XLF deficiency results in reduced N-nucleotide addition during V(D)J recombination

Running title: XLF deficiency reduces junctional diversity

Hanna IJspeert,1 Jacob Rozmus, MD,2 Klaus Schwarz,3,4 René L. Warren,5 David van Zessen,6 Robert A. Holt,5 Ingrid Pico-Knijnenburg,1 Erik Simons,1 Isabel Jerchel,7 Angela Wawer,8 Myriam Lorenz,3 Turkan Patıroğlu,9 Himmet Haluk Akar,9 Ricardo Leite,7 Nicole S. Verkaik,7 Andrew P. Stubbs,6 Dik C. van Gent,7 Jacques J.M van Dongen,1 Mirjam van der Burg,1

1 Dept. of Immunology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands
2 Child & Family Research Institute, BC Children’s Hospital, Vancouver, Canada
3 Institute for Transfusion Medicine, Ulm University, Ulm, Germany
4 Institute for Clinical Transfusion Medicine and Immunogenetics Ulm, German Red Cross Blood Service Baden-Wuerttemberg – Hessen, Ulm, Germany
5 Genome Sciences Centre, British Columbia Cancer Agency, Vancouver, BC V5Z 4S6, Canada
6 Dept. of Bioinformatics, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands
7 Dept. of Molecular Genetics, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands
8 Pediatric Hematology/Oncology, Childrens Hospital, TU Munich, Munich, Germany
9 Dept. of Pediatric Immunology, Erciyes University School of Medicine Kayseri, Turkey

Corresponding author:
Mirjam van der Burg, PhD
Erasmus MC
Dept. of Immunology
Wytemaweg 80
3015 CN Rotterdam
The Netherlands
Tel:+31-10-7043015
Fax: +31-10-7044731
E-mail: m.vanderburg@erasusmc.nl
Scientific category: Immunobiology
Key points

- XLF belongs to the NHEJ ligation complex and has a dual role in DNA double strand break repair and V(D)J recombination
- XLF is involved in N-nucleotide addition and thereby contributes to junctional diversity of the antigen receptors
Abstract

Repair of DNA double strand breaks (DSBs) by the non-homologous end joining pathway (NHEJ) is not only important for repair of spontaneous breaks, but also for breaks induced in developing lymphocytes during V(D)J recombination of their antigen receptor loci in order to create a diverse repertoire. Mutations in the NHEJ factor XLF result in extreme sensitivity for ionizing radiation (IR), microcephaly, and growth retardation comparable to mutations in LIG4 and XRCC4, which together form the NHEJ ligation complex. However, the effect on the immune system is variable (mild-to-severe immunodeficiency) and less prominent than seen in deficiencies of NHEJ factors ARTEMIS and DNA-PKcs with defects in the hairpin opening step, which is crucial and unique for V(D)J recombination. Therefore we aimed to study the role of XLF during V(D)J recombination. We obtained clinical data from nine XLF deficient patients, performed immune-phenotyping, and antigen receptor repertoire analysis of immunoglobulin (IG) and T-cell receptor (TR) rearrangements using next generation sequencing in six patients. The results were compared to XRCC4 and LIG4 deficiency. Both IG and TR rearrangements showed a significant decrease in the number of non-templated (N) nucleotides inserted by terminal deoxynucleotidyl transferase (TdT), which resulted in a decrease of 2-3 amino acids in the CDR3. Such reduction in the number of N-nucleotides has great impact on the junctional diversity and thereby on the total diversity of the IG and TR repertoire. This shows that XLF has an important role during V(D)J recombination in creating diversity of the repertoire by stimulating N-nucleotide insertion.
Introduction

DNA double strand breaks (DSBs) are toxic lesions and when not corrected severely impact the differentiation, growth and function of a cell. DSBs can arise in cells exposed to agents like ionizing radiation (IR), but are also purposely introduced during development and maturation of T and B lymphocytes. Lymphocytes express a unique antigen receptor on their membrane, namely the T-cell receptor or immunoglobulin (TR or IG). Prior to expression, the loci encoding for these receptors are rearranged through recombination of the multiple variable (V), diversity (D) and joining (J) genes. This process is called V(D)J recombination and ensures the generation of a large diversity of both TR and IG repertoires.

V(D)J recombination is initiated by recombination activating gene 1 (RAG1) and RAG2, that introduce DNA DSB near the V, D, and J genes. These RAG1 and RAG2 proteins are lymphoid specific, but the subsequent repair of the DNA DSBs occurs via the ubiquitously expressed non-homologous end joining (NHEJ) DNA repair pathway. After RAG cleavage, the DSBs are recognized by the DNA-PK complex, which consist of the DNA-PK catalytic subunit (DNA-PKcs) and the KU70/KU80 heterodimer. Subsequently, the ARTEMIS protein opens the hairpin-sealed coding ends, which are formed by the RAG proteins. Finally, the coding ends are ligated by the DNA ligase IV (LIG4)/XRCC4 complex in conjunction with Cernunnos/XLF. Recently, a new NHEJ factor called paralog of XRCC4 and XLF (PAXX) has been described. PAXX directly interacts with Ku and promotes DNA ligation, however the exact role of this protein in V(D)J recombination is not yet known. The diversity of the antigen receptors created during V(D)J recombination has been estimated to be over $10^{12}$. Part of the diversity arises from different combinations of the V, D and J genes ($2 \times 10^6$), but to obtain the needed variability in the receptors, the DNA ends are processed during V(D)J recombination. In case of asymmetric opening of the hairpins, palindromic (P) nucleotides can be formed. In addition, nucleotides can be deleted by
exonuclease activity and non-templated (N) nucleotides can be inserted by terminal
deoxynucleotidyl transferase (TdT) at the coupling sites (junctions) of the V, D, and J genes.\textsuperscript{12}

These processing events lead to the junctional diversity of IG and TR rearrangements, which
has an enormous impact on the final repertoire diversity.

Defects in V(D)J recombination block the development of T and B cells, which results in a
(severe) combined immunodeficiency.\textsuperscript{13-15} In addition, patients with defects in NHEJ have
increased sensitivity to IR, and can present with growth retardation and neurological
abnormalities.\textsuperscript{7,16,17}

XLF deficiency in human results in extreme sensitivity to IR, microcephaly, and growth
retardation, but the effect on the immune system is variable (mild-to-severe
immunodeficiency).\textsuperscript{6,7,18} Studies in XLF deficient mouse models have confirmed the severe
sensitivity to IR and chromosomal instability, but the V(D)J recombination defect seemed
relatively mild.\textsuperscript{19,20} So far, the most important action of XLF seems to be the stimulation of
LIG4 activity,\textsuperscript{21,22} probably by stabilizing the XRCC4-LIG4 complex on DNA and promotion
of DNA strand alignment.\textsuperscript{23} However, recently Lescale \textit{et al.} have shown that the RAG post-
cleavage complex (RAG-PCC) and XLF have overlapping functions, and suggested that both
the RAG-PCC and XLF ensure stabilization of DNA ends following DNA cleavage by
RAG.\textsuperscript{24} To further understand the function of XLF during V(D)J recombination and NHEJ,
the characteristics of the immunoglobulin heavy chain (IGH), IG kappa (IGK), IG lambda
(IGL), TR beta (TRB) and TR delta (TRD) gene rearrangements were studied in 6 XLF
deficient patients using next generation sequencing (NGS). We identified an additional role
for XLF in the contribution to junctional diversity during V(D)J recombination.
Methods

Cell samples and flow cytometric immunophenotyping
Peripheral blood, bone marrow and a skin biopsy were obtained with informed consent and according to the guidelines of the local Medical Ethics Committees. Flow cytometric analysis of peripheral blood was performed as previously described.25

Clonogenic survival assay
Clonogenic survival assays with primary skin fibroblasts from patient XLF1 and XLF5 was performed as described previously.26 In short, primary skin fibroblasts in exponential growth were trypsinized, and 500–2,000 cells (5,000–80,000 cells for the highest doses) were seeded into 10 cm plastic dishes (2 dishes per dose) and irradiated at room temperature at a dose of approximately 1 Gy/min. After 12–14 days, the cells were rinsed with 0.9% NaCl and stained with 0.25% methylene blue for survival assessment. Two independent survival experiments were performed. The results were compared to fibroblasts obtained from patients with defects in ARTEMIS (n=7), LIG4 (n=1), and XRCC4 (n=1).27,28

Analysis of DH-JH junctions from bone marrow and IGH, IGK, IGL and TRD junctions from peripheral blood
DNA was isolated from bone marrow mononuclear cells of peripheral blood mononuclear cells. DH-JH coding joints were amplified by PCR with family-specific DH primers and a consensus JH primer as previously described 29 followed by cloning of the PCR products into pGEM-T Easy (Promega). The IGH, IGK, IGL, and TRD junctions were amplified from post-ficoll peripheral blood (PB) mononuclear cells in a multiplex PCR using the BIOMED-2 VH1-6 FR1 and JH consensus (IGH), Vk2 and Vk3 forward and the JK1-5 reverse primers (IGK), Vl1/2 and Vl3 forward and J11/2/3 reverse primers (IGL), and the Vd2 forward and
Jd1 reverse primers (TRD). Primes were adapted for 454 sequencing by adding the forward A or reverse B adaptor, the ‘TCAG’ key and multiplex identifier (MID) adaptor. For illumina sequencing the primers were adapted by adding a MID adaptor to the forward or the reverse primer. PCR products were purified by gel extraction (Qiagen, Valencia, CA) and Agencourt AMPure XP beads (Beckman Coulter, Fullerton, CA, USA). Subsequently, the PCR concentration was measured using the Quant-it Picogreen dsDNA assay (Invitrogen, Carlsbad, CA). The purified PCR products were either sequenced on the 454 GS junior instrument (IGH and part of IGK) according the manufacturer’s recommendations, using the GS junior Titanium emPCR kit (Lib-A), sequencing kit and PicoTiterPlate kit (454 Life Sciences, Roche, Brandford, CT), or using the MiSeq sequencer (IGL, TRD and part of IGK) using low-complexity 300 base pair paired end sequencing (V3 chemistry).

The reads were uploaded to IMGT HighV-Quest. Subsequently these output files were analyzed using the IGGalaxy tool and R studio. For the analysis we only used the unique productive or unique unproductive rearrangements in which unique is defined as junctions having the same V, same J and same nucleotide sequence of the CDR3. Information about the composition of the junctional regions was extracted from the data provided by IMGT HighV-Quest. The majority of the TRD rearrangements we analyzed contained only 1 D gene. Therefore we used only the VDJ rearrangements, except for the analysis used for Figure 4 in which it is indicated that we analyzed either VDJ or VDDJ rearrangements.

**TRB analysis**

TRB rearrangements were amplified and sequenced using 5’RACE and Illumina sequencing as described previously. Data analysis was performed as describe above. The data from the two controls was published before.
**Junction analysis in Xlf-deficient mouse lymphocytes**

Li et al and Vera et al. published nucleotide sequences of DQ52-Jh, V\(\beta\)18D\(\beta\)1, and pMX-RSS12/23 junctions obtained from Xlf deficient lymphocytes.\(^{19,20}\) We reanalyzed the junctions and determined the deletions, N-nucleotides and P-nucleotides.

**TdT expression**

TdT expression was previously determined by microarray analysis in sorted B and T cell subsets.\(^{35,36}\) TdT protein expression levels were determined by flowcytometry using control BM and thymus as previously described.\(^{25,35}\)

**Statistical analysis**

Differences in numbers of N- and P-nucleotides and deletions were analyzed using the nonparametric Mann-Whitney \(U\) test (2-tailed; \(P < 0.05\) was considered significant) in the GraphPad Prism program (GraphPad Software).
Results

**XLF deficiency results in hypersensitivity for ionizing radiation but relative moderate immunodeficiency**

So far, 15 XLF deficient patients have been described in the literature. In this study we describe nine additional patients (Table 1). All patients (except P3) had microcephaly and growth retardation. The XLF deficient patients suffered from recurrent (opportunistic) infections, but in contrast to “classical” SCID patients with defects in RAG or ARTEMIS, in whom the outcome is usual fatal within the first year life unless they receive hematopoietic stem cell transplantation (SCT), the XLF deficient patients survived beyond the first year of life without SCT. In fact, the 12 patients described in literature and in this study in which immunophenotyping was performed were between 11 months and 14 years old at evaluation (Table 2). However, the numbers of lymphocytes, as well as CD4+ and CD8+ T cells were reduced or low. The B cells were strongly reduced or absent, especially in the older patients. (Table 2).

In line with previous studies, the fibroblasts from patient XLF5 and XLFP2 were highly sensitive to IR (Figure 1A). The sensitivity was comparable to LIG4 and XRCC4 deficiency. Interestingly, fibroblasts of ARTEMIS deficient patients are less IR sensitive. These data suggest that XLF is crucial for DNA DSBs in cells given the strongly increased sensitivity to IR in XLF deficiency and that absence of XLF results in reduced B and T cells.

**Antigen receptor repertoire analysis in XLF deficient patients**

To further investigate the V(D)J recombination defect in XLF deficient patients, their antigen receptor repertoire analysis was performed. Vera et al. have shown previously that both Xlf knockout mice and an XLF deficient patient had a skewing of the TRA repertoire toward the
more 3’ (proximal) V and J genes. They hypothesized that the reduced thymocyte lifespan does not allow the T cells to undergo multiple waves of VaJa rearrangements, which can be necessary for positive selection of T cells. At the IGH and TRB loci these subsequent rearrangements do not occur, so we wondered if the combinational diversity was also affected in these rearrangements. Therefore we performed antigen receptor repertoire analysis by NGS of the IGH and TRB rearrangements. Overall the median VH gene location was more downstream (proximal) in the XLF-deficient patients (Figure 1B), however this seemed mainly caused by skewing of the VH gene location in patients XLFP1 and XLFP2, since the other patients are in the same range of the healthy controls (Figure 1C). TRB rearrangements in XLF5 seem slightly skewed towards 5’ (distal) V genes (Supplemental Figure 1A). The IGH and TRB repertoires of XLF deficient patients showed a diverse pattern of V, D and J genes (Supplemental Figures 1B and 2), suggesting that except for TRA, the combinational diversity is intact in XLF deficiency.

Composition of the junctions in an XLF deficient patient and Xlf deficient mice is changed

Besides combinational diversity, the total diversity of the antigen receptors is also heavily influenced by junctional diversity. Therefore we studied DH-JH junction characteristics in XLF deficient bone marrow derived precursor B cells. Recombination of the DH gene to the JH gene is the first rearrangement occurring during IGH recombination, and therefore suitable to study the V(D)J recombination process in recombination deficient patients. Interestingly, DH-JH junctions in XLF deficient patient XLF1 had normal numbers of deletions and P-nucleotides, but strongly reduced numbers (0.1 vs 8.1 in controls) of N-nucleotides (Supplemental Figure 3A). This reduction was larger and significantly different from a LIG4 deficient patient. In contrast to XLF1, the LIG4 patient had a strong increase in the number of deletions, which explains the lower number of N-nucleotides.
In addition, we reanalyzed junctions derived from Xlf deficient mouse lymphocytes which were previously published.\textsuperscript{19,20} Although, it was previously suggested that the junctions in Xlf deficient mice were normal, both the junctions derived from thymocytes and the Igh junctions had a significant decrease of 1.8 (P-value 0.0109) and 1.3 (P-value 0.011) N-nucleotides, respectively (Supplemental Figure 3B).

Analysis of TR and IG rearrangements using next generation sequencing

To further assess the role of XLF in junctional diversity, we analyzed V(D)J junctions using NGS and evaluated three junctional region characteristics: (i) the number non-templated (N) nucleotides, which are inserted by TdT, (ii) the number of nucleotide deletions, and (iii) the number of palindromic (P) nucleotides, which arise due to asymmetric hairpin opening by ARTEMIS. For correct interpretation of the number of N-nucleotides, it is important to note that TdT mRNA and protein expression is highest during early rearrangements (IGH, TRD, TRG) (Figure 2A) and lower during late rearrangements (IGK, IGL and TRB) rearrangements. All rearrangements, except for TRB, were studied at DNA level, allowing to study both the productive rearrangements and the unproductive rearrangements. These unproductive rearrangements are out-of-frame or contain a stop codon, and consequently have not been selected. Depending on the availability of the patient material we used 1-6 patients and 2-10 age matched controls per rearrangements. We analyzed unique junctions, which we defined as a unique combination of V, J and nucleotide CDR3 sequence. The numbers of unique rearrangements are listed in Supplemental table 1.

Overall, the numbers of N-nucleotides in early rearrangements was much higher than in late rearrangements. The numbers for IGK and IGL rearrangements are also lower, because these are single-step rearrangements. In line with the DH-JH junctions and the junctions derived from Xlf deficient mice, the median number of N-nucleotides per junction in XLF deficient
patients was significantly decreased in the productive and unproductive IGH, IGK, IGL and TRD (Figure 2A and 3). Similarly to the TRB junctions in mice (Supplemental Figure 3B) no difference in the number of N-nucleotides was observed in TRB junctions in XLF deficient patients.

Since XLF works together with XRCC4 and LIG4 in the ligation complex, we wanted to know if the reduction of N-nucleotides is specific for XLF deficiency or it is related to a defect in the ligation complex. Therefore we analyzed unique productive and unproductive IGH rearrangements derived from seven XRCC4- and LIG4 deficient patients. These patients also had a significant reduction in the number of N-nucleotides (7nt vs 11nt in the controls) (Figure 2B), however, the reduction was significantly less than observed in XLF deficient patients (3.5nt). This suggests that defects in the ligation complex lead to reduced addition of N-nucleotides, but the largest reduction in the number of N-nucleotides is specific for XLF deficiency.

Besides the reduced number of N-nucleotides, the percentage of rearrangements with low number of N-nucleotides was increased (Figure 4). In TRD and IGH rearrangements the percentage of rearrangements that have low numbers of N-nucleotides (1-3 N-nucleotides) were is much higher compared to the controls. Also in the IGK and IGL rearrangements the percentages of rearrangements with 0 N-nucleotides were much higher (63.4% vs 32.6% and 57.7% vs 33.8%) in XLF deficient patients (Figure 4). Only in TRB rearrangements no difference was observed. These data imply that the junction diversity of TRγδ and IG is much lower in XLF deficient patients.

Reduced N-nucleotides result in a shorter CDR3 length

The diversity and the length of the CDR3 region are influenced by both the insertion of N-nucleotides and deletion of nucleotides during V(D)J recombination. In contrast to the low
numbers of N-nucleotides, the number of deletions and P-nucleotides were normal in all productive and unproductive IGH, IGK, IGL, TRD and TRB rearrangements (Figure 2C and Figure 3B).

The median CDR3 length in IGH rearrangements is 48nt and the median number of N-nucleotides is 11, so the N-nucleotides contribute for approximately a fourth of the total IGH CDR3 length. This holds also true for TRD were the total number of N-nucleotides can even be higher if multiple D genes are used. The median number of N-nucleotides in TRB rearrangements (6.5nt) is lower and very low in IGK (2nt) and IGL (2nt), therefore the number of N-nucleotides contribute less to the CDR3 length in these rearrangements. In line with this, we found that the median CDR3 length in IGK and TRB was not reduced and only 3nt shorter in IGL rearrangements (Figure 5A). However, the XLF deficient patients had the strongest reduction in the number of N-nucleotides in IGH and TRD, and consequently the CDR3 length was strongly decreased with approximately 9nt in IGH and 6nt in TRD (Figure 5A). This implies that the CDR3 protein sequence is 2-3 amino acids shorter. Such reduction in the number of N-nucleotides has great impact on the junctional diversity and thereby on the total diversity of the IG and TRD repertoire.

No difference in bias of nucleotide incorporation by TdT

Although TdT adds random nucleotides to single-stranded DNA, the nucleotide incorporation is biased. Several studies have shown that TdT preferentially uses dGTP and dCTP. 41,42 Therefore, we analyzed the total number of N-nucleotides per rearrangement and determined the percentage of A, T, C and G nucleotides. In line with the previous studies, we observed in all rearrangements that more than half of the N-nucleotides were either dGTP or dCTP (Figure 5B). This was not different in the XLF deficient patients, implicating that only the number of N-nucleotides is reduced.
Discussion

XLF is important for general DSBs repair and for repair of DSBs during V(D)J recombination, however the impact of XLF deficiency on both processes seems different. XLF deficiency has a severe effect on DSB repair, since XLF deficient patients are hypersensitive to IR (Figure 1A) and have a significant 50% reduction of cell survival. The IR sensitivity in XLF deficiency is similar to XRCC4 and LIG4 deficiency, which belong to the NHEJ ligation complex. Patients deficient for ARTEMIS, which is specifically involved in repair of certain DSBs and in hairpin opening during V(D)J recombination, are less sensitive for IR than XLF deficient patients (Figure 1A). In line with this, defects in the NHEJ ligation complex give a severe neurological phenotype and growth failure, which is not seen in ARTEMIS and in a DNA-PKcs deficient patient with a hairpin opening defect. On the contrary, hairpin opening defects have a severe impact on V(D)J recombination and consequently give rise to severe combined immunodeficiency, whereas defects in the NHEJ ligation complex give a milder immunodeficiency, which is even completely absent in XRCC4 deficient patients. Most of the XLF deficient patients survived the first years of life, or even up to 18 years without SCT. So, mutations in the NHEJ ligation complex result in severe neurological complications and growth failure, but a milder immunodeficiency compared to defects in NHEJ factors involved in hairpin opening. The difference in impact on DNA repair and V(D)J recombination in XLF deficiency might be explained by the two-tier synapse model that has recently be proposed by Lescale et al. showing that both the RAG-PCC and XLF ensure stabilization of DNA ends following RAG1/2 cleavage. They showed that the RAG-PCC has functional redundancy with XLF in RAG double strand break stabilization, which indicates that in the presence of RAG DNA end repair can be achieved during V(D)J recombination.
In a previous study we analyzed the precursor B-cell differentiation in bone marrow of patient XLF1\(^{43}\) and showed that XLF1 had considerable numbers of pre-B-II and immature B-cells, although the frequency of pro-B and pre-B-I cells was slightly elevated. So, in XLF deficiency the B-cell differentiation is not completely blocked at the pre-B-I stage where V(D)J recombination takes place as is the case in RAG, ARTEMIS and DNA-PKcs deficiency. This was comparable to the fairly normal precursor B-cell compartment in Xlf deficient mice.\(^{19}\) However, XLF deficient patients had reduced number of T and B cells, especially in the older XLF deficient patients the numbers of B cells were absent. Studies in mice have shown that hematopoietic stem cell dysfunction may contribute to progressive lymphocytopenia in Xlf/-/ mice,\(^{44}\) and that Xlf/-/ thymocytes have increased expression of P53-associated pro-apoptotic genes, which likely makes them more prone to apoptosis.\(^{20}\)

In this study, we have showed that XLF deficiency results in a strong reduction in the number of N-nucleotides. The largest reduction of N-nucleotides was observed in IGH and TRD rearrangements, which are early rearrangements taking place during differentiation stages with high TdT expression and where normally many N-nucleotides are added (Figure 2A). This reduction was not caused by increased removal of the N-nucleotides, since the number of deletions was normal. In XRCC4 and LIG4 deficient patients the number of N-nucleotides was also decreased in IGH rearrangements, but still the total number was significantly higher than in the XLF deficient patients. These data suggest that the ligation complex is important for N-nucleotide addition, and that XLF plays a dominant role.

During IGK, IGL and TRB (late) rearrangements TdT expression is lower, and less N-nucleotides are added per junction. However, we could still observe a reduction in the number of N-nucleotides in IGK and IGL rearrangements in the XLF deficient patients. Surprisingly, no reduction in the number of N-nucleotides was observed in TRB rearrangements in the XLF deficient patients (Figure 2B) and the Xlf deficient mice (Supplemental Figure 3B). In TRB
rearrangements the number of N-nucleotides is lower than in IGH and TRD rearrangements, probably because of lower TdT expression. Similar to IGH rearrangements, TRB rearrangements contain two junctional regions in which N-nucleotides are added. The controls had on average 2-3 N-nucleotides per junction, while TRB rearrangements of patient XLF5 even had a slightly higher number of nucleotides (4nt) per junction. N-nucleotide addition might be less dependent on XLF during TRB rearrangements. Alternatively, thymocytes undergo stringent selection in the thymus, and therefore only thymocytes expressing TRs with a correct CDR3 length might be selected during positive selection. Finally, we cannot exclude that other proteins like polλ and polμ play a (different) role during the different IG and TR rearrangements. Bertocci et al. have shown that polλ is involved in heavy chain rearrangements and polμ in light chain rearrangements.45,46

Recently, Rechavi et al. have shown that fetal derived lymphocytes also have reduced numbers of numbers of N-nucleotides in IGH and TRB rearrangements. In addition, they showed preferential use of the DH-proximal VH genes IGHV6-1 and IGHV1-2, and the JH-proximal IGHD7-27 genes in the fetal IGH rearrangements. In contrast, the use of IGHV6-1, IGHV1-2 and IGHD7-27 was normal in the XLF-deficient patients (data not shown), suggesting that it is not likely that the XLF-deficient patients have an increased frequency of the fetal derived lymphocytes. We rather think that XLF is important for the position, recruitment, or activation of TdT, however the exact mechanism needs to be unraveled. Since its discovery in 2005, several functions have been described for XLF. It stimulates LIG4 activity, 21,22 and it can form long filaments with XRCC4 that keep the DNA ends together in a ligation synapse, 24,47-50 and XLF is essential for gap-filling by polymerase (pol) λ and polμ during NHEJ. The latter function is interesting, since polλ and polμ belong to the same polX family of polymerases as TdT. Loss of N-nucleotide addition was also observed in TdT-/− and Ku80-/− mice.51-53 Purugganan et al. hypothesized that the most likely explanation for the loss
of N-nucleotides in the Ku80 deficient mice was loss of TdT recruitment. Gel mobility shift assays have also shown that Ku together with the Xrcc4-Ligase IV complex required for the recruitment for the two closely related polymerases polμ and polλ.\textsuperscript{54,55} Ku80 deficiency resulted in complete absence of N-nucleotides, while the Xlf deficient mice only have a reduced number of N-nucleotides.

In this study we have shown that XLF has a dual role in DSB repair and V(D)J recombination. Although the effect of XLF deficiency seems more moderate in V(D)J recombination, we could show that XLF has an important role during N-nucleotide addition.
Acknowledgements

The research for this manuscript was (in part) performed within the framework of the Erasmus Postgraduate School Molecular Medicine. The authors acknowledge dr. A.P. Jackson, dr. A.W. Langerak, dr. M. Berkowska and M. Wentink for technical support and advice. This work was supported by grants from the foundation “Sophia Kinderziekenhuis Fonds” (grant 589 H. IJspeert and M. van der Burg), the Dutch Organization for Scientific Research (NWO/ZonMw VIDI grant 91712323 to M. van der Burg) and the Dutch Cancer Foundation (KWF grant EMCR 2008-4045 to D.C. van Gent).

Authorship contributions

HIJ and MvB designed the research and wrote the paper. HIJ, RW, DvZ, RH, IP, ES, IJ, RL,NV performed experiments and/or analyzed data. JR, KS, AW, ML, TP, HA provided patient material, collected clinical data and critically read the manuscript. AS, DvG, and JvD contributed to essential discussion and edited the paper.

Conflict of interest disclosure

The authors declare no conflict of interest.
References


Figure legends

Figure 1. Clonogenic survival assay and VH gene usage

(A) Clonogenic survival after ionizing radiation showed hypersensitivity of the XLF deficient patients (XLF5 and XLFP2) compared to control, ARTEMIS deficient (n=7), LIG4 deficient (n=1), and XRCC4 deficient (n=1) fibroblasts. (B) Median VH gene location according to the relative position in the locus starting from the first functional VH gene (VH4-74) in base pairs (bp). (B) Box-and-whisker representation of VH gene location.

Figure 2. TdT expression levels and junction characteristics of productive IG and TR rearrangements

(A) Expression of TdT is higher during the early IG rearrangements and TR rearrangements, measured at RNA level by microarray (gene expression) and at the protein level by flow cytometry. MFI: mean fluorescence intensity, DN: double negative, ISP: immature single positive, SP: single positive. (B) Median number of N-nucleotides, (C) deletions, and (D) P-nucleotides in productive unique IGH, IGK, IGL, TRD and TRB rearrangements in healthy controls, XRCC4/LIG4- and XLF deficient patients. The numbers of productive unique rearrangements used to calculate the average are listed in Supplemental table 1. *, P≤0.05; **, P≤0.01; ***, P≤0.001 by two-tailed Mann-Whitney test.

Figure 3. Junction characteristics of unproductive IG and TR rearrangements

(A) Median number of N-nucleotides, (B) deletions, and (C) P-nucleotides in unproductive unique IGH, IGK, IGL, TRD and TRB rearrangements in healthy controls, XRCC4/LIG4- and XLF deficient patients. The numbers of productive unique rearrangements used to calculate the average are listed in Supplemental table 1. *, P≤0.05; **, P≤0.01; ***, P≤0.001 by two-tailed Mann-Whitney test.
Figure 4. Number of N-nucleotides per junction

The percentage of productive unique rearrangements with a certain number of total N-nucleotides per junction. The TRD rearrangements were separated based on the presence of 1 D gene (VDJ) or 2 D genes (VDDJ). Controls are indicated in green and the XLF deficient patients are indicated in blue.

Figure 5. Distribution of CDR3 length and nucleotide incorporation

(A). The average percentage of productive unique IGH, IGK, IGL, TRD, and TRB rearrangements with a certain nucleotide CDR3 length in XLF deficient patients (blue) and control (green). The numbers in the IGH and TRD graph indicate the difference in the median CDR3 length between the XLF deficient patients and the controls. (B). Distribution of the total number of N-nucleotides that are either an A, T, C or G nucleotide in controls and XLF deficient patients. nt, nucleotides.
### Table 1 Clinical characteristics of XLF deficient patients

<table>
<thead>
<tr>
<th>Patient</th>
<th>Mutation Allele 1</th>
<th>Allele 2</th>
<th>Microcephaly</th>
<th>Growth retardation</th>
<th>Infections</th>
<th>Autoimmunity</th>
<th>outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1 7</td>
<td>p.R57G</td>
<td>p.C123R</td>
<td>+</td>
<td>+</td>
<td>bacterial and opportunistic</td>
<td>+</td>
<td>died at 18yrs: septic shock</td>
</tr>
<tr>
<td>P3 7</td>
<td>c.177delC, c.177+3A&gt;T</td>
<td>c.177delC, c.177+3A&gt;T</td>
<td>+</td>
<td>+</td>
<td>bacterial and opportunistic</td>
<td>+</td>
<td>alive: 14 yrs</td>
</tr>
<tr>
<td>P4 7</td>
<td>c.177delC, c.177+3A&gt;T</td>
<td>c.177delC, c.177+3A&gt;T</td>
<td>+</td>
<td>+</td>
<td>bacterial and opportunistic</td>
<td>-</td>
<td>alive: 3 yrs</td>
</tr>
<tr>
<td>P5 7</td>
<td>p.R57G</td>
<td>p.R57G</td>
<td>+</td>
<td>+</td>
<td>recurrent respiratory tract</td>
<td>-</td>
<td>alive: 9 yrs old</td>
</tr>
<tr>
<td>P6 38</td>
<td>p.Y167X</td>
<td>p.Y167X</td>
<td>+</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>P7 38</td>
<td>p.R57X</td>
<td>p.R57X</td>
<td>+</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>P8 38,39</td>
<td>g.del ex2-5</td>
<td>g.del ex2-5</td>
<td>+</td>
<td>+</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>2BN 6</td>
<td>p.E5GfsX44</td>
<td>p.E5GfsX44</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>P1 18</td>
<td>p.R57G</td>
<td>p.R57G</td>
<td>+</td>
<td>+</td>
<td>recurrent respiratory tract</td>
<td>-</td>
<td>alive: 12 yrs</td>
</tr>
<tr>
<td>P2 18</td>
<td>p.D166RfsX2</td>
<td>1.9kb deletion</td>
<td>+</td>
<td>+</td>
<td>urinary tract</td>
<td>-</td>
<td>alive: 6 yrs</td>
</tr>
<tr>
<td>P3 18</td>
<td>p.D166RfsX2</td>
<td>1.9kb deletion</td>
<td>-</td>
<td>+</td>
<td>recurrent respiratory tract</td>
<td>-</td>
<td>alive: 2 yrs</td>
</tr>
<tr>
<td>P4 18</td>
<td>p.R176X</td>
<td>6.9kb deletion</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>alive: 8 yrs</td>
<td></td>
</tr>
<tr>
<td>P5 18</td>
<td>p.R178X</td>
<td>p.R178X</td>
<td>+</td>
<td>+</td>
<td>mouth lesions</td>
<td>+</td>
<td>died at 1.5yrs: septic shock</td>
</tr>
<tr>
<td>XLF1</td>
<td>p.Y167X</td>
<td>p.Y167X</td>
<td>+</td>
<td>+</td>
<td>recurrent respiratory tract</td>
<td>-</td>
<td>alive: SCT at 8yrs</td>
</tr>
<tr>
<td>XLF2</td>
<td>p.R178X</td>
<td>p.R178X</td>
<td>+</td>
<td>+</td>
<td>diarrhea</td>
<td>-</td>
<td>alive: SCT at 1.5 months</td>
</tr>
<tr>
<td>XLF3</td>
<td>p.R57X</td>
<td>p.R57X</td>
<td>+</td>
<td>+</td>
<td>urinary and respiratory tract</td>
<td>+</td>
<td>alive: SCT at 15 months</td>
</tr>
<tr>
<td>XLF4</td>
<td>p.R178X</td>
<td>p.R178X</td>
<td>+</td>
<td>+</td>
<td>respiratory tract</td>
<td>-</td>
<td>alive SCT at 10 yrs</td>
</tr>
<tr>
<td>XLF6-1</td>
<td>p.R178X</td>
<td>p.R178X</td>
<td>+</td>
<td>+</td>
<td>Recurrent respiratory fungal</td>
<td>-</td>
<td>alive 2yrs: SCT planned</td>
</tr>
<tr>
<td>XLF6-2</td>
<td>p.R178X</td>
<td>p.R178X</td>
<td>+</td>
<td>+</td>
<td>Recurrent respiratory tract</td>
<td>-</td>
<td>alive 8yrs: SCT planned</td>
</tr>
<tr>
<td>XLFP1</td>
<td>p.R178X</td>
<td>p.R178X</td>
<td>+</td>
<td>+</td>
<td>diarrhea, respiratory tract</td>
<td>-</td>
<td>alive: SCT at 9yrs</td>
</tr>
<tr>
<td>XLFP2</td>
<td>p.R178X</td>
<td>p.R178X</td>
<td>+</td>
<td>+</td>
<td>opportunistic (BCGitis) and bacterial (otitis)</td>
<td>+</td>
<td>alive: SCT at 12 months</td>
</tr>
</tbody>
</table>

ND is Not Described
Table 2 Immunological characteristics of XLF deficient patients

<table>
<thead>
<tr>
<th>Patient</th>
<th>Lymphocytes</th>
<th>T cells</th>
<th>CD4+ T cells</th>
<th>CD8+ T cells</th>
<th>B cells</th>
<th>NK cells</th>
<th>Age at evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>0.8 (1.4-3.3)</td>
<td>0.6 (1.0-2.2)</td>
<td>0.4 (0.5-1.3)</td>
<td>0.2 (0.3-0.9)</td>
<td>0 (0.1-0.6)</td>
<td>n.d.</td>
<td>14 years</td>
</tr>
<tr>
<td>P2</td>
<td>1.2 (2.3-5.4)</td>
<td>0.7 (1.4-3.7)</td>
<td>0.2 (0.7-1.5)</td>
<td>0.6 (0.4-1.1)</td>
<td>0.08 (0.4-1.4)</td>
<td>0.3 (0.09-0.9)</td>
<td>2 years</td>
</tr>
<tr>
<td>P3</td>
<td>1.6 (1.4-3.3)</td>
<td>0.9 (1.0-2.2)</td>
<td>0.5 (0.5-1.3)</td>
<td>0.4 (0.3-0.9)</td>
<td>0 (0.1-0.6)</td>
<td>0.4 (0.04-0.7)</td>
<td>13 years</td>
</tr>
<tr>
<td>P4</td>
<td>2.6 (2.3-5.4)</td>
<td>0.6 (1.4-3.7)</td>
<td>0.3 (0.7-1.5)</td>
<td>0.08 (0.4-1.1)</td>
<td>0.2 (0.4-1.4)</td>
<td>1.1 (0.09-0.9)</td>
<td>2 years</td>
</tr>
<tr>
<td>P5</td>
<td>1.1 (1.9-3.7)</td>
<td>0.7 (1.2-2.6)</td>
<td>0.3 (0.5-1.3)</td>
<td>0.3 (0.3-0.9)</td>
<td>0.04 (0.1-0.6)</td>
<td>0.2 (0.08-0.7)</td>
<td>7 years</td>
</tr>
<tr>
<td>P6</td>
<td>0.6 (2.3-5.4)</td>
<td>1.3 (1.9-3.7)</td>
<td>0.08 (0.4-1.4)</td>
<td>0.04 (0.1-0.6)</td>
<td>0 (0.1-0.6)</td>
<td>6 years</td>
<td></td>
</tr>
<tr>
<td>P7</td>
<td>0.2 (1.4-3.3)</td>
<td>1.3 (1.9-3.7)</td>
<td>0 (0.1-0.6)</td>
<td>0 (0.1-0.6)</td>
<td>0 (0.1-0.6)</td>
<td>6 years</td>
<td></td>
</tr>
<tr>
<td>XLF1</td>
<td>1.4 (1.7-6.9)</td>
<td>1.1 (0.9-4.5)</td>
<td>0.6 (0.5-2.4)</td>
<td>0.4 (0.3-1.6)</td>
<td>0.06 (0.2-2.1)</td>
<td>0.2 (0.1-1.0)</td>
<td>4 years</td>
</tr>
<tr>
<td>XLF5</td>
<td>0.7 (1.1-5.9)</td>
<td>0.5 (0.7-4.2)</td>
<td>0.3 (0.3-2.0)</td>
<td>0.2 (0.3-1.8)</td>
<td>0 (0.2-1.6)</td>
<td>8 years</td>
<td></td>
</tr>
<tr>
<td>XLF6-1</td>
<td>1.5 (2.6-10.4)</td>
<td>0.2 (1.6-6.7)</td>
<td>0 (1.0-4.6)</td>
<td>0.1 (0.4-2.1)</td>
<td>0.04 (0.6-2.7)</td>
<td>0 (0.2-1.2)</td>
<td>11 months</td>
</tr>
<tr>
<td>XLF6-2</td>
<td>0.9 (1.1-5.9)</td>
<td>0.9 (1.7-4.2)</td>
<td>0.3 (0.3-2.0)</td>
<td>0.4 (0.3-1.8)</td>
<td>0.01 (0.2-1.6)</td>
<td>0.1 (0.09-0.9)</td>
<td>8 years</td>
</tr>
</tbody>
</table>

Absolute counts (x10^9/L). The age-matched controls values are shown in italics in parentheses. Reduced values are indicated in bold.
Figures

Figure 1
Figure 2
Figure 3

A

B

C

Figure 3
Figure 4
Figure 5

A

IGH

IGK

IGL

CDR3 length (nt)

% of rearrangements

9nt

3nt

Control

XLF

TRD

TRB

CDR3 length (nt)

% of rearrangements

6nt

B

IGH

IGK

IGL

TRD

TRB

% of N-nucleotides

Control XLF

Control XLF

Control XLF

Control XLF

Control XLF
XLF deficiency results in reduced N-nucleotide addition during V(D)J recombination

Hanna IJspeert, Jacob Rozmus, Klaus Schwarz, René L. Warren, David van Zessen, Robert A. Holt, Ingrid Pico-Knijnenburg, Erik Simons, Isabel Jerchel, Angela Wawer, Myriam Lorenz, Turkan Patiroglu, Himmet Haluk Akar, Ricardo Leite, Nicole S. Verkaik, Andrew P. Stubbs, Dik C. van Gent, Jacques J.M. van Dongen and Mirjam van der Burg