 min at RT. After centrifugation, the parafilm was removed and cells were placed in a 37 $\square$ C incubator over night. The next morning, the cells were centrifuged and the cell pellet was resuspended in fresh complete medium according to the culture conditions of the target cell line.

### **2.2.6. Fluorescent adsorbent cell sorting (FACS)**

High speed cell sorting was performed either on a MoFlow<sup>®</sup> (Cytomation, Fort Collins, CL) or a FACS Vantage<sup>®</sup> (Becton Dickinson, Fremont, CA). In preparation to a FACS, adherent cells were trypsinized and concentrated by centrifugation (1200 x g, 5 min, 4°C). The cells were resuspended in approximately 1 ml of serum free medium and filtered through a 0.45  $\mu$ m cell separator [BD Falcon<sup>®</sup>, BD Biosciences, Bedford, MA]. Except for trypsinizing, suspension cells were prepared the same way. Until sorting, cells were kept on ice. The cells were sorted into FBS or complete culture medium supplemented with 1 mM HEPES buffer. After sorting, cells were washed twice with complete culture medium.

**FACS settings.** To achieve a separation of the GFP signal from the YFP signal during FACS, a 488 nm excitation wavelength at 200 mW was applied. Light passed a 555 long pass beam splitter before the GFP signal in channel 1 passed a 510/20 short pass filter in combination with a 530/30 short pass filter for YFP in channel 2. The DsRed signal was separated from the GFP signal by applying a 530/30 short pass filter in FL1 for GFP and a 580/30 short pass filter for DsRed in FL2. HcRed excitation required a separate 568 nm laser line at 200 mW in combination with a 640 nm long pass filter in the FL2 channel.

### **2.2.8. Analytical flow cytometry**

For analytical flow cytometry, cells were prepared similar to that for a FACS. The analytical flow cytometry was performed on a FACS Calibur<sup>®</sup> [BD Bioscience, San Jose, CA].

## **2.3. Molecularbiological Methods**

### **2.3.1. Strain maintenance and strain conservation**

*E.coli* strains were maintained on Luria broth (LB) agar plates, which were supplemented with the appropriate antibiotics for plasmid selection. The plates were stored up to three weeks at 4°C, before the *E.coli* strains were transferred on fresh LB agar plates.

**Cryo-conservation of *E.coli* strains.** For this procedure, 400 µl of a fresh *E.coli* suspension culture were mixed with 200 µl of 50% (v/v) glycerol (sterile), transferred into a cryo-tube [Nunc, Rochester, NY], and frozen at - 80°C.

### **2.3.2. Making chemically competent *E.coli* cells (RbCl method by Hanahan)**

Ready-to-use competent *E.coli* Top10 cells were purchased from Invitrogen [Invitrogen, Carlsbad, CA] and used according to the manufacturer's manual. Self-made chemically competent *E.coli* cells were produced as described by Hanahan [Hanahan, 1976].

### **2.3.3. Heath shock transformation of competent *E.coli* cells**

Commercially available *E.coli* Top10 cells were transformed according to the manufacture's manual [Invitrogen, Carlsbad, CA]. Self made *E. coli* cells were thawed on ice. Up to 50 ng plasmid DNA were carefully mixed with the competent cells and incubated for 30 min on ice. To heat shock the cells, the sample was incubated for 1 min at 42°C in a water bath. Afterwards, the cells were immediately chilled for at least 2 min on ice. To regenerate the antibiotic resistance encoded by the plasmid, 400 µl SOC medium [Sambrook *et al.*, 1989] were added to the cells, followed by an incubation step of 1 hour at 37°C at 200 rpm. Up to 200 µl of the cell suspension were inoculated on a

LB agar plate supplemented with the appropriate antibiotics and incubated over night at 37°C.

#### **2.3.4. Purification of plasmid DNA**

For standard cloning procedures, plasmid DNA was purified from *E.coli* cells of a 5 ml over-night culture according to the Qiaprep Mini plasmid purification protocol [Qiagen, Valencia, CA]. Cell culture grade plasmid DNA, i.e. for transfection experiments, was obtained by using the Qiaprep Midi- or Maxi- plasmid purification kit [Qiagen, Valencia, CA].

#### **2.3.5. Isolation of chromosomal DNA from mammalian cells**

Chromosomal DNA was isolated from  $5 \times 10^6$  cultured cells using the DNAeasy chromosomal DNA isolation kit [Qiagen, Valencia, CA] according to the manufactures manual.

#### **2.3.6. Isolation of polyA<sup>+</sup> mRNA**

The isolation of polyA<sup>+</sup> mRNA was performed according to the Oligotex<sup>®</sup> mRNA Mini kit protocol [Qiagen, Valencia, CA].

#### **2.3.7. Preparative and analytical gel electrophoresis**

DNA fragments were separated in 0.8% - 2% agarose gels containing ethidium bromide as described by Sambrook *et al.*, [Sambrook *et al.*, 1989]. DNA bands were visualized on an UV-transluminator at a wavelength of 360 nm. In case of a preparative gel electrophoresis, the DNA fragment of choice was cut out of the gel using a scalpel and DNA was subsequently extracted from the gel using the Qiaex II gel extraction kit according to the manufacturer [Qiagen, Valencia, CA].

### **2.3.8. Concentration measurements and purity estimations of plasmid DNA**

The DNA concentration of plasmid preparations was measured in agarose gels after electrophoresis at a wavelength of 360 nm. The fluorescence intensity of the sample band was compared with the fluorescence intensity of a band of known DNA content from either the High DNA Mass Ladder [Invitrogen, Carlsbad, CA] or the Low DNA Mass Ladder [Invitrogen, Carlsbad, CA] [Sambrook *et al.*, 1989]. The purity of DNA was estimated spectrophotometrically in quartz cuvettes by calculating the extinction ratio of the DNA at 260 nm and 280 nm [Müller *et al.*, 1993].

### **2.3.9. Endonuclear restriction digest of DNA**

DNA restriction digests were performed according to the guidelines of the restriction enzyme manufacturer. In compliance with time and temperature optima, 2 - 5 units of restriction enzyme were used per 1  $\mu$ g DNA.

### **2.3.10. DNA-Ligation**

In preparation for a DNA ligation, digested DNA fragments (insert) and vector DNA were purified using the Qiaex II gel extraction kit and the DNA concentration was measured on an agarose gel. Two different kinds of ligations were performed. In a so-called "sticky end" ligation, two DNA fragments with complementary ends were joined. In a "blunt-end" ligation two DNA fragments with straight ends were combined. The phosphodiester bridge between the DNA fragments was built by the T4 DNA-ligase [Qiagen, Valencia, CA], which was used according to the manufactures instructions. In general, a 3:1 molar ratio of "insert"- to vector - DNA was used in a sticky end ligation. For blunt end ligations an insert to vector ratio of 5:1 was used. After ligation, competent *E.coli* cells were transformed with the reaction product.



### 2.3.11. PCR amplification of DNA fragments

Polymerase chain reactions [Saiki *et al.*, 1985] were carried out in 200  $\mu$ l OPS thin wall PCR tubes [Online products for science, Petaluma, CA; www.op4s.com]. A standard PCR contained:

0.01 - 10 ng	template DNA
10 pmol	primer 1
10 pmol	primer 2
1.5 mM	MgCl <sub>2</sub>
0.2 mM/dNTP	dNTP-mix (1.25 mM/dNTP)
10% (v/v)	10x <i>Taq</i> DNA polymerase buffer
1 U	<i>Taq</i> DNA polymerase
ad 50 $\mu$ l	ddH <sub>2</sub> O

Amplifications were performed in Peltier Thermal Cyclers PTC-100 or PTC 200 [MJ Research, Watertown, MA]. Amplification profiles altered with application, but a standard program contained the following steps:

Denaturation of template DNA	94°C / 60 sec
Primer annealing	variable°C / 60 sec
Extension	72°C / 60 sec per 1 kb DNA length

In general, the three program steps were repeated 30 times. The amplification was completed with a 10 min incubation step at 72°C, which ensured fully synthesized PCR products. Annealing temperatures ( $T_p$ ) of primer oligonucleotides were calculated after the empiric formula:

$$T_p = 22 + 1.46 \times [2 \times (G + C) + A + T]$$

Alternatively to this calculation the online software "oligonucleotide properties calculator" was used [[www.biotechlab.nwu.edu/Oligocalc.htm](http://www.biotechlab.nwu.edu/Oligocalc.htm)]. This software is based on thermodynamic calculations by Breslauer *et al.*, [Breslauer *et al.*, 1986].

### **2.3.12. First strand cDNA synthesis**

cDNA was synthesized according to the manufactures guidelines of the SuperScript<sup>III</sup> RNase H<sup>-</sup> RT kit [Invitrogen, Carlsbad, CA].

## **2.4. Biochemical Methods**

### **2.4.1. Protein concentration measurements after Warburg and Christian**

This method by Warburg and Christian [Warburg and Christian, 1942/1942] is based on extinction measurements of protein solutions in equilibrated quartz cuvettes at an UV wavelength of 260 nm and 280 nm. The protein content of the solution is calculated according to the following formula:

$$C_{\text{protein}} \text{ (mg/ml)} = E_{280} \times 1.56 - E_{260} \times 0.76$$

### **2.4.2. Protein measurements after Bradford**

Protein concentrations were determined after the Bradford method [Sambrook *et al.*, 1989].

### **2.4.3. SDS polyacrylamide gelelectrophoresis after Laemmli**

The electrophoretic separation of proteins according to their molecular weight was performed in acrylamide gels as described by Laemmli [Laemmli, 1970]. Homogenous 10 cm wide and 8 cm high (small Biometra gel chamber [Biometra, Horsham, PA]) separating gels with acrylamide concentrations between 7,5% - 15% were used. All

stacking gels contained 5% acrylamide and were 30 mm long. Protein samples were mixed 1:1 (v/v) ratio with 2x Laemmli buffer and denatured by boiling the sample for 5 min at 95°C. Electrophoresis was performed at a current of 10 - 20 mA at RT.

#### **2.4.4. Silver staining of SDS polyacrylamide gels modified after Blum**

In order to silver stain [Blum *et al.*, 1988] a SDS-PAGE after electrophoresis, the gel was fixed over night in solution 1. The next day, the gel was initially incubated for 40 min in solution 2 (50% ethanol), and afterwards washed for 20 min in solution 3 (30% ethanol). After an one minute incubation in solution 4 (thiosulfate) three washing steps (each 30 sec) in double distilled water (ddH<sub>2</sub>O) followed. After a 20 min incubation time in solution 5 (silver/formaldehyde), two further washing steps (each 30 sec) in ddH<sub>2</sub>O followed. The gel was developed in solution 6 (carbonate/formaldehyde/thiosulfate) until the protein bands were clearly visible. The reaction was stopped by one quick wash step in ddH<sub>2</sub>O, immediately followed by a 15 min incubation in solution 7 (methanol/acetic acid). Before the gel was dried on a gel dryer, it was washed three times in water and incubated for 30 min in 3% glycerol.

Solution 1: 50% methanol  
12% acetic acid  
0.5 ml/l formaldehyde (37%)

Solution 2: 50% ethanol

Solution 3: 30% ethanol

Solution 4: 0.2 g/l Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub> x 5 H<sub>2</sub>O

Solution 5: 2,0 g/l AgNO<sub>3</sub>  
0.75 ml/l formaldehyde (37%)

Solution 6: 60 g/l NaCO<sub>3</sub>  
0.5 ml/l formaldehyde (37%)  
4.0 mg/L Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub> x 5 H<sub>2</sub>O

Solution 7: 10% methanol

#### 2.4.5. Coomassie staining of SDS polyacrylamide gels modified after Wilson

SDS polyacrylamide gels were incubated for at least 30 min at RT in Coomassie staining solution [Wilson, 1983]. To reduce the background, gels were incubated in the de-staining solution until clear proteins bands became visible.

Coomassie staining solution:

Serva blue B: 1.75 g  
Methanol: 454 ml  
ddH<sub>2</sub>O: 454 ml  
Acetic acid: 92 ml

De-staining solution:

Methanol: 25%  
Acetic acid: 10%

#### 2.4.6. Western blot modified after Towbin

Western blotting was used for specific identification of low concentrated proteins (in ng range). After the denatured proteins were separated on a SDS-PAGE, the protein bands were transferred on a Trans-Blot<sup>®</sup> pure nitrocellulose membrane [BioRad, Hercules, CA]. For this procedure, a Trans-Blot<sup>®</sup> wet blotting cell [BioRad, Hercules, CA] was used and blotting was performed for one hour at a current of 0.8 mA/cm<sup>2</sup> membrane [Towbin *et al.*, 1979]. After blotting, free binding sites on the nitrocellulose membrane were blocked in TBT buffer [Sambrook *et al.*, 1989] supplemented with 5% (w/v) skim milk powder over-night at 4°C. Next day, the detection antibody (1<sup>st</sup> antibody) was applied to the membrane in a 1:500 dilution (diluted in TBT buffer) and incubated for 3 - 16 h at 4°C. Afterwards, unbound antibody residues were removed by washing the membrane twice for 30 min in TBT buffer at RT. To specifically detect the first antibody,

the membrane was incubated for one hour at 37°C with the corresponding HRP-labeled secondary antibody (diluted 1:5000 in TBT buffer). Again, two washing steps for 30 min at RT in TBT buffer removed excess antibody.

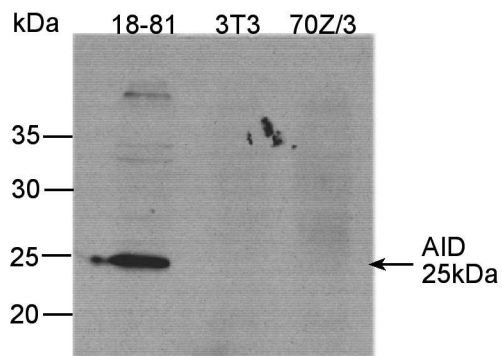
To detect the according protein band, the membrane was incubated with the colormetric substrate of the ECL Western blotting detection system [Amersham Bioscience, Buckingham, UK] according to the maker. The band was made visible on a Kodak Biomax MS film [Eastman Kodak company, Rochester, NY].

### 3. Results

#### 3.1. Specific Aim 1: Develop a retroviral based point mutation indicator system.

##### 3.1.1. Expression of activation induced cytidine deaminase (AID) in 18-81 cells

Activation induced cytidine deaminase (AID) is the absolutely required *trans*-acting factor to induce somatic hypermutation in activated B-cells. To verify that AID was expressed in the continuously hypermutating 18-81 cell line, equal amounts of whole cell lysate from three cell lines was separated on a 10% SDS-PAGE and AID expression was detected by western blot. The NIH/3T3 fibroblast control did not show any AID expression nor did the 70Z/3 hybridoma cell line (Fig. 4). Only the 18-81 cell line expressed the 25 kDa protein. Smaller products were detected in some AID Western blots (data not shown) those products were interpreted as splice variants rather than degradation products. Later, this assumption was verified by McCarthy *et al.* [McCarthy *et al.*, 2003].



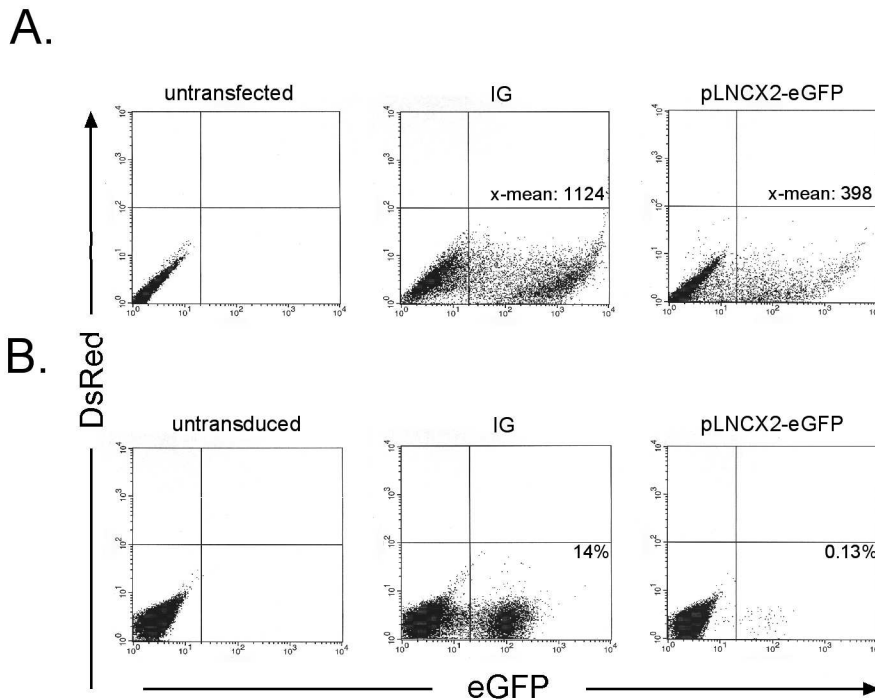
**Figure 4.** Western blot analysis of endogenous AID expression in 18-81 cells, NIH/3T3 cells, and in 70Z/3 cells. 100  $\mu$ g whole cell lysate (estimated by Bradford method) of various cell lines was separated on a 10% SDS-PAGE and AID was detected using a polyclonal rabbit-anti-AID serum. Lane 1, 18-81 lysate; lane 2, NIH/3T3 lysate, lane 3: 70Z/3 lysate. To the left of the blot, molecular weight markers; to the right, position of AID.

### 3.1.2. High virus titer production

Initially, the MoMuLV based retroviral vectors pLNCX2-eGFP and CRU5-IRES-eGFP were tested for high virus titer production in the Phoenix E packaging cell line. The CRU5-IRES-eGFP vector [J.B. Lorens, unpublished] is derived from the p96-dsG plasmid [Lorens *et al.*, 2000]. The CRU5 vector contains a composite CMV promoter fused to the transcriptional start site of the MoMuLV R-U5 region (CRU5) located in the 5'LTR (Fig. 3). Whereas the CMV promoter allows high titer virus production during transient transfection of Phoenix packaging cells, the provirus is expressed by the retroviral MoMuLV LTR promoter after its integration into the host cell genome [Lorens *et al.*, 2000; Swift *et al.*, 1999]. Further, the CRU5 vector contains a so-called extended  $\Psi$  packaging signal for improved encapsidation of retroviral mRNA (Fig. 3). The pLNCX2-eGFP vector is based on the pLNCX2 vector [BD Biosciences Clontech, Palo Alto, CA] and contains a regular MoMuLV 5' LTR promoter for virus production in the packaging cell line and a regular MoMuLV packaging signal (Fig. 3). After integration of the pLNCX2-eGFP provirus into the host cell genome, the CMV immediate early promoter triggers GFP expression (Fig. 3 and Fig. 6).

Virus production rates of the two vector types were compared by transient transfection of the Phoenix E packaging cell line. After three days of incubation, transiently transfected Phoenix cells were analyzed for GFP expression by analytical flow cytometry. The average transfection efficiency of the 102.IG construct was approximately three times higher than for the pLNCX2-eGFP construct (Fig. 5A). Further, the pLNCX2-eGFP construct showed an approximately three times lower relative GFP fluorescence intensity (x-mean = 398) than the 102.IG construct (x-mean = 1124) (Fig. 5A). To investigate if

GFP expression correlates with virus production, the virus containing supernatant from the transiently transfected Phoenix E cells was used to infect 18-81 cells. Seventy-two hours after transduction, 18-81 cells were analyzed for GFP expression by analytical flow cytometry. On average, 0.1% of the 18-81 cells were transduced with the pLNCX2-eGFP construct and 10% of the 18-81 cells were transduced with the 102.IG construct (Fig. 5B). Low transfection and transduction rates of the pLNCX2-eGFP vector were explained by promoter interference between the LTR and the CMV promoter, in that both promoters might compete over shared transcription factors (compare discussion). Further the extended packaging signal of the CRU5 vector may have contributed to improve encapsidation of the retroviral mRNA.



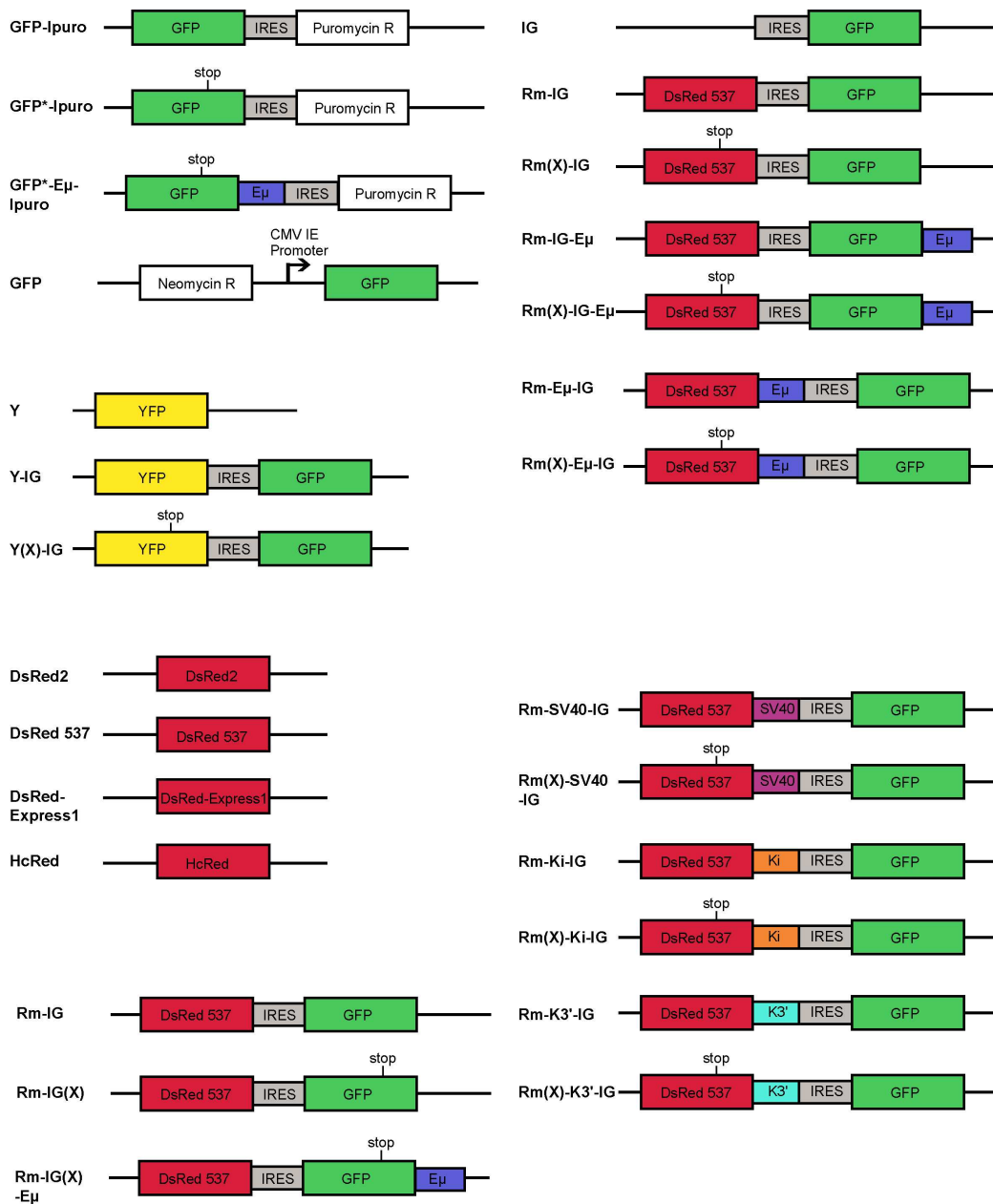
**Figure 5.** Flow cytometry analysis of GFP expressing cells. Fluorescent positive cells were gated, and transduction rates and relative x-means were determined. Transiently transfected Phoenix E cells (**A**) and transduced 18-81 cells (**B**) expressing the retroviral CRU5-IRES-GFP (IG) or pLNCX2-eGFP vector, respectively.



### 3.1.3. Retroviral backbone

Since the CRU5-IRES-eGFP vector produces two orders of magnitude more infectious virus particles as compared to the pLNCX2-eGFP vector, all mutation indicator plasmids described in this thesis and represented in Fig. 6 are based on the retroviral CBRU5-IRES vector backbone. Because retroviral vectors integrate randomly into the host cell genome via their LTR region [van der Putten *et al.*, 1981], positional effects of integration play less of a role. Furthermore, for infections I worked with a high cell to virus particle ratio, so that only one retroviral copy, or less, was integrated per cell. And, by working in large multi-clonal cultures, Luria-Delbrück fluctuation was avoided.

The basic mutation indicator system contained a fluorescence marker with a premature TAG (amber) stop codon. The stop codon is embedded in the RGYW [R = purine, Y = pyrimidine, W = adenine or thymidine] hot spot motif for hypermutation, in which typically the guanine is mutated [Rogozin and Kolchanov, 1992, Bachl *et al.*, 1997]. Upon reversion, a full-length fluorescent protein is expressed. Any single point mutation at the amber stop codon activates the reporter—except the transition from G to A, which only creates the ochre stop codon TAA. In the construct, the reporter gene is followed by an internal ribosomal entry site; and a second marker that allows selection of stably transduced cells. As reporter gene, the green and yellow fluorescence proteins (GFP and YFP, respectively) as well as various proteins with red fluorescence (DsRed) were tested. The second marker was either a drug resistance gene, or a second fluorescent protein. I also introduced various *cis*-acting enhancer elements into the reporter construct, to study the simultaneous activity of enhancers on transcription and hypermutation. Transduced cells were either drug selected, or FACS sorted by fluorescent protein expression.



**Figure 6.** Schematic of CRU5-IRES based indicator vectors. IRES = internal ribosome entry site, DsRed537 = mutated DsRed2 red fluorescence protein, GFP = enhanced green fluorescence protein, YFP = enhanced yellow fluorescence protein, Rm = DsRed537 red fluorescence protein, E $\mu$  = 1 kb immunoglobulin heavy chain intronic enhancer,  $\kappa$ i = 721 bp kappa light chain intronic enhancer fragment,  $\kappa$ 3' = kappa light chain 3' enhancer, SV40 = SV40 enhancer, CMV IE = CMV immediate early promoter, stop, \*, (X) = TAG amber stop codon embedded in RGYW mutational hot spot motif.

In some experiments, pre-existing fluorescent revertants were purged from the culture by sorting.

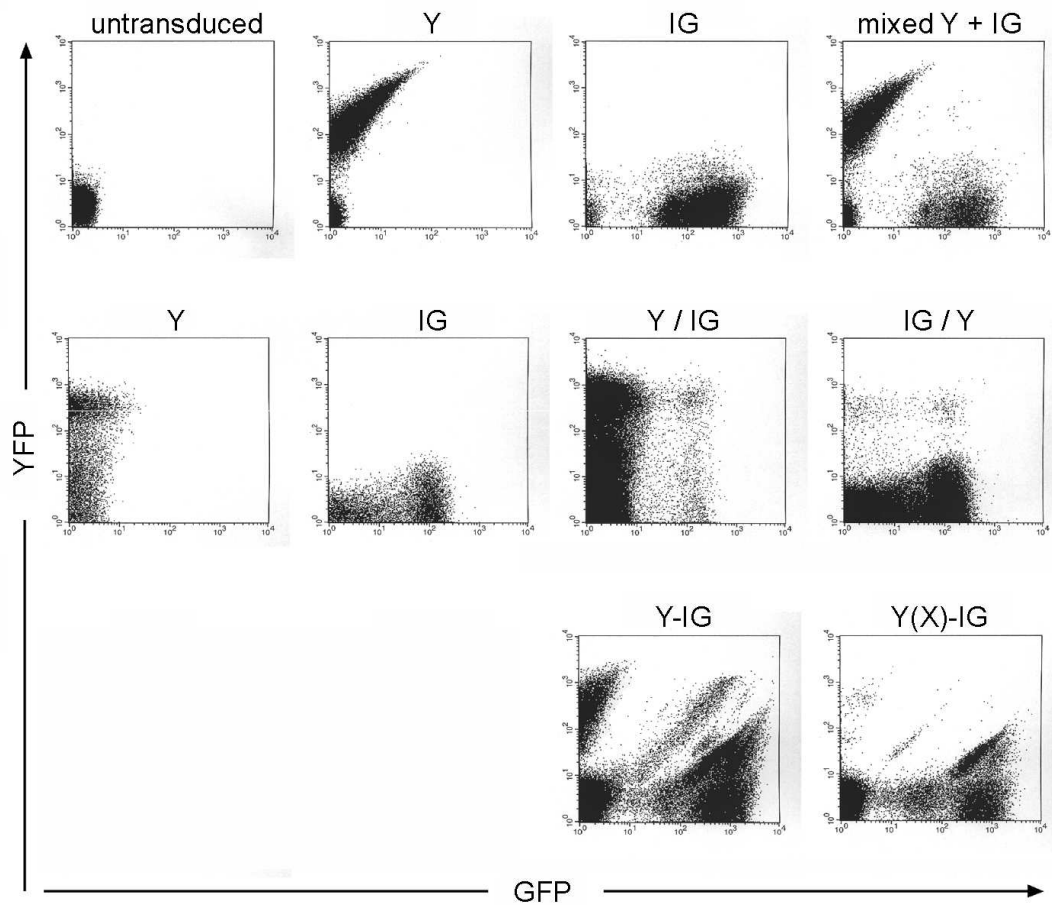
### **3.1.5. GFP family members**

In the past, we and others have used the GFP gene to create a point mutation indicator construct to study hypermutation [Bachl and Olsson, 1999; Wang *et al.*, 2004]. Because of the ease of use in flow cytometry and the familiarity with its characteristics, it seems natural to turn to any of the other GFP family members—yellow fluorescence protein (YFP), cyan fluorescent protein (CFP), and blue fluorescence protein (BFP)—for alternative point mutation indicators (Fig. 6). Indeed, YFP and GFP can be easily separated by flow cytometry, when the proteins are expressed in separate cells. Fig. 7, first row, shows the flow cytometry profiles of NIH/3T3 cells transduced with a vector containing YFP only (denoted Y); a vector containing an IRES followed by GFP (IG); and of a mixture of such cells. Clearly, there is no interference in flow cytometry of cells expressing either YFP or GFP. This also held true when NIH/3T3 cells were transduced with the Y vector first, and then with the IG vector (Y/IG cells), or with the IG vector first, and then with the Y vector. (Fig. 7, second row). The cells were not sorted before or after the second infection. It is curious though, that in the second round of infection much less cells were transduced. Because the virus did not produce envelope, superinfection ought not be inhibited. At any rate, some cells apparently were transduced by both constructs (third profile and fourth profile, upper right quadrant); some cells were not transduced by the Y vector (first), but by the IG vector in the second infection (third profile, lower right quadrant); or not transduced by the IG vector (first), but by the Y vector (fourth profile, upper left quadrant). These results seem to enable the simultaneous

use of YFP and GFP mutation indicators in the same cell. But when it later comes to sequencing of the mutation indicator, template switching during a PCR amplification and cloning step may complicate the identification and characterization of mutants.

In addition to the pitfall mentioned above, the extensive similarity of their DNA sequences precludes the combination of any GFP family genes (YFP, BFP, CFP, GFP) within the same retroviral vector backbone. Even though, such constructs may be useful to monitor transcription and mutability concomitantly. However, during reverse transcription of the viral RNA into cDNA, this similarity causes the retroviral reverse transcriptase to switch templates, resulting in recombinant molecules [Liu *et al.*, 2002; Svarovskaia *et al.*, 2000]. I confirmed this well documented phenomenon by combining the (enhanced) YFP gene with the GFP gene in the same vector. In the constructs, the YFP gene, which either contained a premature termination codon (denoted Y(X)-IG), or did not contain it (denoted Y-IG), was placed 5' to the internal ribosomal entry site (IRES), followed by the (enhanced) GFP gene in the standard retroviral backbone vector used throughout the work described here. These vectors were packaged into viral particles and introduced to the hypermutating 18-81 cells by infection. Furthermore, to exclude RAG-induced recombination and mutator induced reversion of the stop codon, NIH/3T3 fibroblasts, which are both RAG and mutator negative, were also infected. In flow cytometry, both transduced 18-81 cells (data not shown) and NIH/3T3 cells (Fig. 7, third row) showed multiple cell populations, some of which co-expressed YFP and GFP, while others expressed GFP or YFP only. In cells transduced with a Y-IG construct, two major YFP and GFP double-producer populations were present, sandwiching two minor double producer populations. Even in the Y(X)-IG construct, in which a premature stop

codon ought to prevent YFP from being translated, there are two major double producer populations, along with one minor double producer population (at the diagonal), and some YFP only cells. While I attribute the existence of the major double producers to template switching in a way that the recombination occurred before the premature stop in the YFP, I think that the minor populations represent stop codon revertants due to the error prone retroviral reverse transcriptase.



**Figure 7.** Flow cytometry analysis of YFP and GFP expressing NIH/3T3 cells. Cells were transduced with YFP (102.Y) or GFP (102.IG) containing vectors individually (first row), sequentially (second row, panel 3 and 4) or both fluorescence proteins were expressed by one plasmid (third row) (102.Y-IG). Y = eYFP, G = eGFP, I = IRES, Y(X) = eYFP with amber stop codon.

### **3.1.6. A silent point mutation in DsRed resulting in enhanced relative fluorescence intensity**

To avoid the two—quite different from each other—template switch phenomena, I looked at fluorescent proteins that are not derivatives of GFP. Furthermore, I wanted to generate an indicator that can be assayed in the red channel in the flow cytometer. The protein DsRed2 has only 23% DNA sequence similarity to GFP [Yarbrough *et al.*, 2001] and is a candidate for such an indicator. DsRed2 is an improved variant of DsRed, which was originally cloned from a coral of the *Discosoma* genus [Matz *et al.*, 1999]. DsRed2 was shown to be less toxic because of its increased solubility than the original DsRed and the protein matured faster, which reduces the time from transfection to detection [BD Biosciences Clontech, Palo Alto, CA]. Particularly, its emission maximum at 587 nm makes it easy to separate from GFP with an emission maximum at 508 nm. For expression in mammalian systems, the gene was codon optimized according to the human codon usage table for highly expressed genes [Haas *et al.*, 1996], to now have all serines encoded by the predominant UCC codon. In human cells, UCC recruits an abundant serine tRNA with a frequency of 17.8 times per 1000 codons [<http://www.kazusa.or.jp/codon/>]. The UCA codon, which also incorporates serine, is less frequent (12.3 times per 1000 codons). Furthermore, between human and mouse, the UCC to UCA tRNA ratios differ only slightly with 1.46:1 and 1.57:1, respectively.

In my cloning strategy I needed to destroy the *Bst*XI restriction site at nucleotide position 537 in the DsRed2 gene in order to integrate it into the *Bst*XI site of the 102.IG vector. To obtain this mutant, the pDsRed2-1 served as PCR template for site directed PCR mutagenesis in which I mutated the serine codon UCC at aa position 179 to a UCA codon

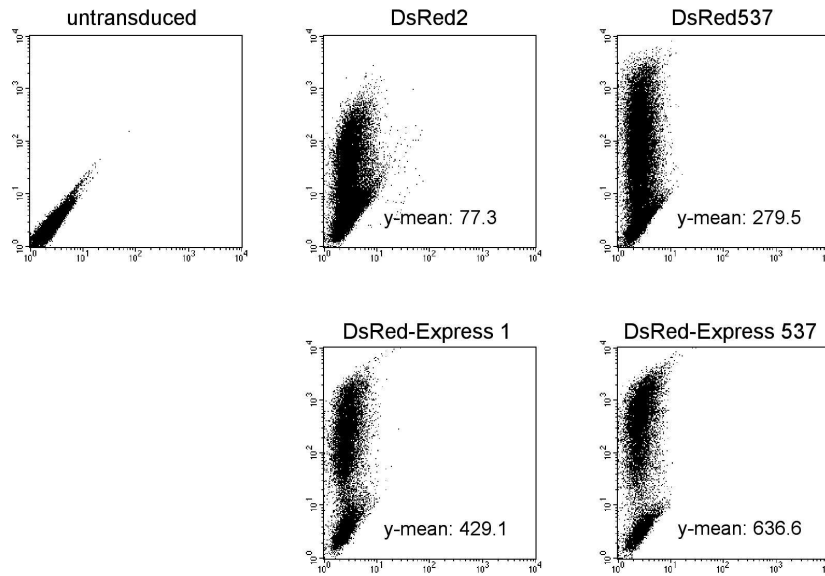
in the DsRed2 gene. Because the procedure replaced the cytidine by an adenosine at nucleotide position 537, the mutated gene was designated DsRed537. To compare mutated and original DsRed, both genes were cloned into the *XhoI* and *NotI* restriction sites of the 102.IG vector replacing the IRES-GFP fragment (Fig. 5). The CRU5 vector allows to either transiently express DsRed directly from the plasmid, driven by the CMV promoter, or by the retroviral MMLV promoter after stable integration of the provirus into the host cell genome. To produce retroviral particles, Phoenix E packaging cells were transiently transfected. For transduction, NIH/3T3 cells and the murine 18-81 lymphocyte line were spin-infected with the virus containing supernatant from transfected Phoenix E cells. The transduced cells were analyzed after 3 and 5 days by flow cytometry on a FACS Calibur<sup>®</sup>. Because the construct copy number per cell influences relative fluorescence intensity, samples with identical transduction rates were compared.

A direct correlation between expression rate and representation of rare versus common tRNAs has been extensively documented [Grosjean and Fiers, 1982; Holm, 1986]. According to these reports, if the switch of the UCC codon to the slightly less favorable UCA codon would directly translate into gene expression levels, one would predict a slight reduction of relative fluorescence intensity for the mutated DsRed protein. But surprisingly, the single base pair substitution resulted in an increase, rather than in a decrease, in relative fluorescence intensity (Fig. 8). The relative fluorescence intensity of DsRed537 was increased 3.6 fold in the NIH/3T3 line (Fig. 8A) and 4.1 fold in the 18-81 line (Fig. 8B). Even in DsRed-Express1, which is an improved variant of DsRed2 in that it shows a much higher fluorescence intensity from the beginning [Bevis and Glick,

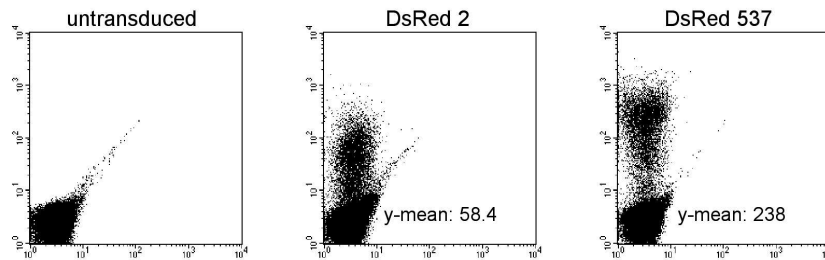
2002], the same single base pair substitution resulted in a 1.5 fold increase in NIH/3T3 cells (Fig 8A, second row). Because the single base pair substitution presumably has no effect on the transcription of the gene, and the amino acid sequence has not changed, the effect ought to be at the mRNA or translational level (compare discussion). Besides several other explanations, it is possible that tRNA depletion occurs during translation of over-expressed genes (compare discussion).



A



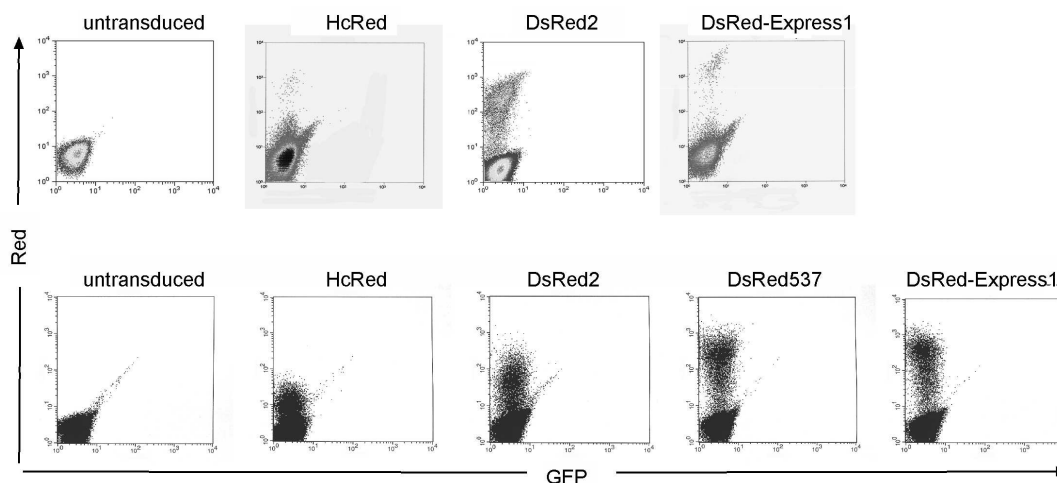
B



**Figure 8.** Flow cytometry analysis of DsRed expressing cells. Fluorescent positive cells were gated, and transduction rates and relative y-means were determined. NIH/3T3 cells (**A**) and 18-81 cells (**B**) transduced with the retroviral CRU5-DsRed2, CRU5-DsRed537, CRU5-DsRed-Express1 and CRU5-DsRed-Express537, respectively.

### **3.1.7. Alternative red fluorescence proteins**

In Fig. 10, I tested three different versions of DsRed—DsRed537, HcRed and DsRed-Express1 [CLONTECHniques, 2001]. The DsRed537 red fluorescent protein is a variant I derived from the DsRed2 protein; it has an increased relative fluorescent intensity [Klasen and Wabl, 2004]. HcRed had the advantage over the tetrameric DsRed537 of being a dimeric protein, which is likely to be more soluble and thus less toxic to the cell [Gurskaya *et al.*, 2001], but it has the disadvantage that its excitation maximum is at a wavelength of 588 nm. This requires a flow cytometer with a separate 568 nm laser line at 200 mW in combination with a 640 nm long pass filter in the FL2 channel. Even under these conditions, the relative fluorescent intensity of a 102.HcRed construct proved to be lower than the one of DsRed537 (Fig. 10). Finally, DsRed-Express1 is an improved variant of DsRed2 reported to be less toxic for the cell, better soluble, having a faster maturation rate, and an improved fluorescent intensity [Bevis and Glick, 2002]. Theoretically, a clear separation from the GFP signal was possible, and when the same flow cytometry conditions as those used for HcRed were applied, the results for DsRed-Express1 looked promising (Fig. 10). I have not yet tested DsRed-Express1 for possible toxicity, though. And quite recently, other fluorescent markers, such as the AsRed and the AmCyan protein (Clontech, San Jose, CA) have become available and might offer better solutions to the separation and toxicity issues.



**Figure 10.** Flow cytometry profiles of 18-81 cells transduced with the retroviral CRU5-HcRed, CRU5-DsRed2, CRU5-DsRed537 and CRU5-DsRed-Express1 vector (Fig. 6), respectively. Cells were analyzed on a MoFlow using a separate 568 nm laser line at 200 mW in combination with a 640 nm long pass filter in the FL2 channel (first row) and on a FACS Calibur using standard settings (second row).

## 3.2 Specific aim 2: Characterize the IgH major intronic enhancer and other enhancer elements, as *cis*-elements in hypermutation.

### 3.2.1. CRU5-Red537-IRES-GFP(X) constructs

In a first set of CRU5 constructs I introduced the amber stop codon into the GFP gene of the IG vector, creating IG(X). The DsRed537 (Rm) gene was placed 5' to the IRES segment, as a selection marker for stably transduced cells (Rm-IG(X)) (Fig. 6). To test the activity of the  $E\mu$  intronic enhancer in this construct, it was placed 5' to the MoMuLV 3'LTR region. Even though the  $E\mu$  enhancer was reported to be the only *cis*-acting element needed to target the mutator, placing it next to the 3'LTR enhancer could extinguish its function by enhancer interference [Bachl *et al.*, 1998]. Yet another

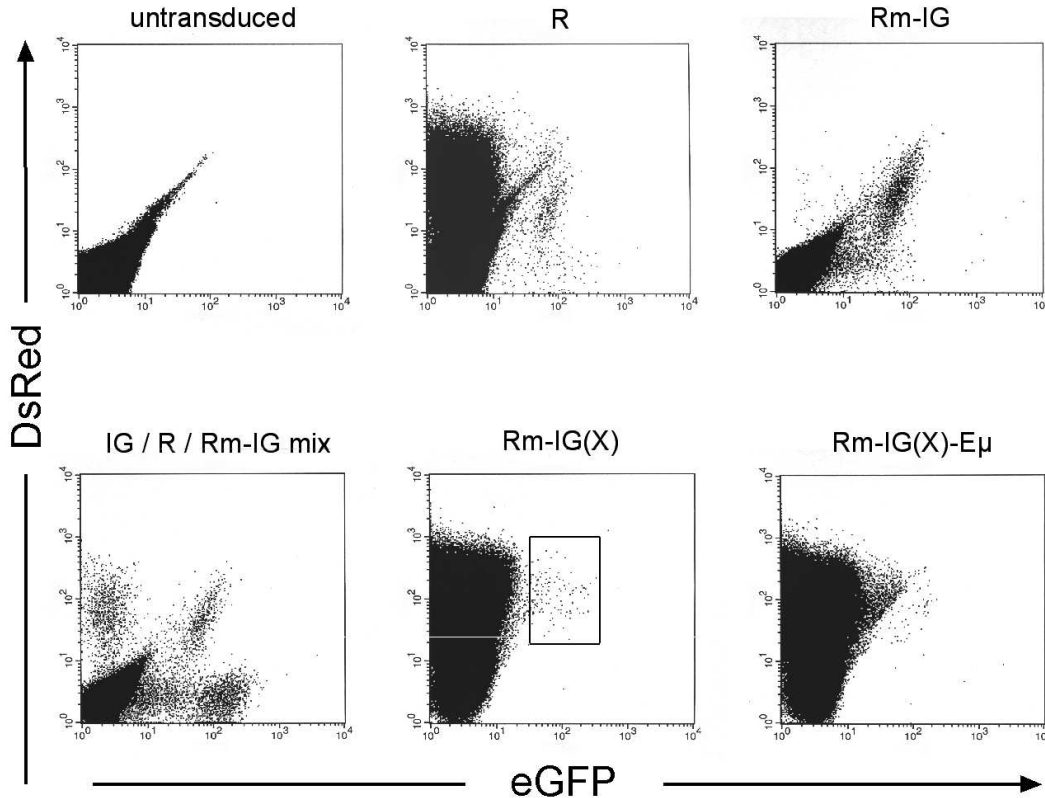
construct contained the Rm gene 5' to the IRES-GFP (Rm-IG) to express both DsRed and GFP (two color control) simultaneously (Fig. 6).

The constructs were used to transduce  $5 \times 10^5$  18-81 cells, which are active in hypermutation [Wabl *et al.*, 1985]. Measured by flow cytometry 3 days after transduction, on average, 20% of all transduced cells expressed DsRed. To measure mutation rates,  $5 \times 10^5$  GFP expressing cells were sorted by FACS, and pre-existing Rm(X) revertants were purged from the culture. All samples were set up in duplicates and sorted cells were seeded into 50 ml medium. Cultures were slowly expanded to a final volume of 400 ml over 14 days. Samples were analyzed every other day by flow cytometry.

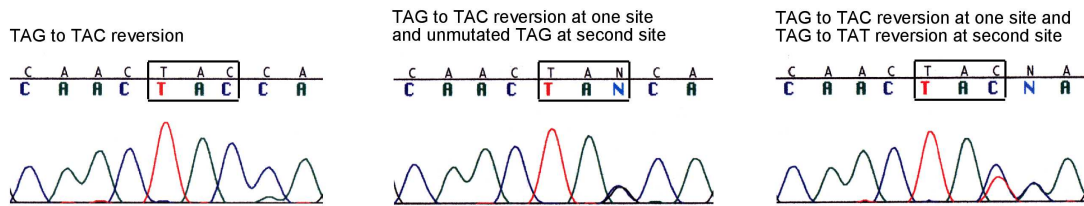
A clear identification of GFP(X) revertants by flow cytometry was not possible in this approach (Fig. 11), because the DsRed protein emits light in the green fluorescence spectrum during protein maturation (Fig. 11, second panel, population in the lower right quadrant) [Baird *et al.*, 2000]. This feature made it impossible to distinguish between maturing DsRed expressing cells and real GFP revertants. Further, the signal of auto-fluorescent dead cells interfered (Fig. 11, first panel). To investigate if real revertants were covered by interfering cells, (maturing DsRed cells or dead cells) a gate was set for potential DsRed537/GFP double positive cells in a Rm-IG(X) sample and single cells were sorted into 96 well plates by FACS (Fig. 11, second row, second panel). Only 17 clones out of 96 seeded single cells grew. A result, which I interpreted as circumstantial evidence, that dead cells have been sorted. The 17 subclones were expanded and genomic DNA was extracted from those cells. Subsequently, the GFP gene was PCR amplified and sequenced. Two of the 17 sequences did not show any homology to the GFP

sequence, which was probably due to unspecific PCR amplification. Six out of fifteen clones showed a clear reversion of the amber stop-codon to the anticipated TAC codon (G → C substitution)[Bachl and Wabl, 1996; Bachl *et al.*, 1997; Bachl and Olson, 1999]. The remaining seven sequences showed a double peak at the stop codon. Six sequences of those seven sequences showed a TAG peak under the TAC peak or vice versa (figure 12). If one excludes the possibility of a plasmid contamination during PCR or DNA sequencing and cell contamination during FACS, the double peaks can be interpreted as two integrated vectors per cell and only one integration site was mutated at the time of the DNA sequencing analysis. Certainly, this finding would have to be proven by i.e. Southern blotting. However, one of the seven clones showed also a double peak, but a TAT codon was detected under a TAC codon. This mutation was particularly interesting since the anticipated G → C substitution was detected for one integration site, only. At the other integration site, the mutator system generated a C → T transition (Fig. 12). Further, in one of the two sequences the TAT or TAC neighboring base pair was mutated. This was the only GFP sequence, which showed additional mutations than the reversion of the stop-codon.

However, since dead cells and DsRed maturing cells interfered with the GFP signal, mutation rate measurements based on a reliable read out by flow cytometry were impossible.



**Figure 11.** Flow cytometry analysis of 18-81 cells transduced with CRU5 indicator vectors. Live cells were gated based on their appearance in the forward / sideward scatter. R = DsRed2, Rm = DsRed537, IG = IRES-GFP, Eμ = 1 kb heavy chain large intronic enhancer, (X) = TAG stop codon. The gate in sample Rm-IG(X) was set for single cell sorting into 94 well plates.



**Figure 12.** Chromatograms of selected DNA sequences from three individual single cell subclones after FACS sorting for GFP(X) revertants. First panel: anticipated stop codon reversion from TAG to TAC (single proviral copy). Second panel: presumably two proviral copies, one TAG to TAC reversion, one TAG unreverted. Third panel: presumably two proviral copies, one reversion from TAG to TAC and one reversion from TAG to TAT. Additional mutation next to the stop codon for one integration site (C → A substitution).

### 3.2.2. CRU5-DsRed537(X)-IRES-GFP vectors

In the following experiments the DsRed537 gene contained the amber stop codon, which is embedded in the RGYW hot spot motif for somatic hypermutation, and the GFP gene was used as selection marker for transduced cells. Initially I tested the activity of the immunoglobulin  $E\mu$  intronic enhancer at two different positions in these reporter constructs. The  $E\mu$  enhancer was placed 3' to the Rm(X) gene to target the mutator complex to the red fluorescent reporter sequence (Fig. 6). In an additional construct, I placed the  $E\mu$  intronic enhancer 5' to the 3'LTR region, where the MoMuLV enhancer is located (Fig. 6). It was anticipated that the function of the  $E\mu$  would be extinguished by interference from the LTR enhancer [Bachl *et al.*, 1998]. Yet another construct, the Rm(X)-IG vector did not contain any immunoglobulin specific DNA elements (Fig. 6). Finally, because it did not contain any dsRed gene, the IG vector served as negative control (Fig. 6).

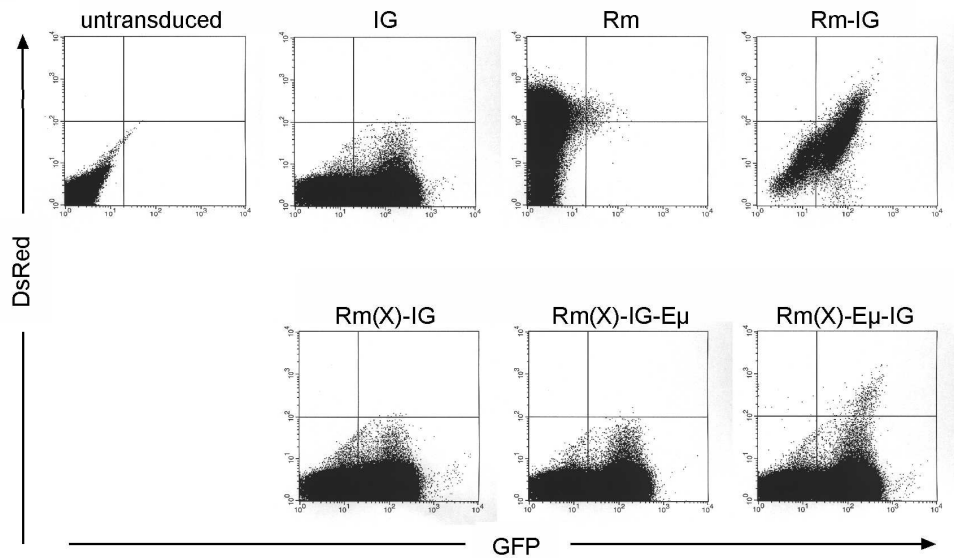
The constructs were used to transduce  $1 \times 10^6$  18-81 cells. Measured by flow cytometry 3 days after transduction, on average, 10% of all transduced cells expressed GFP. To measure mutation rates,  $1 \times 10^6$  GFP expressing cells were sorted by FACS, and pre-existing Rm(X) revertants were purged from the culture (Fig. 13B, day 0). All samples were set up in duplicates and sorted cells were seeded into 50 ml medium. Cultures were slowly expanded to a final volume of 400 ml over 7 days. A sample was removed every 24 hours and the number of DsRed and GFP double positive cells among 1 million GFP expressing cells was determined by flow cytometry (Fig. 13A). The first sample was analyzed on day 2 after the sort and showed an average of 110 revertants for the Rm(X)- $E\mu$ -IG construct. The number of revertants per million cells increased linearly over the

next 5 days, and a mutation rate of  $5 \times 10^{-5}$  mutation/bp/cell generation was calculated. In these experiments, the number of mutants for the Rm(X)-IG and the Rm(X)-IG-E $\mu$  vectors was under the detection limit. But I need to emphasize that due to the cut off line that was set to separate the red cells from green cells bleeding into the red gate (Fig 14, y-axis), approximately 2/3 of all mutants were not recorded (Fig. 13A, panel 4).

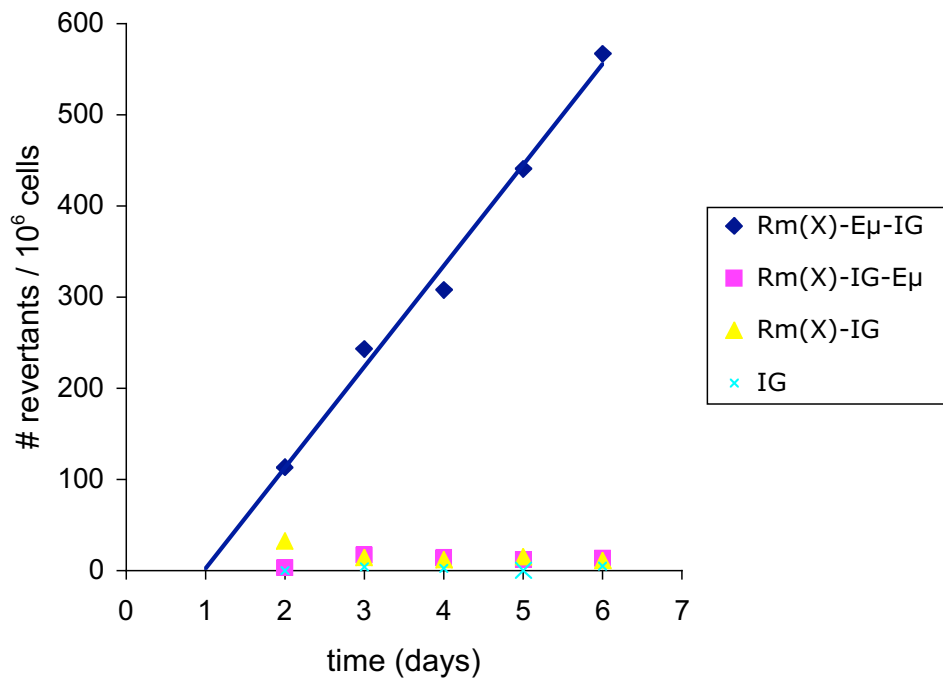
To verify that the revertants detected in the Rm(X)-E $\mu$ -IG samples were due to true reversion of the amber stop codon, GFP/DsRed double positive cells were sorted by FACS, and the chromosomal DNA from these cells was extracted to be used as a PCR template, subcloned into bacterial vectors and sequenced. Of 23 sequences, 22 had the TAG stop codon mutated to TAC, and one to TAT. Additional mutations were found in 4 of the 23 sequences (Fig 14). This pattern has been observed repeatedly in the 18-81 cell line, both at the endogenous IgH locus and at ectopic genes [Bachl *et al.*, 1997; Bachl and Wabl, 1996]. All of the additional mutations were C  $\rightarrow$  T transitions, three of them on the non-template strand and one on the template strand. Interestingly, only 2 out of the 4 additional mutations occurred within the RGYW hot spot motif.



A



B



**Figure 13. (A)** Flow cytometry analysis of Rm(X) revertants in the hypermutating 18-81 line. IG = IRES-GFP, Rm = DsRed537, (X) = amber stop codon, Eμ = 1 kb heavy chain intronic enhancer. **(B)** Somatic hypermutation of the Rm(X) indicator in the presence and absence of the Eμ enhancer in the 18-81 line. x-axis = time in days, y-axis = number of Rm(X) revertants per  $1 \times 10^6$  GFP expressing cells.

dsRed (X)	GTGTACGTGA	AGCACCCCGC	CGACATCCCC	AGGACGGCTG	CTTCATCTAC
clone 7	GTGTACGTGA	AGCACCCCGC	CGACATCCCC	AGGACGGCTG	<span style="border: 1px solid black;">TT</span> TCATCTAC
clone 17	GTGTACGTGA	AGCACCC <span style="border: 1px solid black;">T</span> GC	CGACATCCCC	AGGACGGCTG	CTTCATCTAC
clone 20	GTGTACGTGA	AGCACCCCGC	CGACATCCCC	AGGACGGCTG	CTTCATCTAC
clone 34	GTGTACGTGA	AGCACCCCGC	CGACATCCCC	AGGACGGCTG	CTTCATCTAC
clone 43	GTGTACGTGA	AGCACCCCGC	CGACATCCCC	AGGACGGCTG	CTTCATCTAC
residue (bp)	211		240	341	360

dsRed (X)	CGAGCGCCTG	TACCCCGCG	GAAGGACGGC	GGCCACTAGC	GCAGTGCCC	GGCTACTACT
clone 7	CGAGCGCCTG	TACCCCGCG	GAAGGACGGC	GGCCACTA <span style="border: 1px solid black;">CC</span>	GCAGTGCCC	GGCTACTACT
clone 17	CGAGCGCCTG	TACCCCGCG	GAAGGACGGC	GGCCACTA <span style="border: 1px solid black;">CC</span>	GCAGTGCCC	GGCTACTACT
clone 20	CGAGCGCCTG	TA <span style="border: 1px solid black;">T</span> CCCGCG	GAAGGACGGC	GGCCACTA <span style="border: 1px solid black;">CC</span>	GCAGTGCCC	GGCTACTACT
clone 34	CGAGCGCCTG	TACCCCGCG	GAAGGACGGC	GGCCACTA <span style="border: 1px solid black;">CC</span>	<span style="border: 1px solid black;">T</span> GCAGTGCCC	GGCTACTACT
clone 43	CGAGCGCCTG	TACCCCGCG	GAAGGACGGC	GGCCACTA <span style="border: 1px solid black;">CC</span>	GCAGTGCCC	GGCTACTACT
residue (bp)	441	460	501			540

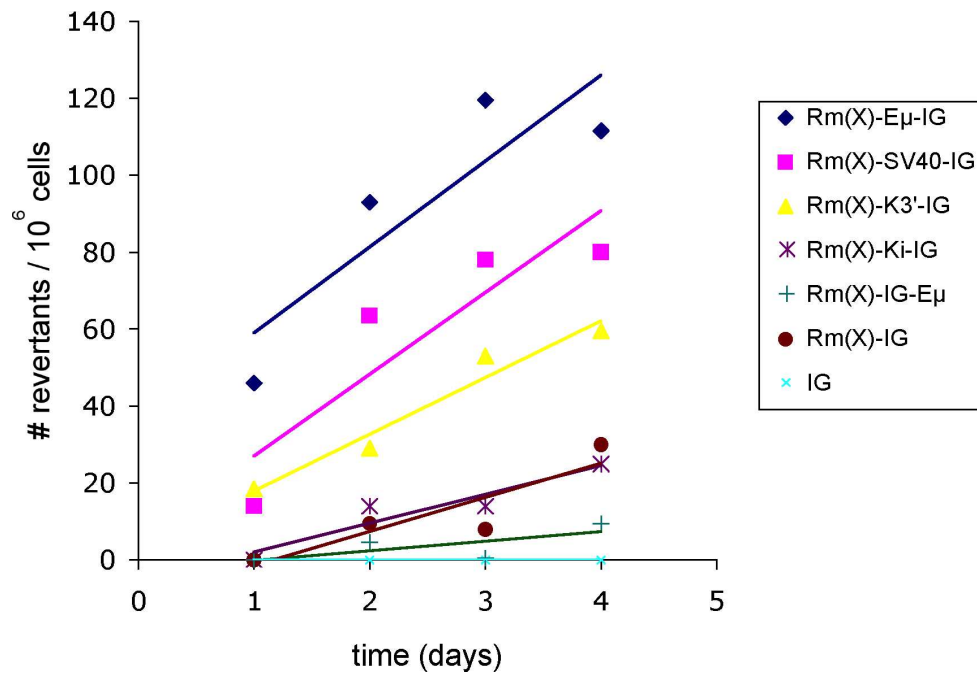
**Figure 14.** Selected sequences from Rm(X) revertants, which show additional mutations. Mutations are boxed. Position of mutations are indicated by the bp residue. DsRed (X) = DsRed537 gene with integrated stop codon at bp 517. Clone # = *E. coli* plasmid number.

### 3.2.3. *Cis*-acting elements

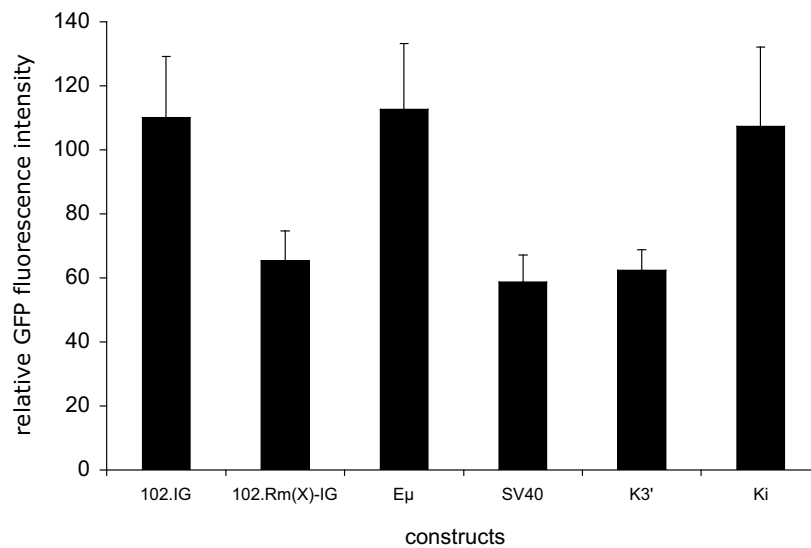
In the next set of experiments I included various other enhancers into the reporter constructs. The  $E\mu$  enhancer of the Rm(X)- $E\mu$ -IG construct was replaced by the SV40 enhancer (to generate the Rm(X)-SV40-IG construct); by the 803 bp kappa 3' enhancer (Rm(X)- $\kappa$ 3'-IG); and a 720 bp fragment of the kappa intronic enhancer, which lacked the 3' matrix attachment region (MAR) to target somatic hypermutation [Klix *et al.*, 1998] (construct Rm(X)- $\kappa$ i-IG). GFP expressing cells, transduced with one of these constructs each, were sorted and seeded in 100 ml medium, and the culture was expanded for 1 day to a final volume of 200 ml. From then on the cultures were split in a 1:3 ratio every 24 hours. Samples were removed and analyzed by flow cytometry (Fig. 15). As expected, the Rm(X)- $E\mu$ -IG samples revealed the highest number of revertants (Fig. 15A). The construct did also show a high transcription rate estimated by the relative x-mean (Fig.

15B,  $\bar{x}$ -mean = 102). But the Rm(X)-SV40-IG samples showed almost as high numbers of revertants as well, despite a lower  $\bar{x}$ -mean (Fig. 15B,  $\bar{x}$ -mean = 59). Inclusion of the SV40 enhancer also gave the construct a higher mutability than the 3'  $\square$  enhancer. And the  $\square$  intronic enhancer had the smallest effect, but the construct was expressed best ( $\bar{x}$ -mean = 107). But even with no enhancer in close proximity to the Rm(X) gene revertants were recorded (Fig 15A, sample Rm(X)-IG).

A



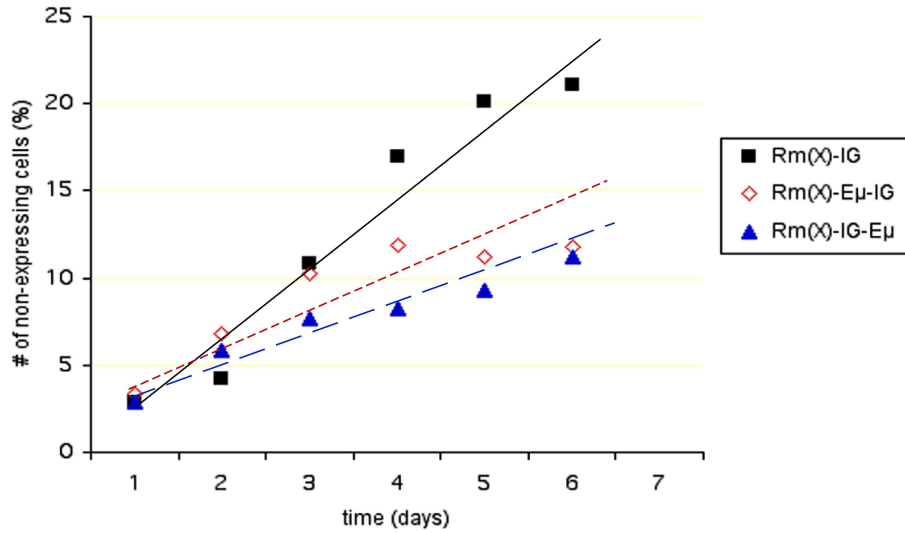
**B**



**Figure 15. (A)** Time course of Rm(X) revertants under influence of various *cis*-acting enhancer elements. Y-axis = number of Rm(X) revertants per  $1 \times 10^6$  GFP expressing cells. X-axis = time in days, IG = IRES-GFP, Rm = DsRed537, E $\mu$  = 1 kb heavy chain intronic enhancer, SV40 = SV40 enhancer,  $\kappa$ i = 721 bp fragment of kappa light chain intronic enhancer,  $\kappa$ 3' = kappa light chain large 3' enhancer, (X) = amber stop codon **(B)** Relative GFP expression rates of reporter constructs containing various enhancers. Y-axis = relative GFP fluorescence intensity. X-axis = construct, E $\mu$  = Rm(X)-E $\mu$ -IG, SV = Rm(X)-SV40-IG, Ki = Rm(X)-Ki-IG, K3' = Rm(X)-K3'-IG constructs

These experiments show that using the retroviral CRU5 reporter constructs to measure mutation rates is reasonably fast and reliable. Yet, the engineered vectors leave some room for potential improvement (compare discussion). Even though the increased relative fluorescent intensity of the DsRed537 variant made a separation from the GFP signal in a standard flow cytometry setting possible, 2/3 of the revertants were still covered by a “hump” of the GFP signal alone. Sequencing the GFP gene from cells in the hump did not show any mutations (data not shown). Further, the number of cells, which do not express the indicator vector (i.e. get silenced) increased over the course of the experiment (Fig. 16). These observations led to switch to indicator vectors which contain the GFP(X)

gene together with an antibiotic resistance marker for selection of transduced cell on a bicistronic message.



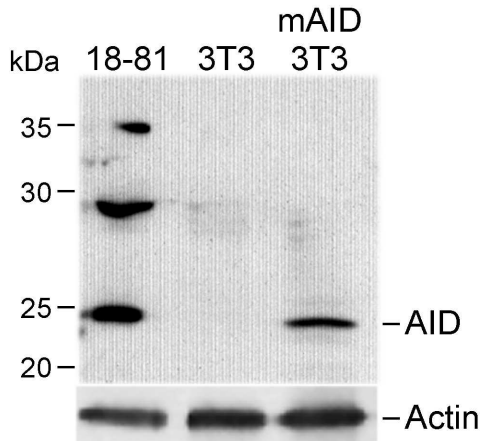
**Figure 16.** Silencing of retroviral reporter constructs over time based on Fig 13. Cell numbers were estimated by setting a gate for non-expressing cells in the flow cytometry analysis. Rm(X) = DsRed537 containing the integrated amber stop codon, Eμ = 1 kb heavy chain intronic enhancer, IG = IRES-GFP, x-axis: time in days, y-axis: number of silenced cells in percentage

### **3.3. Specific Aim 3: Somatic hypermutation and mismatch repair in non-B cells**

With the following experiments, I wanted to answer the question whether mismatch repair plays a role in hypermutation in cells other than B lymphocytes. In one sense, this question is tautological; clearly, if all introduced mutations were repaired, no mutations would be seen. But, somewhat counter-intuitively at first, we and others found that in B lymphocytes the mismatch repair system may be modified in such a way that it actually introduces mutations rather than corrects them. Previously, subclones of the human 293T kidney cell line had been constructed in such a way that the expression of MLH1 can be controlled by tetracycline [Trojan *et al.*, 2002]. This gave us an opportunity to study the influence of MMR in AID mediated hypermutation.

#### **3.3.1. AID expression in NIH/3T3 cells**

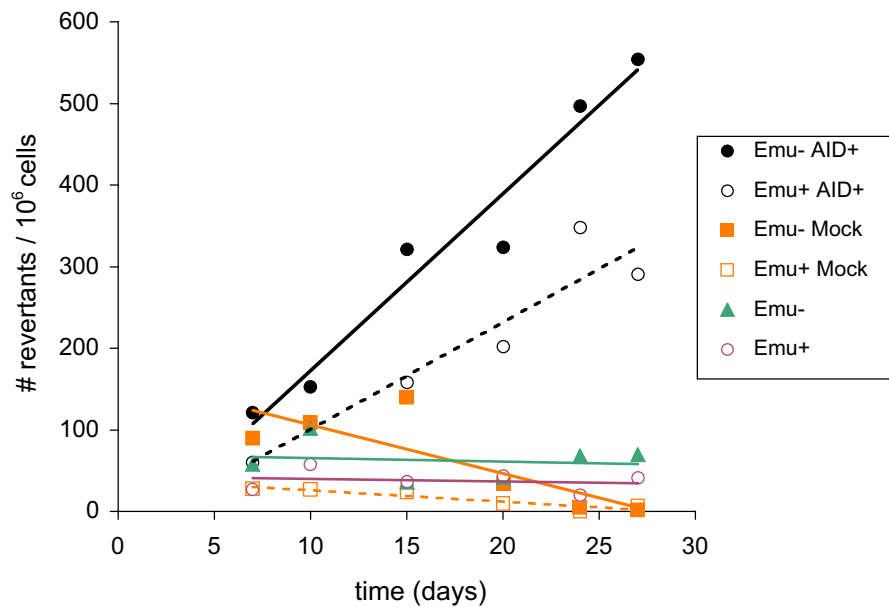
I first wanted to make sure that the AID I cloned from the 18-81 cell line mediates hypermutation, when transduced into non-AID expressing cell lines. I used the pLNCX2-mAID vector for ectopic AID expression in NIH/3T3 cells. AID expression was tested by Western blot (Fig. 17). While the non-transduced NIH/3T3 cells contain no band representing AID (Fig. 17, lane 2), the transduced cells show a strong band at just below 25 kDa representing AID. However, compared to the endogenous AID expression of 18-81 cells (Fig. 17, lane 1), there was less AID protein expressed in the transduced NIH/3T3 cells.



**Figure 17.** Western blot analysis of AID expression in 18-81 cells, and in NIH/3T3 cells transduced with the pLNCX2-mAID retroviral expression vector. Upper portion of the blot, developed with antibody to AID; lower portion, developed with antibody to actin. Lane 1, 18-81 line; lane 2, NIH/3T3 line; lane 3, NIH/3T3 line transduced with AID. To the left of the blot, molecular weight markers; to the right, position of AID and actin, respectively.

To investigate whether the cloned AID would induce somatic hypermutation in non-AID expressing cell lines, I used the reporter constructs pGFP\*-Ipuro (GFP\*) and pGFP\*-Emu-*Ipuro* (GFP\*E) [Wang *et al.*, 2004] (Fig. 5). These reporters contain an enhanced green fluorescent protein gene (GFP) with a premature termination codon (\*), embedded in a RGYW hot spot motif for somatic hypermutation [Rogozin and Kolchanov, 1992; Betz *et al.*, 1993; Bachl *et al.*, 1997; Bachl and Olson, 1999]. Upon reversion of the stop codon, full-length GFP protein is expressed and the cell fluoresces, which is detected by flow cytometry. The reporter constructs encode the puromycin resistance gene on one bicistronic message together with the GFP gene (Fig. 5). Transduced cells were selected with 2.5  $\mu\text{g/ml}$  puromycin. I introduced this reporter into duplicate NIH/3T3 cell cultures, and subsequently, the cells were transduced with the pLNCX2-mAID retrovirus. The empty pLNCX2 vector served as mock control. Stably transduced cells were selected with 100 ng/ml neomycin and selection pressure on both the AID expression vector and the reporter was maintained during the entire experiment. Samples were taken and GFP\* reversion was analyzed by flow cytometry at multiple time points (Fig. 18). In these experiments, I actually used two kind of reporters—one contained the GFP\* as the

essential gene, and another had in addition the large IgH intronic enhancer ( $E\mu$ ). This Ig locus element was described to increase mutation rates for ectopic Ig genes [Bachl *et al.*, 1998], but not the ectopic GFP reporters used here [Wang *et al.*, 2004; Spillmann and Wabl, 2004; Klasen *et al.*, 2004]. In the experiments described here, it did not make much difference whether or not the reporter contained  $E\mu$ . In fact, at  $2 \times 10^{-5}$  per day, the reporter without  $E\mu$  (Fig. 18, open circles) even mutated with a somewhat higher rate than the one with  $E\mu$ , at  $1.2 \times 10^{-5}$  per day (Fig. 18, filled circles).



**Figure 18.** AID mediated somatic hypermutation in NIH/3T3 fibroblasts. Y-axis: number of GFP\* revertants per  $1 \times 10^6$  live cells. X-axis: time in days. Open circles, AID-transduced cells containing GFP\* reporter without  $E\mu$  enhancer (Emu-); filled circles, AID-transduced cells containing GFP\* reporter with  $E\mu$  enhancer (Emu+). Open squares, mock (with pLNCX2) transduced cells containing GFP\* reporter without  $E\mu$  enhancer (Emu-); filled circles, mock transduced cells containing GFP\* reporter with  $E\mu$  enhancer (Emu+). Open and filled triangles, cells containing GFP\* reporter only, without (Emu-), or with  $E\mu$  enhancer (Emu+), respectively.



Importantly, the mutation frequencies increase for both constructs with time, while there is no such increase for cells without AID (Fig. 18, empty and filled squares represent mock-transfected cells; and empty and filled triangles represent cells with reporter only). Taking into account that the fibroblasts divide once in 24 hr, a mutation rate of  $1.2 \times 10^{-5}$  and  $2 \times 10^{-5}$  mutations/bp/ cell generation was calculated. This rate is comparable to the rate of GFP\* in the 18-81 cell line ( $1.0 \times 10^{-5}$ ) [Bachl and Wabl, 1996; Green *et al.*, 1995], which expresses AID endogenously, but ten times lower than the one at the endogenous  $V_H$ , with  $0.3 - 1 \times 10^{-4}$  mutations/bp/cell generation, in the 18-81 cell line [Wabl *et al.*, 1985].

### **3.3.3. Hypermutation in the presence and absence of mismatch repair**

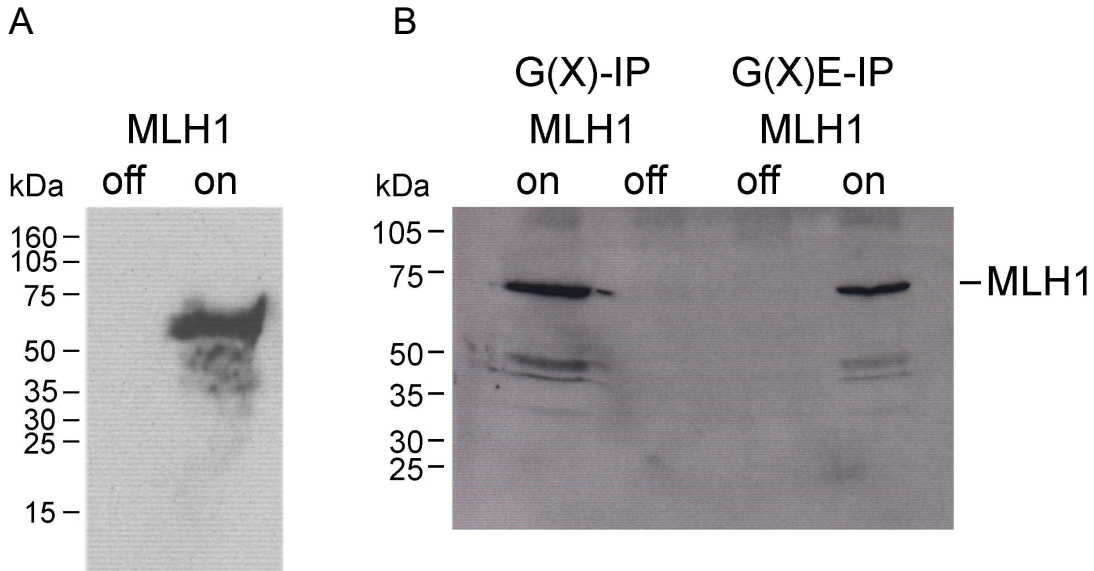
Having successfully tested the activity of our ectopic AID construct in NIH/3T3 cells, I turned to study the influence of the mismatch repair system on somatic hypermutation in cells other than activated B cells. I took advantage of the 293T-Tet-Off-hMLH1 cell line—a kidney cell line, in which the expression of MLH1 can be controlled by tetracycline or its analogs; in this cell line, a transgenic MLH1 gene under the control of the Tet-off inducible system has replaced endogenous MLH1 gene expression [Trojan *et al.*, 2002]. MLH1 repression was tested by Western blot by adding  $50 \mu\text{g/ml}$  doxycyclin to the culture medium for 7 days (Fig. 19A, left lane). To detect small amounts of MLH1, the SDS-gel was overloaded with  $400 \mu\text{g}$  whole cell lysate. Under these conditions, no MLH1 protein could be detected, although in the absence of doxycyclin MLH1 was well expressed (Fig. 19A, right lane).

Next the modified 293T cell line was transduced with mutation reporter constructs and selected for stable integration. Then the culture was split into two parts, and for five days,

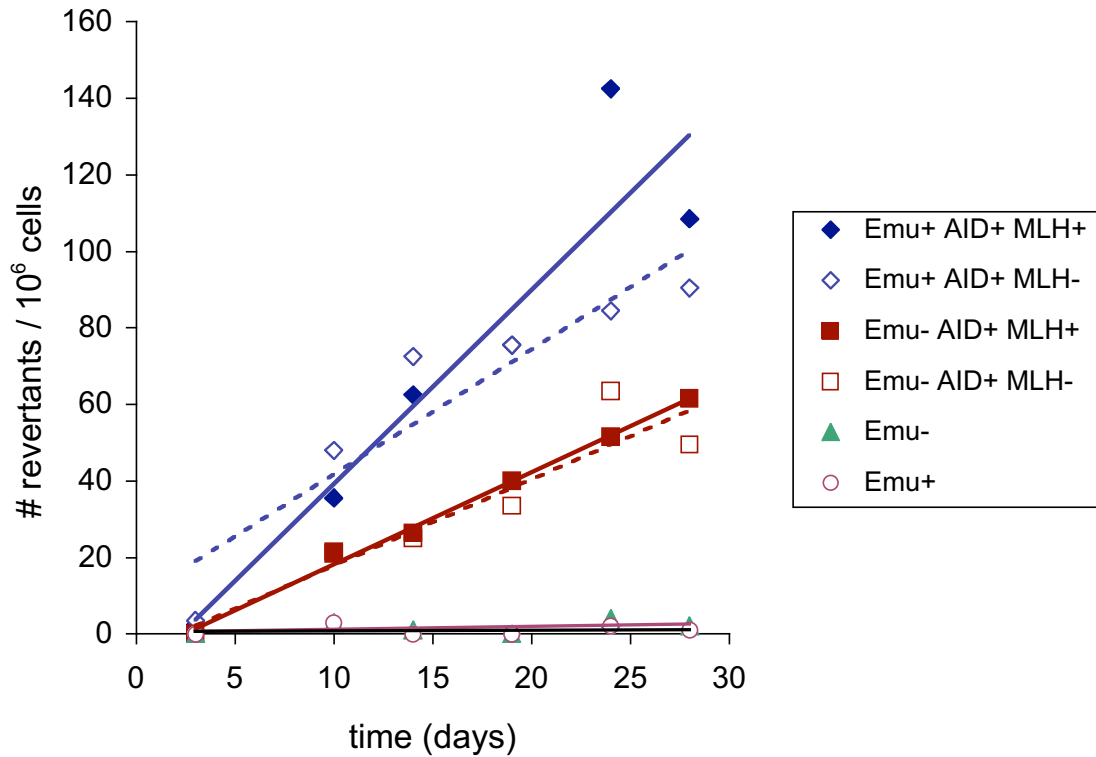
the one aliquot was grown in medium containing doxycycline to repress MLH1 expression, whereas the medium of the other aliquot contained no doxycycline. Both of these cell cultures—one expressing MLH1 and the other not—were transduced with the retroviral construct that contains AID (pLNCX2-mAID). All transduced cells were kept under selection with 0.5  $\mu\text{g/ml}$  puromycin, for the indicator plasmid, and with 500 ng/ml neomycin, for the AID expression vector, during the entire experiment. All cultures were also set up in duplicates, and samples were taken for the determination of mutation frequencies by flow cytometry over a period of 28 days. On day 30 of the experiment, the performance of the Tet-off system was tested by Western blot analysis again; i.e., was the MLH1 expressed in the appropriate cell lines? Fig. 19B shows that this indeed was the case. No MLH1 protein was detectable in doxycyclin treated cultures (Fig. 19B, lanes 2 and 3), but MLH1 was well expressed in cultures without doxycyclin (Fig. 19B, lanes 1 and 4).

In these experiments, parallel culture sets were tested with reporter constructs with (GFP\*E) and without (GFP\*) the large intronic enhancer  $E\mu$ . I took flow cytometry measurements of GFP revertant frequencies in the various cultures over a period of 28 days. Unlike in the experiments with the NIH/3T3 line, where addition of the  $E\mu$  had a slight negative effect on the mutability of the reporter gene, in the 293T line, the  $E\mu$  increased the mutation frequency (Fig. 20, filled and empty diamonds;  $4 \times 10^{-6}/\text{day}$ ) approximately two-fold over the one in the construct without  $E\mu$  (rectangles;  $2.1 \times 10^{-6}/\text{day}$ ). However, MLH1 expression did not have any influence on the mutation rate in this system. Thus, whether (Fig. 20, filled diamonds and rectangles) or not (Fig. 20, open

diamonds and rectangles) the cells expressed MLH1, there was no difference in mutation frequencies.



**Figure 19.** Western blot analysis of MLH 1 expression in human 293T-Tet-Off-hMLH1 cells, developed with antibody to MLH1. **(A)**, before the experiment 18-81 cells; and **(B)**, at day 30 of the experiment. To the left of the blots, molecular weight markers; to the right of blot **(B)**, position of MLH1. Same protein amounts were applied in all lanes. Above the individual lanes: “off”, 50 µg/ml doxycyclin; “on”, no doxycyclin. GFP\*, GFP reporter without Eµ enhancer; and GFP\*E, reporter with Eµ enhancer.



**Figure 20.** AID mediated somatic hypermutation in the presence and absence of MLH1. Y-axis: number of GFP\* revertants per  $1 \times 10^6$  live cells. X-axis: time in days. Diamonds, cells containing GFP\* reporter with E $\mu$  enhancer (Emu+); AID-transduced (AID+); and MLH expressing (MLH+; filled diamonds), or MLH-non-expressing (MLH-; open diamonds). Squares, cells containing GFP\* reporter without E $\mu$  enhancer (Emu-); AID-transduced (AID+); and MLH expressing (MLH+; filled squares), or MLH-non-expressing (MLH-; open squares). Open and filled triangles, cells containing GFP\* reporter only, without (Emu-), or with E $\mu$  enhancer (Emu+), respectively.

## **4. Discussion and conclusions**

### **4.1. Specific Aim 1: Develop a retroviral based point mutation indicator system.**

#### **4.1.1. High virus titer production**

The results of the comparison between the pLNCX2 vectors and the CRU5 vectors were not surprising since the CRU5 vector was optimized to gain high virus titers in the Phoenix cell line [Lorens *et al.*, 2000; Swift *et al.*, 1999; Pear *et al.*, 1993; Danos and Mulligan, 1988]. On the other hand it was not anticipated that the pLNCX2-eGFP plasmid would produce that little amounts of virus particles. It can be assumed that promoter/promoter and promoter/enhancer interference contributed to the poor pLNCX2-eGFP virus titers, in that, the MMLV and the CMV promoter would compete over common transcription factors [Lorens *et al.*, 2000]. This would explain the comparable low relative fluorescence GFP intensity of the pLNCX2-eGFP construct in the Phoenix cell line and the low infection rate of 18-81 cells with the pLNCX2-eGFP virus. Because less retroviral transcripts were made in the first place, which ultimately results in low virus titers. Additionally, transcription initiation at the CMV promoter could interfere with the elongation of the full size retroviral transcript initiated by the MMLV promoter of the pLNCX2-eGFP construct during virus production in the Phoenix line. Another source for interference could be the short GFP mRNA product generated under the CMV promoter, which could hybridize with the template DNA strand. In this respect a bicistronic strategy as developed for the CRU5 vector, in which only one transcript is made, has major advantages during virus particle production in the Phoenix cell line. Certainly, the extended  $\Psi$  packaging signal of the CRU5 vector does improve the

encapsidation of the viral mRNA in the Phoenix packaging cell line further, which also contributes to high virus titers [Lorens *et al.*, 2000]. These results clearly favor the CRU5 bicistronic retroviral vector over the pLNCX2 vector.

#### **4.1.2. CRU5 based retroviral reporter constructs**

Using the retroviral CRU5 reporter constructs we could show that measuring mutation rates is reasonably fast and reliable. The large number of integration sites mitigates a position effect that invariably comes into play in cultures of cells with a single integration site, as is the case with stable transfectants. Transducing large numbers of cells circumvents time consuming subcloning and compartmentalization tests, which are necessary when low numbers cause fluctuation. However, there certainly is room for potential improvement on this type of vectors. For example, silencing of the provirus by methylation of the LTR promoter/enhancer region is an extensively reported host defense mechanism against retroviruses [Svoboda *et al.*, 2000; Robbins *et al.*, 1998]. This has no influence on constructs that can be kept under drug selection pressure, but it has been an issue for the construct without such selection.

An improvement might be achieved with so-called "sin-vectors", in which the LTR promoter/enhancer is deleted in the provirus. In these vectors, silencing via methylation is thought to be reduced, and, in some cases, not even detectable [Lorens *et al.*, 2000; Pfeifer *et al.*, 2002]. Expression of the mutation indicator cassette would then be triggered by a promoter located immediately 5' of it [Lorens *et al.*, 2000]. This approach would have two other advantages. First, it would be possible to reduce the distance between promoter and indicator sequence. In vivo, the mutation rate at the V(D)J and VJ region is highest in close proximity to the promoter (with an approx. 200 bp spacer

between promoter and target sequence) and decreases with increasing distance from the promoter [Lebecque and Gearhart, 1990; Rada *et al.*, 1997]. In our CRU5 constructs, the MoMuLV promoter is separated from the indicator gene by the 1283 base pair  $\Psi$  packaging signal, which may explain the lower mutation rates measured with this indicator plasmids, as compared to other measurements in the 18-81 cell line [Bachl and Wabl, 1996; Bachl *et al.*, 1999; Bachl and Olsson, 1999; Bachl *et al.*, 2001]. Second, an enhancer-less promoter, such as the thymidine kinase promoter, would allow a better controlled study of putative mutation enhancers.

On the other hand, sin-vectors would encode two promoters in the plasmid stage (the 3'LTR promoter and the internal promoter) when expressed in the Phoenix packaging line. This can result in low virus titers due to promoter/promoter interference, transcription elongation, and short mRNA interference as discussed for the pLNCX2 vector. In this respect a bicistronic strategy as developed for the CRU5 vector, in which only one transcript is made, has major advantages during virus particle production in the Phoenix cell line.

Alternatively, lentiviruses may be beneficial since they yield higher transduction rates, and they can infect non-dividing cells, which may help in studies on the influence of cell division on hypermutation [Logan *et al.*, 2002]. To completely exclude positional effects, episomal adenoviral vectors might be applicable, because they do not integrate into the host genome [Ehrhardt *et al.*, 2003]. Today these vectors still do have the disadvantage of introducing high plasmid copy numbers per transfected cell (1 - 100 copies) and low transfection rates [Colosimo *et al.*, 2000].

#### **4.1.3. Fluorescence marker genes versus antibiotic selection markers**

These and other studies showed that drug selection seems to be the most straightforward way to measure relative mutation rates on stably expressed indicator genes [Wang *et al.*, 2004; Spillmann and Wabl, 2004; Klasen *et al.*, 2004]. However, because transcription rates correlate directly with hypermutation rates, for a limited antibiotic dosage range high mutation rates would change with the degree of antibiotic selection pressure. Furthermore, increased antibiotic selection pressure may increase the cytotoxicity of the reporter gene, as is the case for the DsRed tetramer.

Further, certain proviral integration sites in the host genome could support expression of the selection cassette, which would ultimately lead into a growth advantage of those cells. After a few cell generations one would measure mutation rates of some predominant clones, which support proviral expression. By using fluorescence marker genes this selection criteria were circumvented. Unless, expression of a fluorescence marker would increase cell proliferation, but this is a purely theoretical scenario which has not been reported so far.

#### **4.1.4. DsRed**

Used alone, DsRed is a viable alternative to the GFP. However, in combination with (all cells also expressing), the GFP signal bleeding into the DsRed channel during flow analysis can pose a problem. Even though the increased relative fluorescent intensity of the DsRed537 variant made a separation from the GFP signal in a standard flow cytometry setting possible, 2/3 of the revertants were still covered by a “hump” of the GFP signal alone. Sequencing the GFP gene from cells in the hump did not show any mutations (data not shown). I hypothesize that the hump might be due to dying cells,



even though the cells would still be recorded as live cells in the forward/sideward scatter of the flow cytometer. Thus, the increased relative fluorescent intensity of the improved DsRed537 was still not optimal. As potential alternatives, HcRed and DsRed-Express1 have become commercially available [CLONTECHniques, 2001]. The results for DsRed-Express1 looked promising since a clear separation from the GFP signal was theoretically possible. However, DsRed toxicity combined with silencing of the retroviral reporter may be problematic in long-term experiments.

#### **4.1.5. A silent point mutation in DsRed resulting in enhanced relative fluorescence intensity**

Since the single base pair substitution in DsRed537 has not changed the amino acid sequence and presumably has no effect on the transcription of the gene, the effect ought to be at the mRNA or translational level.

Initiation of mRNA translation is the same for both transcripts, but the elongation at the mutated site might be affected. The point mutation could have altered tertiary mRNA structure and, therefore, have improved processivity of the ribosomes. However, the mRNA prediction software *mfold* [<http://www.bioinfo.rpi.edu/applications/mfold/old/rna/>] did not show any changes of mRNA structure. Another possibility could be an increase in mRNA stability; but theoretically, the less frequently used codons decrease it [Kennell, 1986]. Other explanations include the influence of flanking codons on codon usage at particular sites. The so-called "codon-context effect" has been found in various species; it describes the finding that some major codons are not used preferentially at all sites [Morton and So, 2000]. The effect is presumably due to codon-codon interaction during translation [Shpaer, 1986]. A more complex explanation for our findings would be

decreased differences in the energy of codon-anticodon pairing during the elongation phase [Thomas *et al.*, 1988].

Besides these explanations, it is possible that tRNA depletion occurs during translation of over-expressed genes. I here suggest that in highly expressed genes, the tRNA with the anticodon corresponding to a commonly used codon might become depleted. As a consequence, overall less protein is synthesized, but mRNAs using less frequent codons at some positions may be translated more efficiently. So far, the tRNA drop-off phenomenon has been shown only under starvation conditions for aminoacylatable tRNA in *E. coli* and *in vitro* [Menninger, 1978], but it seems to be an attractive explanation for this study. Because all serines in the DsRed gene are encoded by the UCC codon, substituting one UCC by an UCA codon makes available an entirely unused tRNA pool. At any rate, besides the general codon optimization rules, my data suggest a more detailed consideration of tRNA kinetics, especially, when tRNA pools are represented at equal levels.

## **4.2. Specific Aim 2: Characterize the IgH major intronic enhancer and other**

### **4.2.1. *Cis*-acting elements in somatic hypermutation**

Because none of our constructs needed to contain an Ig sequence, we confirmed our previous findings that no Ig specific sequence is needed to target the mutator [Wang *et al.*, 2004], at least on exogenous substrates. The question whether (Ig and non-Ig) enhancers have a direct effect on hypermutation, or only via their effect on transcription, was not answered in a simple way. Because all our reporter constructs contain the

MoMuLV enhancer in the LTR regions, we could not rule out that enhancer elements are needed for hypermutation in a direct way. As described above, working with sin-vectors would resolve this issue. But are enhancer elements needed that are Ig-locus specific? It was suggested by Michael *et al.* that E-box motifs have an influence on mutation rates [Michael *et al.*, 2003]. E-box motifs are distributed in the IRES and LTR enhancer sequences in the reporter constructs, and thus one cannot rule out that they are directly responsible for a high mutation rate. But certainly, these motifs are not unique to the Ig enhancers. Besides their presence in the IRES and LTR enhancer, additional E-box motifs were also introduced with the E $\mu$ ,  $\mu$  intronic and  $\mu$  3' enhancer elements, but not with the SV40 enhancer. Still, the SV40 enhancer, with a lower expression rate than the  $\mu$  intronic enhancer, gave a higher mutation rate.

A comparison between the Rm(X)-IG constructs, with positive selection by FACS, and between the G(X)-IP and Rm(X)-IN constructs [Klasen *et al.*, 2004], with negative selection by drugs, showed that enhancers had an influence on the mutation rate in the case where there was no drug selection. A simple explanation for this could be that the enhancer elements increase transcription rates which, in turn, results in higher mutation rates. In drug selection, a high selection pressure results in a high transcription rate, because cells that are not resistant just die. However, although in my experiments it did not increase the transcription rate of the reporter construct (as measured by the relative GFP fluorescent intensity), the SV40 enhancer mediated a high mutation rate. Because the SV40 enhancer was placed directly 3' to the reporter DsRed gene and 5' to the IRES-GFP selection marker, it might be that the SV40 enhancer increases transcription initiation at the promoter, and at the same time inhibits the initiation of translation at the

IRES element. Indeed, interference between IRES elements and other *cis*-acting elements has been reported [Chappell *et al.*, 2000]. On the other hand enhancer elements, e. g., matrix attachment regions, may form secondary structures that lead to a pause in transcription [Fernandez *et al.*, 2001; Wiersma *et al.*, 1999]. Hypothetically, structurally framing a target DNA sequence between a promoter and an enhancer element could cause a pause in transcription. This would keep AID in close proximity to an exposed ssDNA within the transcription bubble for a longer period of time, which could result in an increased mutational load per transcription initiation [Peters and Storb, 1996; Pham *et al.*, 2003]. But very high transcription rates could compensate for the necessity of an enhancer. The different influences of the  $E\mu$  enhancer in our antibiotic resistance selected constructs compared to the GFP selected constructs may be seen as circumstantial evidence for this hypothesis. In conclusion, depending on their position in these ectopically expressed constructs and the transcription rate, enhancers can have positive or negative effects on hypermutation.

### **4.3. Specific Aim 3: Somatic hypermutation and mismatch repair in non-B cells**

#### **4.3.1. Hypermutation rate in the 293T cell line**

By regulating expression of MLH1, which is essential in mismatch repair, in the 293T kidney cell line that had been transduced by an AID containing vector no difference in mutation rates of an indicator gene was found, whether or not MLH1 was expressed.

In the 293T cell line, AID is not phosphorylated [Chaudhuri *et al.*, 2004]. In a cell free system, AID purified from ectopic expression in the 239T cell line could deaminate

single stranded DNA, but had little deamination activity on the RGYW motif [Chaudhuri *et al.*, 2004]. One would, therefore, expect a low mutation rate in this cell line. With  $4 \times 10^{-6}$  mutations per bp per day, the frequency does not seem to be so low. However, to estimate a rate one needs to take into account (i) the copy number of the GFP\* reporter gene; and (ii) the number of cell divisions per day. For my experiment, the average GFP\* copy number per cell was measured by quantitative PCR [Spillmann, personal communication] and it was found that it contained 1.6 times as many GFP\* gene copies as our standard transduced 18-81 cell line, with one copy of GFP\* per cell. I thus divide the frequency of  $4 \times 10^{-6}$  mutations per bp per day by the factor 1.6, which yields a frequency of  $2.5 \times 10^{-6}$  mutations per bp per day.

Now the question arises whether to give rates per day or per cell generation. It is generally believed that in B cells undergoing Ig somatic hypermutation, more cell divisions result in more mutations. This is because DNA synthesis and replication is thought to play a role in the known mutation mechanisms—cytidine deamination and subsequent conversion to thymidine, uracil-DNA glycosylase-mediated repair, DNA mismatch repair, and DNA synthesis by error prone polymerases. If the above number of  $2.5 \times 10^{-6}$  mutations per bp per day was converted into a rate per cell generation, one would have to divide it by the number of cell generation per day. Due to the drug selection pressure, the transduced 239T cells divided every 74 hours, i.e., with 0.32 generations per day. I thus would arrive at  $7.8 \times 10^{-6}$  mutations per bp per cell generation—still much higher than the normal, spontaneous point mutation rate. However, in recent experiments we found that the frequency of mutants does not necessarily increase with the number of cell generations [Wang *et al.*, 2004; Bachl *et al.*,

1999; Wang, personal communication]. On the contrary, a greater number of divisions can lead to a lower frequency of mutants. Our results indicate that cell division is not a rate-limiting step in the hypermutation process. We, therefore, suggested that the mutation rate is governed primarily by the rate of DNA transcription and the activity of AID. If so, one needs to compare the rate of  $2.5 \times 10^{-6}$  mutations per bp per day to the one of  $1 \times 10^{-5}$  mutations per bp per day for the ectopic GFP\* in the 18-81 cell line [Bachl and Wabl, 1996].

#### **4.3.2. Conclusion**

In this thesis, rather than on the activity of AID, we focused on the function of MMR on hypermutation in cells other than activated B lymphocytes. In hypermutating B cells, the loss of MMR proteins, such as MLH1 and PMS2, results in a reduction of the mutation frequencies and in a change of the mutation pattern [Winter *et al.*, 1998, Chaudhuri *et al.*, 2004; Phung *et al.*, 1999]. Consequently, we and others suggested an active contribution of MMR to hypermutation, for example fixing [Cascalho *et al.*, 1998; Shannon and Weigert, 1998] or introducing [Li *et al.*, 2004; Woo *et al.*, 2003] base mismatches. Physical association of MLH1 and Exo1 to the Ig V-region, but not the Ig C-region in hypermutating BL2 cells support this idea [Li *et al.*, 2004; Woo *et al.*, 2003]. However, there is no reason to expect “reverse repair” in non-B lymphocytes. Here I find no difference in mutation frequencies, whether or not MMR is functional. Because in both cases mutations do persist at a rate still much higher than the normal, spontaneous mutation rate, I conclude that in cells other than hypermutating B cells, MMR is overwhelmed and thus does not correct mutations in a noticeable way. A corollary of this

conclusion would be that hypermutating B cells contain specific factor(s) that modify mismatch repair.

## 5. Summary

Somatic hypermutation in the Ig genes is a paradigm of a site-specific, stage-specific, lineage-specific "mutator system" that generates point mutations at high rates at the Ig locus in an AID mediated and transcription dependent manner during affinity maturation of activated B-cells. The questions in the field of hypermutation today are (i) what targets the mutator factors to the Ig genes? And (ii) what are these factors?

To address some aspects of these questions, I engineered and compared retroviral vectors to monitor hypermutation. Retroviral vectors combine a high transduction rate with integration at random sites within the host cell genome, thus equalizing positional effects on the reporter gene. The vectors contain a reporter gene with a premature TAG termination codon; upon reversion, a full-length fluorescent protein is expressed. The number of fluorescent cells can be easily measured in flow cytometry, and thus mutation frequencies can be determined with accuracy. I tested the green and yellow fluorescence proteins (GFP and YFP); and various proteins with red fluorescence (dsRed). Furthermore, I tested various selection markers to select for transduced cells. I found that GFP as a reporter, combined with a drug selection marker, gave the most consistent and convenient mutation rate measurements. DsRed is a good alternative to GFP, but variants with greater fluorescence intensity are needed when combined with green fluorescence measurements.

To study the activity of enhancers on transcription and hypermutation, I introduced various *cis*-acting enhancer elements into the reporter construct. Using such constructs I found that no immunoglobulin specific sequence is needed to target hypermutation. Also,



depending on their position in these ectopically expressed constructs and on the transcription rate, enhancers can have positive or negative effects on hypermutation.

I also applied the indicator system to investigate whether or not mismatch repair influences AID mediated hypermutation in a non-B lymphocyte line. To do so MLH 1 expression, which is essential in mismatch repair, was regulated in a non-lymphocyte cell line that had been transduced by an AID containing vector. Whether or not MLH1 was expressed, we found no difference in mutation rates of an indicator gene. This is in contrast to activated B cells, where the absence of mismatch repair results in a reduction of hypermutation. I, therefore, conclude that in order to contribute to hypermutation, mismatch repair needs additional factors that are present in activated B lymphocytes but absent in the cell line investigated.

During my cloning efforts to create DsRed containing indicator vectors, I found that a single base substitution that does not change the amino acid sequence in the gene encoding DsRed2 resulted in an increase in relative fluorescence intensity of the protein. Surprisingly, the mutated codon is slightly less favored than the original one and, therefore, would be expected to have a negative effect on fluorescence intensity, if any at all. I suggest that in highly expressed genes, the tRNA with the anticodon corresponding to a commonly used codon might become depleted. As a consequence, overall less protein is synthesized, but mRNAs using less frequent codons at some positions may be translated more efficiently.

## 6. References

- Alani, E. The *Saccharomyces cerevisiae* MSH2 and MSH6 proteins form a complex that specifically binds to duplex oligonucleotides containing mismatched DNA base pairs. *Mol. Cell Biol.* 1996. **16**: 5604 - 5615
- Allen, D., Cumano, A., Dildrop, R., Kocks, C., Rajewski, K., Rajewski, N., Roes, J., Sablitzky, F., Siekevitz, M. Timing, Genetic requirements and functional consequences of somatic hypermutation during B-cell development. *Immunol. Rev.* 1987. **96**: 5 - 22
- Amin, N.S., Nguyen, M.N., Oh, S., Kolodner, R.D. exo1-dependent mutator mutations: model system for studying functional interactions in mismatch repair. *Mol. Cell Biol.* 2001. **21**(15): 5142 - 5155
- Bachl, J., Wabl, M. Enhancers of hypermutation. *Immunogenetics.* 1996. **45**: 59 - 64
- Bachl, J., Steinberg, C., Wabl, M. Critical test of hot spot motives for immunoglobulin hypermutation. *Eur. J. Immunol.* 1997. **12**: 3398 - 3403
- Bachl, J., Olsson, C., Chitkara, N., Wabl, M. The Ig mutator is dependent on the presence, position, and orientation of the large intron enhancer. *Proc. Natl. Acad. Sci.* 1998. **95**: 2396 - 2399
- Bachl, J., Dessing, M., Olsson, C., von Borstel, R.C., Steinberg, C. An experimental solution for the Luria-Delbrueck fluctuation problem in measuring hypermutation rates. *Proc. Natl. Acad. Sci.* 1999. **96**: 6847 - 6849
- Bachl, J., Olsson, C. Hypermutation targets a green fluorescent protein-encoding transgene in the presence of the immunoglobulin enhancers. *Eur. J. Immunol.* 1999. **29**: 1383 - 1389
- Bachl, J., Carlson, C., Gray-Schopfer, V., Dessing, M., Olsson, C. Increased transcription levels induce higher mutation rates in a hypermutating cell line. *J. Immunol.* 2001. **166**: 5051 - 5057
- Baird, S.G., Zacharias, D.A., Tsien, R.Y. Biochemistry, mutagenesis, and oligomerization of DsRed, a red fluorescent protein from coral. *Proc. Natl. Acad. Sci.* 2000. **97**: 11984 - 11989
- Barreto, V., Reina-San-Martin, B., Ramiro, A.R., McBride, K.M., Nussenzweig, M.C. C-terminal deletion of AID uncouples class switch recombination from somatic hypermutation and gene conversion. *Mol Cell.* 2003. **12** (2): 501 - 508
- Berek, C., Berger, A., Apel, M. Maturation of immune response in germinal centers. *Cell.* 1991. **67**: 1121 - 1129

Betz, A.G., Neuberger, M.S., Milstein, C. Discriminating intrinsic and antigen-selected mutational hotspots in immunoglobulin V genes. *Immunol. Today*. 1993. **14**: 405 - 411

Betz, A.G., Milstein, C., Gonzales-Fernandez, A., Pannell, R., Larson, T., Neuberger, M.S. Elements regulating somatic hypermutation of an immunoglobulin  $\kappa$  gene: critical role for the intronic enhancer/matrix attachment region. *Cell*. 1994. **77**: 239 - 245

Bevis, B.J., Glick, B.S. Rapidly maturing variants of the *Discosoma* red fluorescent protein (DsRed). *Nat. Biotechnol.* 2002. **20**: 83 - 87

Blum, H., Beier, H., Gross, H.J. Improved silver staining of plant proteins, RNA and DNA in polyacrylamid gels. *Electrophoresis*. 1987. **8**: 93 - 99

Boursier, L., Su, W., Spencer, J. Analysis of strand biased 'GC' hypermutation in human immunoglobulin V( $\lambda$ ) gene segments suggests that both DNA strands are targets for deamination by activation-induced cytidine deaminase. *Mol. Immunol.* 2004. **40** (17): 1273 - 1278.

Bransteitter, R., Pham, P., Scharff, M.D., Goodman, M.F. Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. *Proc. Natl. Acad. Sci.* 2003. **100**: 4102 - 4107

Breslauer, K.J., Frank, R., Blocker, H., Marky, L.A. Predicting DNA duplex stability from the base sequence. *Proc. Natl. Acad. Sci.* 1986. **83**: 3746 - 3750

Bronner, C.E., Baker, S.M., Morrison, P.T., Warren, G., Smith, L.G., Lescoe, M.K., Kane, M., Earabino, C., Lipford, J., Lindblom, A., Tannergård, P., Bollag, R.J., Godwin, A.R., Ward, D.C., Nordenskjöld, M., Fishel, R., Kolodner, R., Liskay, R.M. Mutation in the DNA mismatch repair gene homologue hMLH 1 is associated with hereditary non-polyposis colon cancer. *Nature*. 1994. **368**: 258 - 261

Bross, L., Fukita, Y., McBlane, F., Demolliere, C., Rajewsky, K., Jacobs, H. DNA double-strand breaks in immunoglobulin genes undergoing somatic hypermutation. *Immunity*. 2000. **13**: 589 - 597

Buermeyer, A.B., Deschenes, S.M., Baker, S.M., Liskay, R.M. Mammalian mismatch repair. *Annu. Rev. Genet.* 1999. **33**: 533 - 564

Bullock, W.O., Fernandez, J.M., Short, J.M. XL1-Blue: a high efficiency plasmid transforming *recA Escherichia coli* strain with beta-galactosidase selection. *BioTechniques*. 1987. **5** (4): 376 - 378

Cascalho, M., Wong, J., Steinberg, C., Wabl, M. Mismatch repair co-opted by hypermutation. *Science*. 1998. **279**: 1207 - 1210

Chappell, S.A., Edelman, G.M., Mauro, V.P. A 9-nt segment of a cellular mRNA can function as an internal ribosome entry site (IRES) and when present in linked multiple copies greatly enhances IRES activity. *Proc. Natl. Acad. Sci.* 2000. **97** (4):1536 - 1541

Chaudhuri, J., Tian, M., Khuong, C., Chua, K., Pinaud, E., Alt, F.W. Transcription-targeted DNA deamination by the AID antibody diversification enzyme. *Nature*. 2003. **422**: 726 - 730

Chaudhuri, J., Khuong, C., Alt, F.W. Replication protein A interacts with AID to promote deamination of somatic hypermutation targets. *Nature*. 2004. **430**: 992 - 998

CLONTECHniques. Living Colors<sup>®</sup> DsRed 2. BD Biosciences. 2001. XVI: 2 - 3

Colosimo, A., Goncz, K.K., Holmes, A.R., Kunzelmann, K., Novelli, G., Malone, R.W., Bennett, M.J., Gruenert, D.C. Transfer and expression of foreign genes in mammalian cells. *BioTechniques* 2000. **29** (2): 314 - 318

Danos, O., Mulligan, R.C. Safe and efficient generation of recombinant retroviruses with amphotropic and ecotropic host range. *Proc. Natl. Acad. Sci.* 1988. **85** (17): 6460 - 6464

Diaz, M., Verkoczy, L.K., Flajnik, M.F., Klinman, N.R. Decreased frequency of somatic hypermutation and impaired affinity maturation but intact germinal center formation in mice expressing antisense RNA to DNA polymerase  $\delta$ . *J. Immunol.* 2001. **167**: 327 - 335

Dickerson, S.K., Market, E., Besmer, E., F. Nina Papavasiliou, F.N. AID mediates hypermutation by deaminating single stranded DNA. *J. Exp. Med.* 2003. **197** (10): 1291 - 1296

Di Noia, J. Neuberger, M.S. Altering the pathway of immunoglobulin hypermutation by inhibiting uracil-DNA glycosylase. *Nature*. 2002. **419**: 43 - 48

Doi, T., Kinoshita, K., Ikegawa, M., Muramatsu, M., Honjo, T. *De novo* protein synthesis is required for the activation-induced cytidine deaminase function in class-switch recombination. *Proc. Natl. Acad. Sci.* 2003. **100** (5): 2634 - 2638

Ehrhardt, A., Xu, H., Kay, M.A. Episomal persistence of recombinant adenoviral vector genomes during the cell cycle in vivo. *J. Virol.* 2003. **77** (13): 7689 - 7695

Fernandez L.A., Winkler, M., Grosschedel, R. Matrix attachment region-dependent function of the immunoglobulin  $\mu$  enhancer involves histone acetylation at a distance without changes in enhancer occupancy. *Mol. Cell. Biol.* 2001. **21** (1): 196 - 208

Frey, S., Bertocci, B., Delbos, F., Quint, L., Weill, J.C., Reynaud, C.A. Mismatch repair deficiency interferes with the accumulation of mutations in chronically stimulated B cells and not with the hypermutation process. *Immunity*. 1998. **9** (1): 127 - 134.

Gearhart, P.J., Johnson, N.D., Douglas, R., Hood, L. IgG antibodies to phosphorylcholine exhibit more diversity than their IgM counterparts. *Nature*. 1981. **291**: 29 - 34

Gearhart, P.J. The effect of somatic hypermutation on antibody affinity. *Ann. N.Y. Acad. Sci.* 1983. **418**: 171 - 176

Gearhart, P.J., Bogenhagen, D.F. Clusters of point mutations are found exclusively around rearranged antibody variable genes. *Proc. Natl. Acad. Sci.* 1983. **11**: 3439 - 3443

Genschel, J., Littman, S.J., Drummond, J.T., Modrich, P. Isolation of MutS $\square$  from human cells and comparison of the mismatch repair specificities of MutS $\square$  and MutS $\square$ . *J. Biol. Chem.* 1998. **273**: 19895 - 19901

Genschel, J., Bazemore, L.R., Modrich, P. Human exonuclease 1 is required for 5' and 3' mismatch repair. *J. Biol. Chem.* 2002. **277**: 13302 - 13311

Gordon, M.S., Kanegai, C.M., Doerr, J.R., Wall, R. Somatic hypermutation of the B cell receptor genes B29 (Igbeta, CD79b) and mb1 (Igalpha, CD79a). *Proc. Natl. Acad. Sci.* 2003. **100** (7): 4126 - 4131

Green, N.S., Rabinowitz, J.L., Zhu, M., Kobrin, B.J., Scharff, M.D. Immunoglobulin variable region hypermutation in hybrids derived from a pre-B- and a myeloma cell line. *Proc. Natl. Acad. Sci.* 1995. **92**: 6304 - 6308

Grosjean H., Fiers, W. Preferential codon usage in prokaryotic genes: the optimal codon-anticodon interaction energy and selective codon usage in efficiently expressed genes. *Gene*. 1982. **18** (3): 199 - 209

Grosschedl, R., Weaver, D., Baltimore, D., Costantini, F. Introduction of a mu immunoglobulin gene into the mouse germ line: specific expression in lymphoid cells and synthesis of functional antibody. *Cell*. 1984. **38**: 647 - 658

Gurskaya, N.G., Fradkov, A.F., Terskikh, A., Matz, M.V., Labas, Y.A., Martynov, V.I., Yanushevich, Y.G., Lukyanov, K.A., Lukyanov, S.A. GFP-like chromoproteins as a source of far-red fluorescent proteins. *FEBS Letters*. 2001. **507**: 16 - 20

Haas, J., Park, E.C., Seed, B. Codon usage limitations in the expression of HIV-1 envelope glycoprotein. *Curr. Biol.* 1996. **6** (3): 315 - 324

Hanahan, D. Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* 1983. **166**: 23

Hesslein, D.G., and Schatz, D.G. Factors and forces controlling V(D)J recombination. *Adv. Immunol.* 2001. **78**: 169 - 232

Holm L. Codon usage and gene expression. *Nucleic Acids Res.* 1986. **14**: 3075 - 3087

Honjo, T. Does AID need another aid? *Nature Immunol.* 2002. **3** (9): 800 - 801

Honjo, T., Muramatsu, M., Fagarasan, S. AID: how does it aid antibody diversity? *Immunity.* 2004. **20** (6): 659 - 68

Ito, S., Nagaoka, H., Shinkura, R., Begum, N., Muramatsu, M., Nakata, M., Honjo, T. Activation-induced cytidine deaminase shuttles between nucleus and cytoplasm like apolipoprotein B mRNA editing catalytic polypeptide 1. *Proc. Natl. Acad. Sci.* 2004. **101** (7): 1975 - 1980

Jacobs, H., Fukita, Y., van der Horst, G.T., de Boer, J., Weeda, G., Essers, J., de Wind, N., Engelward, B.P., Samson, L., Verbeek, S., de Murcia, J.M., de Murcia, G., te Riele, H., Rajewsky, K. Hypermutation of immunoglobulin genes in memory B cells of DNA repair-deficient mice. *J. Exp. Med.* 1998. **187** (11): 1735 - 1743

Jiricny, J. Replication errors: cha(lle)nging the genome. *EMBO J.* 1998. **17**: 6427 - 6436

Jolly, C.J., Wagner, S.D., Rada, C., Klix, N., Milstein, C. The targeting of somatic hypermutation. *Sem. Immunol.* 1996. **8**: 159 - 168

Kennell, D.E. 1986. The instability of messenger RNA in bacteria. 101-142. In W. Reznikoff and L. Gold. *Maximizing gene expression.* Butterworth, Boston.

Kim, N., Bozek, G., Lo, J.C., Storb, U. Different mismatch repair deficiencies all have the same effects on somatic hypermutation: Intact primary mechanism accompanied by secondary modifications. *J. Exp. Med.* 1999. **190** (1): 21 - 30

Kinoshita, K., Honjo, T. Linking class-switch recombination with somatic hypermutation. *Nat. Rev. Mol. Cell. Biol.* 2001. **2**: 493 - 503

Klasen, M., Wabl, M. Silent point mutation in DsRed resulting in enhanced relative fluorescence intensity. *BioTechniques.* 2004. **36** (2): 236 - 238

Klasen, M., Spillmann, F.J., Lorens, J.B., Wabl, M. Retroviral vectors to monitor somatic hypermutation. 2004. In submission.

Klein, U., Gossens, T., Fischer, M., Kanzler, H., Braeuninger, A. Somatic hypermutation in normal and transformed human B-cells. *Immunol. Rev.* 1998. **162**: 261 - 280

Klix, N., Jolly, C.J., Davies, S.L., Bruggemann, M., Williams, G.T., Neuberger, M.S. Multiple sequences from downstream of the J kappa cluster can combine to recruit somatic hypermutation to a heterologous, upstream mutation domain. *Eur. J. Immunol.* 1998. **28** (1): 317 - 326

Kolodner, R. Biochemistry and genetics of eukaryotic mismatch repair. *Genes & Dev.* 1996. **10**: 1433 - 1442

Kong, Q., Harris, R.S., Maizels, N. Recombination-based mechanisms for somatic hypermutation. *Immunol. Rev.* 1998. **162**: 67 - 76

Kong, Q., Maizels, N. DNA breaks in hypermutating immunoglobulin genes: Evidence for a break-and-repair pathway of somatic hypermutation. *Genetics.* 2001. **1**: 369 - 378

Kotani, H., Newton, P.B. 3rd, Zhang, S., Chiang, Y.L., Otto, E., Weaver, L., Blaese, R.M., Anderson, W.F., McGarrity, G.J. Improved methods of retroviral vector transduction and production for gene therapy. *Hum. Gene Ther.* 1994. **5** (1): 19 - 28

Laemmli, U.K. Cleavage of structural proteins during assembly of the head of bacteriophage T4. *Nature (London).* 1970. **227**: 680 - 685

Lebecque, S.G., Gearhart, P.J. Boundaries of somatic mutation in rearranged immunoglobulin genes: 5' boundary is near the promoter and 3' boundary is approximately 1 kb from V(D)J gene. *J. Exp. Med.* 1990. **172**: 1717 - 1727

Lederberg, J. Genes and Antibodies. *Science.* 1959. **129**: 1649 - 1653

Li, Z., Woo, C.J., Iglesias-Ussel, M.D., Ronai, D., Scharff, M.D. The generation of antibody diversity through somatic hypermutation and class switch recombination. *Genes & Dev.* 2004. **18**: 1 - 11

Liu, M., Deora, R., Doulatov, S.R., Gingery, M., Eiserling, F.A., Preston, A., Maskell, D.J., Simons, R.W., Cotter, P.A., Parkhill, J., Miller, J.F. Reverse transcriptase-mediated tropism switching in Bordetella bacteriophage. *Science.* 2002. **295**: 2091 - 2094

Logan, A.C., Lutzko, C., Kohn, D.B. Advances in lentiviral vector design for gene-modification in hematopoietic stem cells. *Curr. Opin. Biotechnol.* 2002. **13**: 429 - 436

Lorens, J.B., Jang, Y., Rossi, A.B., Payan, D.G., Bogenberger, J.M. Optimization of regulated LTR-mediated expression. *Virology.* 2000. **272**: 7 - 15

Luria, S.E., Delbrück, M. Mutations of bacteria from virus sensitivity to virus resistance. *Genetics.* 1943. **28**: 491 - 511

Manser, T. The efficiency of antibody affinity maturation: Can the rate of B-cell division be limiting? *Immunol. Today.* 1990. **9**: 305 - 308

Marsischky, G.T., Filosi, N., Kane, M.F., Kolodner, R. Redundancy of *Saccharomyces cerevisiae* MSH3 and MSH6 in MSH2-dependent mismatch repair. *Genes & Dev.* 1998. **10**: 407 - 420

Martin, A., Bardwell, P.D., Woo, C.J., Fan, M., Shulman, M.J., Scharff, M.D. Activation-induced cytidine deaminase turns on somatic hypermutation in hybridomas. *Nature.* 2002. **415**: 802 - 806

Matz, M.V., Fradkov, A.F., Labas, Y.A., Savitsky, A.P., Zaraisky, A.G., Markelov, M.L., Lukyanov, S.A.. Fluorescent proteins from nonbioluminescent Anthozoa species. *Nature Biotech.* 1999. **17** (10): 969 - 973

Max, E.E., Maizel, J.V. Jr., Leder, P. The nucleotide sequence of a 5.5-kilobase DNA segment containing the mouse kappa immunoglobulin J and C region genes. *J. Biol. Chem.* 1981. **25** (10): 5116 - 5120

McBride, K.M., Barreto, V., Ramiro, A.R., Stavropoulos, P., Nussenzweig, M.C. Somatic hypermutation is limited by CRM1-dependent nuclear export of activation-induced deaminase. *J. Exp. Med.* 2004. **199** (9): 1235 - 1244.

McCarthy, H., Wierda, W.G., Barron, L.L., Cromwell, C.C., Wang, J., Coombes, K.R., Rangel, R., Elenitoba-Johnson, K.S., Keating, M.J., Abruzzo, L.V. High expression of activation-induced cytidine deaminase (AID) and splice variants is a distinctive feature of poor-prognosis chronic lymphocytic leukemia. *Blood.* 2003. **101** (12): 4903 - 4908

McKean, D., Huppi, K., Bell, M., Staudt, L., Gerhard, W., Weigert, M. Generation of antibody diversity in the immune response of BALB/C mice to Influenza virus hemagglutinin. *Proc. Natl. Acad. Sci.* 1984. **81**: 3180 - 3184

Menninger J.R. The Accumulation as peptidyl-transfer RNA of isoaccepting transfer RNA families in *Escherichia coli* with temperature-sensitive peptidyl transfer RNA hydrolase. *J. Biol. Chem.* 1978. **253** (19): 6808 - 6813

Michael, N., Shen, H.M., Longerich, S., Kim, N., Longacre, A., Storb, U. The E box motif CAGGTG enhances somatic hypermutation without enhancing transcription. *Immunity.* 2003. **19**: 8791 - 8794

Milstein, C., Neuberger, M.S., Staden, R. Both DNA strands of antibody genes are hypermutation targets. *Proc. Natl. Acad. Sci.* 1998. **95** (15): 8791 - 8794

Modrich, P. Mechanisms and biological effects of mismatch repair. *Annu. Rev. Genet.* 1991. **25**: 229 - 253.

Morton B.R., So, B.G. Codon usage in plastid genes is correlated with context, position within the gene, and amino acid content. *J. Mol. Evol.* 2000. **50**: 184-193

Muschen, M., Rajewsky, K., Kronke, M., Küppers, R. *CD95*-gene mutations in B-cell lymphoma. *Trends Immunol.* 2002. **23** (2): 75 - 80

Müller, H., Ziegler, B., Schweizer, B. UV/VIS-Spektroskopie in der Nukleinsäureanalytik. *BioTechnology.* 1993. **4**: 25 - 29



Muramatsu, M., Sankaranand, V. S., Anant, S., Sugai, M., Kinoshita, K., Davidson, N.O., Honjo, T. Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. *J. Biol. Chem.* 1999. **274** (26): 18470 - 18476

Muramatsu, M., Kinoshita, K., Fagarasan, S., Yamada, S., Shinkai, Y., Honjo, T. Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell.* 2000. **102**: 553 - 563

Neuberger, M.S., Milstein, C. Somatic hypermutation. *Curr. Opin. Immunol.* 1995. **7**: 248 - 254

Neuberger, M.S., Ehrenstein, M.R., Klix, N., Jolly, C.J., Yelamos, J. Monitoring and interpreting the intrinsic features of somatic hypermutation. *Immunol. Rev.* 1998. **162**: 107 - 116

Neuberger, M.S., Harris, R.S., Di Noia, J., Petersen-Mahrt, S.K. Immunity through DNA deamination. *Trends Biochem. Sci.* 2003. **28** (5): 305 - 312

Papavasiliou, F.N., Schatz, D.G. Cell-cycle-regulated DNA double-strand breaks in somatic hypermutation of immunoglobulin genes. *Nature.* 2000. **9**: 216 - 221

Pear, W.S., Nolan, G.P., Scott, M.L., Baltimore, D. Production of high-titer-helper-free retroviruses by transient transfection. *Proc. Natl. Acad. Sci.* 1993. **90** (18): 8392 - 8396

Peltomaki, P., Vasen, H.F. Mutations predisposing to hereditary nonpolyposis colorectal cancer: databases and results of a collaborative study. The international collaborative group on hereditary nonpolyposis colorectal cancer. *Gastroenterology.* 1997. **113**: 1146 - 1158

Peters, A., Storb, U. Somatic hypermutation of immunoglobulin genes is linked to transcription initiation. *Immunity.* 1996. **1**: 57 - 65

Petersen-Mahrt, S.K., Harris, R.S., Neuberger, M.S. AID mutates *E.coli* suggesting a DNA deamination mechanism for antibody diversification. *Nature.* 2002. 418: 99 - 103

Pfeifer, A., Ikawa, M., Dayn, Y., Verma, I.M. Transgenesis by lentiviral vectors: Lack of gene silencing in mammalian embryonic stem cells and preimplantation embryos. *Proc. Natl. Sci. Acad.* 2002. **99** (4): 2140 - 2145

Pham, P., Bransteitter, R., Petruska, J., Goodman, M.F. Processive AID-catalysed cytosine deamination on single-stranded DNA simulates somatic hypermutation. *Nature.* 2003. **424**: 103 - 107

Phung, Q.H., Winter, D.B., Cranston, A., Tarone, R.E., Bohr, V.A., Fishel, R., Gearhart, P.J. Increased hypermutation at G and C nucleotides in immunoglobulin variable genes from mice deficient in the MSH2 mismatch repair protein. *J. Exp. Med.* 1998. **187** (11): 1745 - 1751.

Phung, Q.H., Winter, D.B., Alrefai, R., Gearhart, P.J. Hypermutation in Ig V genes from mice deficient in the MLH1 mismatch repair protein. *J. Immunol.* 1999. **162**: 3121 - 3124

Plotz, G., Raedle, J., Brieger, A., Trojan, J., Zeuzem, S. N-terminus of hMLH1 confers interaction of hMutL $\alpha$  and hMutL $\beta$  with hMutS $\alpha$ . *Nucleic Acids Res.* 2003. **31**(12): 3217 - 3226

Poltoratsky, V., Woo, C.J., Tippin, B., Martin, A., Goodman, M.F., Scharff, D. Expression of error-prone polymerases in BL2 cells activated for Ig somatic hypermutation. *Proc. Natl. Acad. Sci.* 2001. **98**: 7976 - 7981

Quartier, P., Bustamante, J., Sanal, O., Plebani, A., Debre, M., Deville, A., Litzman, J., Levy, J., Ferman, J.P., Lane, P., Horneff, G., Aksu, G., Yalcin, I., Davies, G., Tezcan, I., Ersoy, F., Catalan, N., Imai, K., Fischer, A., Durandy, A. Clinical, immunologic and genetic analysis of 29 patients with autosomal recessive hyper-IgM syndrome due to activation-induced cytidine deaminase deficiency. *Clin. Immunol.* 2004. **110**: 22 - 29

Rada, C., Yelamos, J., Dean, W., Milstein, C. The 5' hypermutation boundary of k chains is independent of local and neighbouring sequences and related to the distance from the initiation of transcription. *Eur. J. Immunol.* 1997. **27**: 3115 - 3120

Rada, C., Ehrenstein, M.R., Neuberger, M.S., Milstein, C. Hot spot focusing of somatic hypermutation in MSH2-deficient mice suggests two stages of mutational targeting. *Immunity.* 1998. **9** (1): 135 - 141

Rada, C., Jarvis, J.M., Milstein, C. AID-GFP chimeric protein increases hypermutation of Ig genes with no evidence of nuclear localization. *Proc. Natl. Acad. Sci.* 2002a. **99** (10): 7003 - 7008

Rada, C., Williams, G.T., Nilsen, H., Barnes, D.E., Lindahl, T., Neuberger, M.S. Immunoglobulin isotype switching is inhibited and somatic hypermutation perturbed in UNG-deficient mice. *Curr. Biol.* 2002b. **12**: 1748 - 1755

Ramiro, A.R., Stavropoulos, P., Jankovic, M., Nussenzweig, M.C. Transcription enhances AID-mediated cytidine deamination by exposing single-stranded DNA on the nontemplate strand. *Nat. Immunol.* 2003. **4** (5): 452 - 456

Revy, P., Muto, T., Levy, Y., Geissmann, F., Plebani, A., Sanal, O., Catalan, N., Forveille, M., Dufourcq-Labelouse, R., Gennery, A., Tezcan, I., Ersoy, F., Kayserili, H., Ugazio, A.G., Brousse, N., Muramatsu, M., Notarangelo, L.D., Kinoshita, K., Honjo, T., Fischer, A., Durandy, A. Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the hyper-IgM syndrome (HIGM2). *Cell*. 2000. **102**: 565 - 575

Robbins, P.B., Skelton, D.C., Yu, X.J., Halene, S., Leonard, E.H., Kohn, D.B. Consistent, persistent expression from modified retroviral vectors in murine hematopoietic stem cells. *Proc. Natl. Acad. Sci.* 1998. **95** (17): 10182 - 10187

Rogerson, B., Hackett J. Jr., Peters, A., Haasch, D., Storb, U. Mutation pattern of immunoglobulin transgenes is compatible with a model of somatic hypermutation in which targeting of the mutator is linked to the direction of DNA replication. *EMBO J.* 1991. **13**: 4331 - 4341

Rogozin, I.B., Kolchanov, N.A. Somatic hypermutagenesis in immunoglobulin genes. II. Influence of neighbouring base sequences on mutagenesis. *Biochem. Biophys. Acta.* 1992. **1171**: 11 - 18

Rogozin, I.B., Sredneva, N.E., Kolchanov, N.A. Somatic hypermutagenesis in immunoglobulin genes. III. Somatic mutations in the chicken light chain locus. *Biochem. Biophys. Acta.* 1996. **1306**: 171 - 178

Rogozin, I.B., Pavlov, Y.I., Bebenek, K., Matsuda, T., Kunkel, T.A. Somatic mutations hotspots correlate with DNA polymerase  $\epsilon$  error spectrum. *Nature Immunol.* 2001. **6**: 530 - 536

Saiki, R.K., Scharf, S., Falcoona, F., Mullis, K.B., Horn, G.T., Ehrlich, H.A., Arnheim, N. Enzymatic amplification of  $\beta$ -globin genomic sequences and restriction site analysis for diagnostic of sickle cell anemia. *Science*. 1985. **230**: 1350 - 1354

Sambrook, J., Fritsch, E.F., Maniatis, T. 1989. *Molecular Cloning - A laboratory manual*. 2<sup>nd</sup> Edition. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Schmutte, C., Sadoff, M.M., Shim, K.S., Acharya, S., Fishel, R. The interaction of DNA mismatch repair proteins with human exonuclease 1. *J. Biol. Chem.* 2001. **276**: 33011 - 33018

Shannon, M., Weigert, M. Fixing Mismatches. *Science*. 1998. **279**: 1159 - 1160

Shen, H.M., Peters, A., Baron, B., Zhu, X., Storb, U. Mutation of *BCL-6* gene in normal B cells by the process of somatic hypermutation of Ig genes. *Science*. 1998. **280**: 1750 - 1752

Shen, H.M., Michael, N., Kim, N., Storb, U. The TATA binding protein, c-Myc and survivin genes are not somatically hypermutated, while *Ig* and *BCL-6* gene are hypermutated in human memory B cells. *Int. Immunol.* 2000. **12**: 1085 - 1093

Shinkura, R., Ito, S., Begum, N.A., Nagaoka, H., Muramatsu, M., Kinoshita, K., Sakakibara, Y., Hijikata, H., Honjo, T. Separate domains of AID are required for somatic hypermutation and class-switch recombination. *Nat. Immunol.* 2004. **5** (7): 707 - 712

Shpaer E.G. Constraints on codon context in *Escherichia coli* genes. Their possible role in modulating the efficiency of translation. *J. Mol. Biol.* 1986. **188**: 555 - 564

Sia, E.A., Kokoska, R.J., Dominska, M., Greenwell, P., Petes, T.D. Microsatellite instability in yeast: dependence on repeat unit size and DNA mismatch repair genes. *Mol. Cell Biol.* 1997. **17**: 2851 - 2858

Sohail, A., Klapacz, J., Samaranayake, M., Ullah, A., Bhagwat, A.S. Human activation-induced cytidine deaminase causes transcription-dependent, strand-biased C to U deaminations. *Nucleic Acids Res.* 2003. **31** (12): 2990 - 2994

Storb, U., Peters, A., Klotz, E., Kim, N., Shen, H.M. Cis-acting sequences that affect somatic hypermutation of *Ig* genes. *Immunol. Rev.* 1998. **162**: 153 - 160

Spillmann, F.J., Wabl, M. Endogenous expression of activation-induced cytidine deaminase in cell line WEHI-231. *J. Immunol.* 2004. **173** (3): 1858 - 1867

Svarovskaia, E.S., Delviks, K.A., Hwang, C.K., Pathak, V.K. Structural determinants of murine leukemia virus reverse transcriptase that affect the frequency of template switching. *J. Virol.* 2000. **74** (15): 7171 - 7178

Svoboda, J., Hejnar, J., Geryk, J., Elleder, D., Vernerova, Z. Retroviruses in foreign species and the problem of provirus silencing. *Gene.* 2000. **261** (1): 181 - 188

Swift, S.E., Lorens, J.B., Achacoso, P., Nolan, G.P. 1999. Rapid production of retroviruses for efficient gene delivery to mammalian cells using 293T cell-based systems. Vol. 10.17C, 1-17. In J. R.C. Coligan, E., Kruisbeek, M. A., Margulies, D.H., Shevach, E.M., Strober, W. *Current protocols in immunology*. Wiley, New York.

Ta, V. T., Nagaoka, H., Catalan, N., Durandy, A., Fischer, A., Imai, K., Nonoyama, S., Tashiro, J., Ikegawa, M., Ito, S., Kinoshita, K., Muramatsu, M., Honjo, T. AID mutant analysis indicate requirement for class-switch-specific cofactor. *Nature Immunol.* 2003. **4** (9): 843 - 849

Thomas, L.K., Dix, D.B., Thompson, R.C. Codon choice and gene expression: Synonymous codons differ in their ability to direct aminoacyl-transfer RNA binding to ribosomes in vitro. *Proc. Natl. Acad. Sci.* 1988. **85**: 4242 - 4246

Tishkoff, D.X., Boerger, A.L., Bertrand, P., Filosi, N., Gaida, G.M., Kane, M.F., Kolodner, R.D. Identification and characterization of *Saccharomyces cerevisiae* EXO1, a gene encoding an exonuclease that interacts with MSH2. Proc. Natl. Acad. Sci. 1997. **94**: 7487 - 7492

Tonegawa, S. Somatic generation of antibody diversity. Nature. 1983. **302**: 575 - 581

Towbin, H., Staehlin, T., Gordon, J. Electrophoretic transfer of proteins from polyacrylamid gels to nitrocellulose sheets: Procedure and some applications. Proc. Natl. Acad. Sci. 1979. **76**: 4350 - 4354

Trojan, J., Zeuzem, S., Randolph, A., Hemmerle, C., Brieger, A., Raedle, J., Plotz, G., Jiricny, J., Marra, G. Functional analysis of hMLH1 variants and HNPCC-related mutations using a human expression system. Gastroenterology. 2002. **122** (1): 211 - 219

Umar, A., Risinger, J.I., Glaab, W.E., Tindall, K.R., Barrett, J.C., Kunkel, T.A. Functional overlap in mismatch repair by human MSH3 and MSH6. Genetics. 1998. **148**: 1637 - 1646

van der Putten, H., Quint, W., van Raaij, J., Maandag, E.R., Verma, I.M., Berns, A. M-MuLV-induced leukemogenesis: integration and structure of recombinant proviruses in tumors. Cell. 1981. **24** (3): 729 - 739

von Borstel, R.C., Cain, K.T., Steinberg, C.M. Inheritance of spontaneous mutability in yeast. Genetics. 1971. **69** (1): 17 - 27

Wabl M.R., Beck-Engeser, G.B., Burrows, P.D. Allelic inclusion in the pre-B-cell line 18-81. Proc. Natl. Acad. Sci. 1984. **81** (3): 867 - 870

Wabl, M., Burrows, P.D., von Gabain, A., Steinberg, C. Hypermutation at the immunoglobulin heavy chain locus in a pre-B-cell line. Proc. Natl. Acad. Sci. 1985. **82**: 479 - 482

Wabl, M., Jaeck, H.M., Meyer, J., Beck-Engeser, G., von Borstel, R.C., Steinberg, C. Measurements of mutation rates in B lymphocytes. Immunol. Rev. 1987. **96**: 91 - 107

Wabl, M., Cascalho, M., Steinberg, C. Hypermutation in antibody affinity maturation. Curr. Opin. Immunol. 1999. **11**: 186 - 189

Wang, C., Harper R.A., Wabl, M. Genome wide hypermutation. Proc. Natl. Acad. Sci. 2004. **101** (19): 7352 - 73526

Warburg, O., Christian, W. Isolierung und Kristallisation des Gärungsfermentes Enolase. Biochem. Z. 1941/42. **314**: 544 - 546

Wei, K., Kucherlapati, R., Edelman, W. Mouse models for human DNA mismatch-repair gene defects. *Trends. Mol. Med.* 2002. **8**: 346 - 353

Weigert, M., Cesari, I.M., Yonkovich, S.J., Cohn, M. Variability in the lambda light chain sequence of mouse antibody. *Nature.* 1970. **228**: 1045 - 1047

Wiersma, E.J., Ronai, D., Berru, M., Tsui, F.W.L., Shulman, M.J. Role of the intronic elements in the endogenous immunoglobulin heavy chain locus. *J. Biol. Chem.* 1999. **274** (8): 4858 - 4862

Wilson, C.M. Staining of proteins on gels: comparison of dyes and procedures. *Meth. Enzymol.* 1983. **91**: 236 - 247

Winter, D.B., Gearhart, P.J. Dual enigma of somatic hypermutation of immunoglobulin variable genes: targeting and mechanism. *Immunol. Rev.* 1998. **162**: 86 - 96

Winter, D.B., Phung, Q.H., Umar, A., Baker, S.M., Tarone, R.E., Tanaka, K., Liskay, R.M., Kunkel, T.A., Bohr V.A., Gearhart, P.J. Altered spectra of hypermutation in antibodies from mice deficient for the DNA mismatch repair protein PMS2. *Immunol.* 1998. **95**: 6953 - 6958

Woo, C. J., Martin, A., Scharff, M. D. Induction of somatic hypermutation is associated with modifications in immunoglobulin variable region chromatin. *Immunity.* 2003. **19**: 479 - 489

Yarbrough, D., Wachter, R.M., Kallio, K., Matz, M.V., Remington, J. Refined crystal structure of DsRed, a red fluorescent protein from coral, at 2.0-Å resolution. *Proc. Natl. Acad. Sci.* 2001. **98** (2): 462 - 467

Yelamos, J., Klix, B., Goyenechea, B., Lozano, F., Chui, Y.L. Targeting of non-Ig sequences in place of the V segment by somatic hypermutation. *Nature.* 1995. **376**: 225 - 229

Yoshikawa, K., Okazaki, I.M., Eto, T., Kinoshita, K., Muramatsu, M., Nagaoka, H., Honjo, T. AID enzyme-induced hypermutation in an actively transcribed gene in fibroblasts. *Science.* 2002. **296**: 2033 - 2036

Zeng, X., Winter, D.B., Kasmer, C., Kraemer, K.H., Lehmann, A.R., Gearhart, P.J. DNA Polymerase  $\beta$  is a A-T mutator in somatic hypermutation of immunoglobulin variable genes. *Nature Immunol.* 2001. **2**: 537 - 541

## 7. Acknowledgements

This thesis could only be written with the collaboration of many dedicated and special people, and I am very grateful to them.

For the introduction into the topic, his open mind, his willingness to discuss, his incredible eye on substance, his never ending patience and help, I would like to thank Matthias Wabl, especially.

I would like to thank Elisabeth Weiss for supporting this thesis, her constructive critique, and her protective hand.

For financial and personal support, my gratitude goes to the Boehringer Ingelheim Fonds.

The Wabl lab offered me such a great time at UCSF. Here, I would like to thank Tobi, my old buddy, for always listening; Freia for her enthusiasm and drive; Ryan for just staying cool about almost everything; Thomas for his being a real colleague; Claudia for reminding me of things other than science; Desirée for always being a cheerleader, Cliff for great scientific suggestions; and my interns Barbara, Julia, Helena and Terri for having a great time and cheering me up. Further thanks go to Sarah, Cliff and Shuwei for trusting me with their FACS machines.

Especially, I would like to thank my "little one", Nilka, for her support and warm heart, as well as my parents for always being there for me.

# Curriculum vitae

## Angaben zur Person:

Name: Maik Klasen

Geburtstag: 08.03.1973

Geburtsort: Wolfsburg

Geschlecht: männlich

Status: alleinstehend

Nationalität: Deutsch

## Ausbildung

- 08.2000 - 12.2004      Doktorrand an der Ludwig-Maximilians-Universität München, in Kollaboration mit der University of California San Francisco (UCSF) (Abschluss: Dr. rer. nat.)
- 10.1995 - 02.2000      Studium der Biologie an der Universität Hannover (Abschluss: Diplom)
- 04.1992 - 12.1995      Berufsausbildung zum Rettungsassistenten (Abschluss: Staatsexamen)
- 04.1992 - 10.1993      Zivildienst RSH-Rettungsdienst in Lehre
- 1986 - 1992              Abitur am Theodor-Heuss Gymnasium in Wolfsburg

## Arbeits- und Projekterfahrung

02. 2004 - gegenwärtig      Unternehmensberater für M.G.P. Biotechnologies LLC, Long Beach, Kalifornien (President: Prof. R.A. Acey) (Freiberuflich)
- Erarbeitung von Marktanalysen und Marketingstrategien
  - Technische Beratung in der Produktentwicklung
  - Strategie-Beratung und Kosten-Nutzenkalkulation während der non-operativen Phase von MGP
- 08.2000 - 12.2004      Doktorarbeit: "*Cis* and *trans*-acting elements in somatic hypermutation" (Immunogenetik), University of California, San Francisco. In Kollaboration mit der Ludwig-Maximilians-Universität München (Betreuer: Prof. Wabl (USA) und Prof. Weiss (München))
- 02.1996 - 06.2000      Expansion des familiären Immobiliengeschäftes
- Aquisition und Aufwertung von Grund und Boden
  - Finanzierung, Planung und Ausführung von Bauvorhaben als Bauträger
  - Marketing, Vertragsverhandlungen, Verkauf, sowie Organisation von Vermietung und Verpachtung



- 02.2000 - 08.2000      Wissenschaftliche Hilfskraft am Institut für Mikrobiologie, Universität Hannover (Betreuer: Dr. Meens und Prof. Diekmann)
- Industrieprojekt in Kollaboration mit Novo Nordisk: Expression von sekretierten Proteinen in *Streptomyces*
  - Etablierung eines "16S rRNA-Sequencing" Services
  - Unterrichten von Biologiestudenten im Fach "Molekularbiologie"
- 05.1999 - 02.2000      Diplomarbeit: "Anwendung der Phage-display Technik zur Gewinnung von scFv-Antikörpern gegen verschiedene Chitinasen aus *Streptomyces olivaceoviridis* ATTC 11238". Institut für Mikrobiologie, Universität Hannover. In Kollaboration mit RWTH Aachen (Betreuer: Dr. Meens (Hannover) und Dr. Finnern (Aachen))
- 1997                      Erteilung des "Kategorie (B) – Zertifikates" der Federation of European Laboratory Animal Science Associations (FELASA)
- Befähigung zur selbstständigen Durchführung von Tierversuchen innerhalb der EU.
- 06.1997 - 07.1997      Automatisierung eines computergestützten high-through put Systems zur Detektion und Quantifizierung von "knospenden" Hefezellen. Institut für Mikrobiologie, Universität Hannover (Betreuer: Prof. Diekmann)
- 09.1997 - 10.1997      Praxis-Seminar in „Formulation, Implementation and Control of Competitive Strategy“. Institut für Wirtschaftswissenschaften, Universität Hannover
- 08.1994 - 11.1994      Sommerpraktikum als Vertriebs- und Versicherungs-Assistent. Volkswagen Leasing AG Braunschweig
- Vorbereitung von Leasing- und Kaufverträgen sowie Kundenbetreuung
  - Assistierend in der Bearbeitung von Kfz-Schadensfällen
- 04.1992 - 08.2000      Rettungsassistent, Rettungswache Grossburgwedel und andere (Teilzeit)
- Truppführer der Rettungs- und Notarztwagenbesatzung,
  - Ausbildung von Rettungssanitätern und Rettungsassistenten
  - Eigenverantwortliche Behandlung von Patienten im Rahmen der Notfallkompetenz

### **Ehrungen und Auszeichnungen**

- 2004 - gegenwärtig      Alumni des Boehringer Ingelheim Fonds
- 01.2002 - 01.2004      Stipendiat des Boehringer Ingelheim Fonds zur Förderung von Doktoranden im Ausland

### **Publikationen**

Klasen, M., Spillmann, F., and Wabl, M. Somatic hypermutation and mismatch repair in non-B cells. 2004. In submission.

Klasen, M., Spillmann, F., Lorens, J.B., and Wabl, M. Retroviral vectors to monitor somatic hypermutation. 2004. In submission.

Klasen, M. and Wabl, M. Silent point mutation in DsRed resulting in enhanced relative fluorescence intensity. 2004. *BioTechniques* 36 (2): 236 - 238

Baker, B. C., Chapman, J., Judus, M. R., Williams, C., Klasen, M., and Acey, R. A. Identification and partial cloning of a novel gene transcript from activated rat macrophages. 2003. FASEB J. 17: 17 - 18

Klasen, M. The E $\mu$  inronic enhancer in somatic hypermutation. 2002. FUTURA 17: 92 - 94

**Weitere Fähigkeiten**

Deutsch: Muttersprachler  
Englisch: fließend  
Spanisch: grundlagen

Sehr gute EDV-Anwenderkenntnisse in allen MS-Office Applikationen  
Sehr gute Kenntnisse im Gebrauch von DNA Datenbanken

**Hobbies**

Jazz-Bassist und Komponist, Fussball, Snow boarding, Literatur